Measuring the Economically Active in Population Censuses: A Handbook

United Nations

International Labor Office

United Nations
Department of Economic and Social Affairs

International Labour Office
Department of Statistics

Studies in Methods   Series F, No. 102

Measuring the Economically Active in Population Censuses: A Handbook

United Nations   International
                        Labour Office

New York, 2010

Note

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of either the Secretariat of the United Nations or of the International Labour Office concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Preface

The United Nations, as part of its world programme on population and housing censuses, has provided technical guidelines to its Member States primarily through the preparation of (*a*) Principles and Recommendations for Population and Housing Censuses;[1] (*b*) handbooks[2] on general census operations and on the collection of data on specific topics, such as economic activity; and (*c*) supporting technical materials.

The first version of *Principles and Recommendations for Population and Housing Censuses* (United Nations, 1980) was issued for the 1980 round of censuses (1975-1984). It has been supplemented and updated by further issues (United Nations, 1990b; 1998; and 2008b) for the 1990, 2000 and 2010 rounds of censuses that take into account changes in relevant international standard concepts and classifications as well as accumulated experiences from national census exercises. For the 1990 round of censuses, a handbook on economic activity status was also issued, presenting the content of census questionnaires and national practices in the collection of information on economic characteristics of the population during the 1970 and 1980 round of censuses.[3] The review documented which countries had included specific items in their population censuses. In 2002, a report entitled "Collection of economic characteristics in population censuses: technical report" was issued in draft form and was circulated for comments and suggestions for improvement with a view to the preparation of a handbook.

Given the broad focus of *Principles and Recommendations*, it was not possible or practical to include in that publication details that would guide countries operationally in the implementation of those concepts and standards in censuses. The present volume, *Measuring the Economically Active in Population Censuses: A Handbook*, is based on the above-mentioned technical report and provides suggestions on the implementation of guidelines on the measurement of economic characteristics in population censuses, based on relevant experiences of countries, with a particular focus on the questions used and the requirements for processing of responses. A separate part of the *Handbook* considers the role that the census results on economic characteristics can play in the planning of statistical surveys. The *Handbook* is intended to provide census planners with a variety of approaches to assess the questions and methods of collecting economic characteristics used in their national census, as they evaluate the performance in the past decade and plan for the current decade round of censuses (2005-2014). Users of census results may also find the present text useful when evaluating the quality of census results.

The publication of the *Handbook* is the result of a collaboration between the Department of Statistics of the International Labour Office and the United Nations Statistics Division. The original technical report on which it is based was first drafted

[1] *Principles and Recommendations for Population and Housing Censuses* (United Nations publication, Sales No. E.80.XVII.8); *Revision 1* (Sales No. E.98.XVII.8); *Revision 2* (Sales No. E.07.XVII.8). *Supplementary Principles and Recommendations for Population and Housing Censuses* (United Nations publication, Sales No. E.90.XVII.9).

[2] Examples include *Handbook of Population and Housing Censuses—Part I: Planning, Organization and Administration of Population and Housing Censuses* (United Nations publication, Sales No. E.92.XVII.8); *Part II: Demographic and Social Characteristics* (United Nations publication, Sales No. E.91.XVII.9); and *Part IV: Economic Activity Status* (United Nations publication, Sales No. E.96.XVII.13). See also *Handbook on Census Management for Population and Housing Censuses* (United Nations publication, Sales No. E.00.XVII.15 Rev.1).

[3] *Handbook of Population and Housing Censuses— Part IV: Economic Activity Status* (United Nations publication, Sales No. E.96.XVII.13). In 1990, 1996 and 2003 the International Labour Office published more detailed descriptions in its publication *Sources and Methods: Labour*

under the supervision of the Office,[4] reviewed by a group of experts convened by the Division and the Office and finalized in 2002 under the supervision of the Office and the Division. That report benefited from inputs by several writers. Special mention is made of Reginald Gilbert, who contributed to what are now chapters IV to VIII of the *Handbook* and prepared most of the suggestions for the questions to be tested. An initial draft of the *Handbook* was prepared by the International Labour Office and finalized by a consultant. That draft was then discussed at a meeting of experts organized by the United Nations Statistics Division and the International Labour Office in November 2007. On the basis of comments by participants in the meeting, the Division and the Office finalized the *Handbook*.

The United Nations Statistics Division and the International Labour Office invite feedback from the users of the present *Handbook* on the following: national experiences besides those presented here; the relevance and practicality of the suggested (untested) questions; how well the approaches presented have worked; and the circumstances under which they were found to be most effective. Such feedback will be useful in preparations for subsequent census rounds.

Comments or suggestions may be sent to:

Director	*Director*
Statistics Division	Department of Statistics
Department of Economic	International Labour Office
and Social Affairs	CH-1211 Geneva 22
United Nations	Switzerland
New York, NY 10017	E-mail: stat@ilo.org
United States of America	Fax: + 41 22 799 6957
Attn: Demographic	
and Social Statistics Branch	
E-mail: socialstat@un.org	
Fax: + 1-212-963-1940	

Contents

Part Three
Measurement of characteristics of jobs, establishments and persons

Part Six
Using population censuses to improve labour force and related statistics

Part One
General issues

Chapter I
Focus and content—
rationale and structure

A. Overview

1. For decades, countries have been collecting, through population censuses, data on economic characteristics of the population. For many countries the population census remains the only source for nationally representative statistics on economic characteristics, while for others the population census complements other sources (if available) such as household and labour force surveys, establishment surveys, and administrative registers. As the census is conducted only once every 10 years (in a few countries once in every 5 years), economic characteristics are often competing with demands for coverage of a wide range of other topics.

2. The decisions on content and topics to be covered are generally guided not only by the availability of resources (which determines the size of the questionnaire and number of questions that can be accommodated), but also by the availability of other sources of similar statistics and their reliability and scope, as well as by the potential usefulness of the statistics that the population census would generate. This is the context in which the question of whether to include economic characteristics in a census is considered.

3. The population census potentially covers the entire population. Therefore, in the absence of a comprehensive population and administrative register, it is the only source that can provide information on each individual, and thus finely detailed statistics on occupations and the like for the country as a whole as well as for small geographical areas. However, some of the basic concepts relating to economic characteristics are complex and will, in principle, require a depth in interviewing and probing that is not easily achieved in the population census, except where such considerations are taken into account in the planning and execution of the census operation.

4. For more complete coverage of economic characteristics, enquiries designed for in-depth interviews, based on a sample of the population, are generally more suitable for the collection of economic characteristics than a population census. Such enquiries may be attached to the population census or may be household-based sample surveys such as labour force surveys. The labour force survey can provide more extensive coverage of the topic because of its specific focus, but will not be able to provide statistics with satisfactory precision for many of the groups that may be of interest for policy formulation and implementation purposes or for analysis of the economic and social structure of the country and its regions.

5. Consequently, the population census (or for some countries, population and administrative registers) is an indispensable source for obtaining broad-based estimates on selected economic characteristics (such as the employed and unemployed, and some characteristics of the job or of the employing establishment). It also provides benchmark statistics for labour policies and programmes, such as labour force participation and unemployment rates, and occupational and industrial distributions, particularly at the regional, district

and other levels of administration. Population censuses are also useful for evaluating and producing national estimates from sample-based enquiries of those characteristics. It is therefore important to produce accurate and reliable estimates through population censuses and provide better results, not only for the countries for which the census is the only source of data on economic characteristics but also for those that carry out labour force surveys. An important underlying issue is how population census design and procedures may be chosen so as to maximize its usefulness for the above purposes.

6.   Despite their many benefits, population censuses have some disadvantages, including the high cost to the collection agency, the heavy reporting burden on the community and a certain vagueness in the measurement of topics that would require many questions to correctly classify particular groups of the population. One must also reconcile population census results with those from other data collections. Such reconciliation is facilitated if comparable concepts, definitions, units and classifications are used.

7.   *Measuring the Economically Active in Population Censuses: A Handbook* complements other resource materials that are already available to help measure economic characteristics, namely the following:

  (*a*)   International standards on concepts and definitions:

   (i)   Resolution concerning statistics of the economically active population, employment, unemployment and underemployment (see ILO, 2000);

   (ii)   Definition of System of National Accounts production boundary and related concepts, such as the institutional sector, in *System of National Accounts, 1993* (Commission of the European Communities and others, 1993);

   (iii)   Resolution concerning statistics of employment in the informal sector (see ILO, 2000);

   (iv)   Guidelines concerning treatment in employment and unemployment statistics of persons on extended absences from work (see ILO, 2000);

   (v)   Guidelines concerning a statistical definition of informal employment;[1]

  (*b*)   International standard classifications of industry, occupations and status in employment, as follows:

   (i)   International Standard Classification of Occupations (ISCO);[2]

   (ii)   International Standard Industrial Classification of All Economic Activities (ISIC);[3]

   (iii)   International Classification of Status in Employment (ICSE-93) (ILO, 1993a);

   (iv)   In general, it is preferable to adapt the international classifications for national or regional use rather than to use them directly in national statistical collections;

  (*c*)   Guidelines relating to methods on economic characteristics:

   (i)   Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods (Hussmanns, Mehran and Verma, 1990);

   (ii)   *Principles and Recommendations for Population and Housing Censuses, Revision 2* (United Nations, 2008b).

8.   The *Handbook* focuses specifically on population censuses. It explores the strengths and weaknesses of population censuses as a source of data on economic characteristics and emphasizes the uses of those data and their role in the national statistical system.

1 *Seventeenth International Conference of Labour Statisticians, 24 November–3 December 2003: Report of the Conference* (Geneva, ILO, 2004), part one, chap. 3, annex.

2 At the time of the release of the present *Handbook*, the most recent version was ISCO-08 (forthcoming). Available from http://www.ilo.org/public/english/bureau/stat/isco/index.htm. For the previous version (ISCO-88), see International Labour Office (1990a).

3 The most recent version is the *International Standard Industrial Classification of All Economic Activities, Revision 4*, endorsed by the United Nations Statistical Commission at its thirty-seventh session in 2006 (United Nations, 2008a). Available from http://unstats.un.org/unsd/cr/registry/isic-4.asp. See also United Nations (2002) for the previous version.

It reviews general census operations, highlighting aspects that have direct relevance to the collection of data on economic characteristics, and how those operations could take greater account of the requirements of such data. It also reviews some of the critical considerations for including items on economic characteristics in the census, such as issues related to the ordering of items and formulating questions. The *Handbook* explains how to operationalize the conceptual underpinnings of the items on economic characteristics that are presented in *Principles and Recommendations, Revision 2* and highlights, with illustrative examples from national census questionnaires, aspects of question formulation that could provide more accurate representation of the relevant concepts. In some cases, the *Handbook* proposes samples of questions (untested) that might be considered for trial and testing. Procedures for coding industry and occupations and the necessary tools to accomplish those tasks are also presented. The *Handbook* also deals with the comparability and complementarities of population censuses and household-based sample surveys that measure economic characteristics and reviews specific ways in which sample estimates could be enhanced with population census data.

B. Purpose of *Measuring the Economically Active in Population Censuses: A Handbook*

9.   The primary aim of the *Handbook* is to draw on the experiences of countries in measuring economic characteristics and provide guidance on the implementation of the recommendations on the economic characteristics outlined in *Principles and Recommendations for Population and Housing Censuses, Revision 2*. Revision 2 updates the 1998 edition of the same publication, which the United Nations undertook to take into account changes and new developments in related international standards on topics that are commonly included in national population and housing censuses.

10.   The main changes specific to economic characteristics that have taken place since the publication of the 1980 edition of *Principles and Recommendations*, and which are reflected in the 1998 edition as well as in the second revision, are as follows: (*a*) the revised definition of the concept of the economically active;[4] (*b*) a revised delineation of the boundary of production in the *System of National Accounts, 1993*; (*c*) revised versions of the standard international classifications of industry (ISIC Revision 3), occupation (ISCO-88) and status in employment (ICSE-1993); and (*d*) the adoption in 1993 of a framework for defining the informal sector by the International Conference of Labour Statisticians.[5] The first two revisions (that is, of the concept of economic activity and of the SNA production boundary) have profound implications for the definition as well as for the measurement of economic activity, and necessitate a review of questions and methods used by many national census offices to collect data on economic characteristics.

11.   More than 95 per cent of the countries[6] have included items on economic characteristics (in particular, employment and unemployment, status in employment, industry, occupation and, less frequently, place of work) in their recent census taking, as was the case in the past (see United Nations, 1996; ILO, 1996a). It is expected that while many users of census results will want statistics on employment in the informal sector and on informal employment for analysis of local labour markets situations, only a few census operations will consider including those topics, as the identification of those employed in such activities is highly complex and will require more probing than is possible in many census operations. However, by outlining how it could be done in a census, the *Handbook* provides countries with the option of including the informal sector and informal employment and also gives guidance on the extent to which other characteristics included in the census may be used to derive estimates of employment in the informal sector and informal employment.

[4] "Resolution concerning statistics of the economically active population, employment, unemployment and underemployment", in *Current International Recommendations on Labour Statistics, Second Edition* (Geneva, International Labour Office, 2000).

[5] "Resolution concerning statistics of employment in the informal sector", in *Current International Recommendations on Labour Statistics, Second Edition* (Geneva, International Labour Office, 2000).

[6] Based on a review undertaken by the United Nations Statistics Division in 2007 of 176 national census questionnaires available from the 2000 round of population censuses.

12.   As a supporting document to *Principles and Recommendations, Revision 2*, the *Handbook* provides more detailed and technical operational guidance concerning recommended economic characteristics than is contained in the former, in the following three respects: (*a*) elaborating on the internationally recommended concepts in light of the practical possibilities; (*b*) examining illustrative examples of questions used in national population census and their limitations; and (*c*) suggesting suitable procedures for collecting and processing the relevant information, drawing on national and international experiences.

13.   The *Handbook* does not discuss the theoretical justifications for the underlying concepts and their definitions but concentrates instead on the practical issues of identifying and describing the economically active population in a population census. It aims to translate the current SNA-93 boundary into questions that can be used in population census enumerations in which the economic characteristics of the population are being collected. While the *Handbook* focuses primarily on techniques and methods to be used in population censuses, the suggestions made are also applicable to household sample surveys that include economic characteristics as background variables. They are also applicable to specialized labour force surveys and when sampling is employed in censuses to investigate economic characteristics. There is, however, wider latitude in terms of the type and number of questions that can be included in a sample survey that focuses on the measurement of employment characteristics. Other manuals on the topic of surveys[7] need therefore to be consulted for additional and more specific guidance when making preparations for such specialized surveys.

14.   In addition to *Principles and Recommendations for Population and Housing Censuses*,[8] the International Labour Office publication entitled *Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods* (Hussmanns, Mehran and Verma, 1990) explores extensively the concepts as well as the conceptual basis for many of the decisions related to the parameters that are needed to make the concepts operational. The 1990 ILO manual was written before the SNA-93 was finalized and uses a concept of economic activity that was updated in the SNA-93. An article published in the ILO *Bulletin of Labour Statistics* (Hussmanns, 2007) updates the definition of economic activity to be consistent with the SNA-93. In addition to *Principles and Recommendations* and the ILO *Manual*, several other United Nations and International Labour Office publications provide some relevant information and guidance on some of the topics and issues addressed in the *Handbook*. They include an ILO working paper entitled "What kind of work do you do? Data collection and processing strategies when measuring 'occupation' for statistical surveys and administrative records" (Hoffmann and others, 1995); *Handbook of Population and Housing Censuses—Part I: Planning, Organization and Administration of Population and Housing Censuses* (United Nations, 1992); and *Handbook of Population and Housing Censuses—Part IV: Economic Activity Status* (United Nations, 1996).

C.   Users and uses of the *Handbook*

15.   The *Handbook* is addressed mainly to census planners; to managers responsible for designing and carrying out the census; and to other producers of statistics, including survey experts, data processing staff and trainers. For those producers, the *Handbook* is intended as a basic reference for (*a*) consulting basic documentation on current standards and practices in the measurement of the economically active population using population censuses; (*b*) identifying potential problems in measuring specific characteristics and in the formulation of certain types of questions; (*c*) considering a range of possible questions or improving upon previously used questions; and (*d*) deciding on the extent

7 See, in particular, Hussmanns, Mehran and Verma (1990).

8 The most recent version was adopted by the United Nations Statistical Commission at its thirty-eighth session in 2007 and published in 2008 as *Principles and Recommendations for Population and Housing Censuses, Revision 2* (United Nations publication, Sales No. E.07.XVII.8). For ease of reference, throughout the present handbook it will be cited simply as *Principles and Recommendations, Revision 2.*

to which economic characteristics are to be covered in the population census. It is a guide not only for the application of international standards for data collection but also for coding and adapting the standards to national classification requirements. Many of the issues raised are also important considerations in the training of field staff and in developing manuals for enumerators and their supervisors.

16. The *Handbook* is also addressed to users of census statistics on economic characteristics. The main users of those statistics have generally been various national government ministries, committees and agencies concerned with the formulation, implementation and evaluation of economic, social and labour policies. Provincial and local government authorities with similar responsibilities and concerns are also a core set of users. Development programme officers also have need of such data, for example, when planning and implementing—at the national, regional and local levels—projects for the building of roads and for water supply, electricity and telephones as well as for waste disposal; when deciding on the location of special zones for industrial development; and so on. There are in addition, at least potentially, a wide range of different users, including the following:

(*a*) Those who are interested in understanding the economic and social structure of a country as well as of its small and large regions and population groups;

(*b*) Employers' organizations and managers, planners and business leaders in private sector enterprises who need to understand their markets and the potential for recruitment of their most important resource: people to work in their establishments;

(*c*) Members of the research community, at universities and in economic, labour and social research institutes; political leaders and other policymakers; social activists and representatives of special interest groups; journalists; and international agencies;

(*d*) Those concerned with the welfare of workers and their households, such as labour unions, and those providing financial, marketing and social services to either management or labour.

17. Increasingly, advocates of gender issues and equal opportunity questions are becoming more involved as a growing user group with very specific data availability and quality concerns, particularly in relation to the different work status of women and men, as of girls and boys.

18. Users of statistics need a broad knowledge of the issues of data collection in order to engage in and maintain a fruitful dialogue with producers at all stages of the census operation, and with them to arrive at a mutual understanding of the kinds of statistics for which the population census is well suited, as well as of the limitations of such statistics. The *Handbook* will acquaint users of those statistics with the concepts, definitions and methods as well as the constraints that producers of statistics face while responding to user needs. With that understanding, census planners and users of census results may enter into a mutually beneficial dialogue about the content of and procedures for the census in the planning stages. Equally important, informed users can meaningfully communicate to the census authorities information about the form that tabulations and other census outputs should take to best meet the needs of users and provide feedback on the adequacy of the data relative to their data requirements.

D. Organization of the *Handbook*

19. The *Handbook* is divided into six main parts. Part one covers the planning and design of population censuses (see chapter II) in addition to the points covered above.

20. Part two reviews the measurement frameworks and general issues related to the inclusion of economic characteristics in the population and housing census, presents conceptual definitions associated with activity status and reviews some of the constraints on

measuring economic characteristics from population censuses. It also addresses in more detail practical considerations in collecting information for characteristics relating to economic activity that are included in *Principles and Recommendations, Revision 2*. Part two illustrates those details with examples of questions used in national censuses, indicating some aspects of the questions that might be modified to enhance their effectiveness. This part also emphasizes the role of training and presents relevant issues to be investigated.

21. Part three continues the more detailed presentation of definitions, examples and suggestions for enhancing effectiveness but focuses on the measurement of the characteristics of jobs, establishments and persons.

22. Part four deals with data processing, analysis and tabulation. It describes the procedures for checking for errors and assuring quality control. The observations on tabulation outline those tables that are considered essential separately from those that are recommended and those that are only additional.

23. Part five focuses on two main characteristics for which there are international standard classifications, namely occupation and industry. It covers the strategies that the census planners could adopt for the inclusion of those variables and their classifications, as well as the implications for enumerators or office coders in coding responses to the questions on the census questionnaire. It describes the tools to be used for coding operations, such as coding indices and instructions, and explains how to develop and test those tools.

24. Part six covers issues of comparability between census and labour force and/or household survey results. It reviews some specific uses of population census results, such as in preparing a sampling frame for labour force surveys, in evaluating labour force data and for small area and small group estimation. It reviews sources of differences between labour force data collected through the population census and the labour force survey, with respect to the scope, coverage, timing, survey instruments, extent of application of the concepts and definitions, methods, sampling and non-sampling errors and the like, and proposes ways to maximize the effective use of census results in improving sample designs and estimates from labour force surveys.

Chapter II

Planning and design of population censuses for the collection of data on economic characteristics

Uses and sources of statistics on economic characteristics

A. Uses of statistics on economic characteristics

25.   Population censuses can generate statistics on economic activity that meet the needs of many users (see para. 5). Social and demographic information on the population engaged in the production of goods and services is vital to the analysis of the economic performance of a country as well as of the regions of which it is composed. One of the main strengths of the population census as a source of data is the provision of locality-specific statistics. As a first step to including questions on economic characteristics in a census questionnaire, census authorities and potential users of census results need to work together to develop lists of the actual and anticipated uses of census statistics in their own country. Such lists will be helpful in planning census content and types of outputs, and in mobilizing support for the census.

26.   Statistics on the scope and structure of the economic activities of the population provide a useful basis for the following:

(*a*)   Formulation and planning of economic and social policies and campaigns: for example, planning for employment promotion in the public and private sectors, developing special remedial programmes to reduce unemployment, underemployment and other decent work deficits, and making decisions about the location of new economic activities and markets;

(*b*)   Administration and management of programmes, such as unemployment benefits and welfare schemes;

(*c*)   Setting of quantitative targets and thresholds to be achieved in the implementation of or as a result of policies, study of labour market behaviour and local labour market situations as a basis for making investment, zoning and community development decisions, and identification of needs in education and technical and vocational training, as well as understanding of linkages with poverty alleviation;

(*d*)   Evaluation of national or local employment or social welfare programmes and policies, including assessment of the current state of the economy and providing indications of differences in economic opportunities between different population sub-groups;

(*e*)   Conduct of scientific research and hypothesis testing, such as construction of working life tables and related estimations used to design pension schemes and other social insurance programmes;

(*f*)   Analysis of labour-related migration;

(*g*)   Poverty mapping;

(*h*)   Teaching about and general descriptions of the economic and social situation, and analysis of the labour market situation for major population groups or sub-groups in different localities, regions or the whole country.[1]

27.   In addition to the examples of the direct use of census results, the census is also used as a basis for preparing more intensive sample surveys to study special issues. The census can provide the basis for the following:

(*a*)   Developing national statistical capability, particularly in sample surveys, and providing a convenient and efficient basis for launching ongoing survey programmes;

(*b*)   Preparing population projections and benchmarks for survey results;

(*c*)   Constructing the sampling frame, and designing and selecting the sample, not only for population surveys but also for other types of surveys, such as agricultural surveys of farming households and holdings, and economic surveys of establishments and of economic activities in the informal sector;

(*d*)   Stratifying the sample for a survey, for example, to ensure that the sample will include sufficient number of relatively rare circumstances, which may be of interest to planners;

(*e*)   Evaluating sample survey results;

(*f*)   Producing small-area estimates.

A more detailed explanation of the subject is presented in part six.

28.   In spite of their potential, the usefulness to particular users of census statistics on economic characteristics may be limited by such factors as the timeliness of the census results; the form in which they are disseminated to the users; and the number of topics on which data are available. In addition, users' need for specialized cross-tabulations (with multiple variables) may not be fully taken into account, or the statistics may not be made available in formats that lend themselves to further analysis. This means that it is important to consider carefully not only which characteristics to include in the census, but also who should be the main targets for the dissemination of the results. It also means that it is necessary to plan carefully the effective and timely processing of the data that have been collected and the dissemination of the results.

B.   Population census compared with other sources of data

29.   Sound planning and design of a population census should take into account other available sources of statistics on the economic characteristics of the population (such as labour force surveys) and be consistent with them. In addition to the population census, a fully developed national statistical system will also involve (*a*) one or more national survey programmes based on samples of households or persons, including periodic labour force and income and expenditure surveys; (*b*) industrial or establishment-based censuses and surveys; and (*c*) statistics based on a range of administrative reporting systems, such as population registers, social security and pension records, income tax records and unemployment compensation reports.[2] Census designers can then focus their attention on those types of data needs for which the census is best suited or for which it represents the only alternative.

30.   If a country has held recent household-based labour force surveys, then it is important that there be consistency in scope, definitions, concepts, classifications, reference periods, age cut-offs, timing, and the like, in order to facilitate comparison of results (including validation of the census results as mentioned in paragraph 552). Results from

[1]   See United Nations (1998, 1996) and Hussmanns, Mehran and Verma (1990) for additional information on the uses of data on economic characteristics.

[2]   See Hussmanns, Mehran and Verma (1990), chap. 10, sect. 2, for further information concerning the comparative strengths and weaknesses of the various sources.

different sources will generally differ to some extent, but, for countries where the population census statisticians are not the same as the labour force survey statisticians, it is recommended that the two groups of statisticians work together to prepare notes that would guide politicians, users and the general public on the various sources of difference between the two sets of results. Such measures are especially important if the results differ so much that user credibility in statistical output will be affected.

31.   In many countries the population census is the only form of measurement for obtaining complete information on the economic activity rates for different population groups, and on their occupational and industrial distributions. Even though the population census is generally not the most effective device for collecting information on economic characteristics, in particular when the statistics needed are mainly for large groups at the national level, in many countries it is still the vehicle most used for collecting such information. Questions on economic characteristics were included in almost all population census enumeration exercises around the world in the 2000 round of population censuses.[3] The major advantages of the census compared to alternative sources are as follows: (*a*) its national scope and consistent geography; (*b*) the availability of other demographic and socio-economic variables for the same individuals; (*c*) its inclusion of all persons, permitting extensive and detailed cross-tabulations and the testing of complex models; and (*d*) the possibility of examining long-term socio-economic and demographic trends through comparisons of statistics from two or more censuses (assuming a certain minimum degree of conceptual and operational consistency between the censuses).

32.   All else being held equal, a population census has the potential to cover the entire population in all locations in the country, including specific types of economic activity, the rarest occupations and the most highly localized branches of activity, for even the smallest population group as determined by other variables. However, in practice, reporting and processing errors may limit the potential of the census in that regard (see paras. 33-35). Nevertheless, a census can produce, for example, summary measures of labour force participation for all civil or administrative divisions of the country, consistent with those for the country as a whole. Tabulations can be generated not only to the lowest geographical detail but also to the most detailed occupational and industrial distributions without having to be concerned about the precision of the estimates, as expressed by the sampling error that would be inherent in similar statistics derived from a sample survey. Moreover, multivariate classifications and tabulations are feasible, within the limits of the census budget.

33.   The main disadvantage of the population census as a source of statistics on economic characteristics is a function of its strength: because of its unparalleled geographical coverage it is also the major source of baseline data on the characteristics of the population in general. Economic characteristics thus represent just a few of the numerous topics that compete for coverage in a population census. Moreover, since the population census aims to cover the entire population in a country, the scale of operation is great, and that presents a major constraint on the number of questions that can be included. Such considerations substantially influence decisions on the following:

(*a*)   Whether economic characteristics should be covered or not;

(*b*)   How many questions can be included on economic characteristics;

(*c*)   Who is to be the respondent;

(*d*)   What depth of interviewing and probing can be realistically used to elicit the correct information on each individual.

34.   To ensure that the recorded information and the resulting statistics from the census are accurate and give meaningful representations of economic activity in the

[3] Based on a review, undertaken by the United Nations Statistics Division in 2007, of 176 national census questionnaires from the 2000 round of population censuses.

country, a lot of effort is required when designing the instruments for their collection and when training enumerators. Typically, however, most population censuses have to abstain from the use of the detailed questioning and probing required to identify whether some individuals are economically active. Moreover, since the census interview relies on proxy respondents and since census staff usually have less experience with economic activity concepts than staff engaged in labour force surveys, a census does not always capture reasonably accurate information on certain characteristics that only the individual might know, such as occupation, industry, hours worked, income and sector of employment, although to a large extent it shares those problems with the other statistical sources mentioned. It may also be that the measurement of certain variables (such as economic activity status, institutional sector, occupation, industry or the informal sector) can be problematic, hence requiring particular attention during questionnaire design and interviewer training. Similarly, attention needs to be given, especially in questionnaire design, training and field operations, to minimizing gender, ethnic or cultural bias in the measurement of economic activity status.

35.   As a consequence, the reliability of the measurement of each respondent's characteristics as obtained in a well-run sample survey is likely to be higher than in a census, with consequent implications for comparability of results. By the same token, administrative reporting systems are more likely than the census to track closely the concepts and data items that are important for the agency and staff responsible for the administrative reporting system.

36.   Census offices in some countries are administering long forms on a sample basis to investigate certain topics in greater depth. In such cases, the basic census questionnaire is often the short form that is applied to all individuals, while the designated sample receives the long form, which covers the questions on the short form as well as several others that may be investigated in greater depth. Data collected for a sample in the census exercise, however, have limitations similar to those of any sample survey data. In addition to non-sampling errors, such as under-coverage of the target population, non-response, recall lapses and inaccuracies in reporting, which are common to both the population census and sample surveys, the estimates based on a sample rather than the whole population will suffer from imprecision due to the sampling. Moreover, the potential for generating estimates at the smallest level of geographical detail and for coverage of rare occurrences of certain characteristics, such as certain occupations and types of economic activity, are greatly diminished with the use of sampling. However, the strengths of household survey-based statistics include the capacity of the data collection instrument to accommodate a larger number of questions on a variety of topics and also the possibility of estimating parameters comparable to those measured in the population census, within the margins of error associated with the sample design.[4]

37.   Specialized surveys such as time-use surveys, which are increasingly being used in many countries, can also be designed to provide comprehensive and accurate measures of the extent of economic activity in a country, for large groups of the population. Time-use methodology may be used as a tool for improving statistics on unremunerated work of women, men, girls and boys. At the present stage, such studies have not been extensively used for that purpose. Many of them, particularly the few that have been carried out in developing countries, have been small studies restricted to a single or a few localities; therefore it is not possible to use their results for national estimates. Nevertheless, the experiences with those studies can serve to sharpen the instruments and questions used to measure economic activity in national censuses and household surveys. A guide on collecting time-use statistics, with a proposed international classification for time-use statistics, has been issued by the United Nations to support this area of work.[5]

[4] See Hussmanns, Mehran and Verma (1990), chaps. 1 and 10-14, for additional information on sample surveys.

[5] Guide to Producing Statistics on Time Use: Measuring Paid and Unpaid Work (United Nations publication, Sales No. E.04.XVII.7).

Census administration from the perspective of economic characteristics

C. General considerations in census planning and design

38.   The search, in census planning and design, for the proper balance between substantive and operational considerations is not an easy task in the face of conflicting demands, resource constraints, the interrelated nature of various aspects of the entire census process and uncertainties about how full-scale operations will proceed. The publication entitled Handbook on Census Management for Population and Housing Censuses[6] covers in detail issues related to the planning, organization and administration of population and housing censuses. Section C highlights issues that have a direct bearing on obtaining data on economic characteristics, and decisions on items that have significant implications for overall census planning are presented in the context of measuring economic activity.

39.   Census administration involves establishing the organizational structure, which has typically included organizing the central, regional and field offices; designating personnel and recruiting additional staff, as necessary; constituting a planning group; and convening technical advisory committees. It is important that the census-planning group have at its disposal expertise in the collection, compilation, processing and analysis of statistics on economic characteristics. Ideally, especially given the intricacies of the concepts associated with economic characteristics and the difficulties they present in their application, it would be necessary to have a labour statistics expert on the planning team when any of the economic characteristics were among the topics to be covered in the census enumeration. That may be ensured, at the minimum, by having experts in the field of labour statistics on one or more of the advisory committees (depending on how the committees are organized).

40.   Two major forms of organization are proposed in the Handbook on Census Management: one is to form committees by subject-matter areas, in which case one of the committees would need to be on economic characteristics. The other is to set up a committee with subject-matter sub-groups, again with one such sub-group devoted to economic characteristics. If the former is chosen, it is important to provide opportunities for all the committees to meet together, to ensure consistency and coherence in approaches and better coordination. Advisory committees provide professional input and represent the user perspective. They also strengthen census planning by providing a forum for certain decisions to be debated among a wider expert audience.

41.   **Building on past experiences.** In planning any new population census the logical starting point is the previous population census. Accordingly, users should endeavour to obtain relevant documentation relating to the previous census. At a minimum this would involve reviewing the questionnaire(s) or schedule(s) used in the last census, reading about the basic census procedures and reviewing the main census outputs related to the economic characteristics of the population. In addition, evaluation studies of the previous census and any census monographs focusing on the economic characteristics that may have been produced can provide helpful information. In that regard, it is of great importance that census organizers prepare full documentation of the implementation of the census programme, not only of successful experiences but also of failures, so that both the users and current and future census planners benefit from them. Where documentation is limited, the production of such documentation needs to be considered and planned for early in the census-planning period.

42.   For countries that have prior experience in the conduct of labour force surveys, it is recommended that census planning and design build upon that experience, including in the design of questions and manuals, field operations and techniques, etc.

6   Handbook on Census Management for Population and Housing Censuses (United Nations publication, Sales No. E.00.XVII.15 Rev.1).

43.   In countries with ongoing programmes of labour force surveys, the timing of the population census may affect survey operations. Some countries even suspend survey operations during the preparations for and conduct of the population census. However, if resources permit, it is preferable to continue survey operations in order to maintain a continuous comparable series of survey estimates (see also chap. XIII, sect. D).

44.   In choosing the date for the conduct of the population census in which current economic activity is being measured (as opposed to usual economic activity: see chapters IV and V), one must be careful to select a date that does not fall during an abnormal period of economic activity. In some countries, the census is held during school vacation periods so that teachers may be used as field staff. However, the timing may have an impact on the measurement of current economic activity, since students who would normally be at school may take the opportunity of the school vacation to work and, if work is not available, may be classified as currently unemployed.

45.   **Producer-user consultations.** The importance of communicating information about census plans, operations and outputs to user groups, leaders of public opinion and the general public throughout the entire census process cannot be overemphasized (see United Nations, 2008b, paras. 1.112-1.116). The primary initial responsibility for those communication and publicity efforts lies with the census office itself. At a certain point, however, major public and private sector users can make critically important contributions to such efforts. The users of statistics generated by the census are particularly well situated to provide reinforcement, through their own formal and informal channels of communication, on the value of the census and the importance of cooperation by the public.

46.   Users of census statistics have valuable inputs to make with respect to the relevance of the concepts and methods, the adequacy of the statistics that will be produced, the content of the questionnaire and the forms of dissemination, and can state how the statistics could best meet their needs. The users include not only those from Government, such as ministries and government departments, but also researchers and academicians, representatives of employers and workers' organizations, other non-governmental institutions and special interest groups (see para. 16). Groups of users who should be represented in advisory committees include researchers who have analysed statistics on economic characteristics and gender statisticians and analysts. In recent decades, many of those users have drawn attention to the limitations of available statistics on women's economic activity, and they will provide invaluable assessment and feedback on how to avoid some of those shortcomings in the statistics that will be available in the census.

47.   Producer-user dialogue should also serve as a forum for educating the users about other data sources and their availability. The mechanisms for the dialogue may vary and may include informal contact as well as the technical groups mentioned previously. Many countries make public calls for submissions on the content of the census questionnaire and the output to be produced. To avoid suggestions that are not practicable, it is desirable to give users parameters within which they can provide comments. Users need also to be informed of the outputs of the major alternative sources of statistics on the economic characteristics of the population, such as household and establishment surveys and administrative reporting systems. From the users' perspective the sharing of census and other experiences will most likely result in improvements in terms of content, detailing, validity, reliability and timeliness, and, for some users, in addressing the concern that consistency be preserved between different data sources and from census to census. In many individual circumstances, improvements in terms of subject-matter detail, validity, reliability and timeliness cannot, within the resources available, be obtained simultaneously, so that trade-offs between different kinds of improvements will be involved. Moreover, improvements are often accomplished at a price, in terms of consistency over

time and with statistics from other sources. Here again, trade-offs are involved. Informed users can assist in determining whether and how those trade-offs will be resolved in ways that are acceptable to the most important users of the resulting statistics.

48.   **The role of users.** Forward-looking leaders of the financial and business community, who use census data to help with investment planning and marketing, may often be able to convince others in the financial and business community of the value of the census and census results. Their support can be useful in securing political support for the census in its preparatory phases as well as more general public support during the actual enumeration, and in some instances in mobilizing resources from the private sector for census activities. Labour organizations are also in a strong position to provide political and public support for the census, in both the planning and the enumeration stages. In both cases, it is useful that support for the census also continue during the post-enumeration stage of the census, such as processing, analysis and dissemination. Periodic specialized reports about the census plan and its progress, and outputs tailored to different sets of users are important tools for mobilizing continued support, both in the business community and among labour representatives. The likelihood that users will take an active role in supporting the census is directly related to their involvement in consultations on census content and outputs.

49.   There is considerable variation among countries as to who makes the final decision on census content. The responsibility may be given to the census commissioner, a minister or a parliamentary body, or it may be provided for in a presidential decree. However, regardless of who has ultimate authority in the matter, the process of providing input into that decision should be as broad as possible. Users of census data on economic characteristics should be invited to play an active role in the consultation process and, once invited, should be encouraged to participate actively. It is important for all users and potential users to understand from the start that the process is subject to considerable give and take, with a considerable amount of compromising involved. As mentioned elsewhere, many trade-offs are involved not only between subject-matter interests, but also between subject-matter concerns and operational considerations. Moreover, owing to the high degree of interrelatedness among the various aspects of census design, it is important for the number of economic items included in the census to be balanced with other topics to be covered, to assure the production of reliable data and the cost-effectiveness of the census operation.

50.   Finally, in respect of general considerations in census planning and design, it is important not to overlook the measurement of economic activity in collective housing units. They may be particularly important in some countries (for example, mine workers in dormitories). Conceivably, the identification of collective housing units could also be used as an indicator of economic well-being.

D.   Preparation for enumeration

51.   In addition to the user-producer dialogue mentioned above, some other pre-enumeration activities, ranging from the determination of enumeration areas to the carrying out of organizational pre-tests, apply to the census exercise as a whole.

52.   **Mapping.** One aspect of the mapping exercise that needs particular attention is the mapping of non-typical areas, such as large tracts with sparse population, rapidly growing areas, congested urban areas and areas with ill-defined boundaries (See United Nations, 1992, paras. 189-193). Mapping is intended to ensure the full coverage of the population in a census, but it is also of particular importance for measuring economic activity. In many agricultural societies, some farmers or farm workers reside on part of their farmland. Those tracts of land may be difficult to reach or, lacking clearly defined

boundaries, difficult to identify. Such problems could lead to the under-enumeration of persons employed and residing on those types of farms. Areas with a very dense population such as the slums in some major cities may also present a problem for the measurement of economic characteristics: many of the residents of such areas are involved in marginal activities that tend not to be reported with the traditional census questions on economic activities. In addition, the crowding in those areas could present difficulties in the listing of households, leading to an undercount of the population and therefore the under-enumeration of those marginal activities, as well as to the unequal division of the enumeration areas and workload. Finally, care must be taken to ensure that collective housing units are correctly identified since the residents of collective housing may exhibit particular economic activity characteristics.

53.   **Pre-tests.** The use of pre-tests to determine the efficacy of the data collection instrument, the field operations and even the data processing is a critical issue to consider for every census round since the circumstances and characteristics of the population change in a way that could affect the way respondents receive and respond to questions. Both new formulations of questions and those that were previously used need to be subjected to pre-tests. It is particularly indispensable, in the light of the changes in the scope of the production concept and consequently in the intended understanding of "economic activity", for pre-tests to be conducted to ensure the correct measurement of economic activity.

E.   Decisions about the coverage and format of the enumeration

1.   Scale of the census and use of sampling to expand content

54.   Census authorities often wrestle with pressure to expand the list of topics covered in the census beyond what they consider to be consistent with preserving public good will and a reasonable level of data quality in the overall census results. The pressures are particularly acute in countries that lack alternate sources of data, for example, continuing programmes of sample surveys, to provide statistics on many important topics. However, in general it is the statistically most developed countries, in other words, those with a range of different sources for national statistics, that also tend to have the most comprehensive and complex population censuses. That suggests that considerations of the costs of the census and the capacities of the census organization will normally be decisive in determining the scope of the census.

55.   One strategy used in some countries, developed and developing, to ensure that the basic census form is relatively brief is to use sampling as part of the census operation itself. That approach, which is described in Principles and Recommendations, Revision 2 (United Nations, 2008b, paras. 1.422-1.427), involves obtaining information on some topics for only a sample of the population enumerated in the census (see para. 36).

56.   When sampling has been used in national population censuses (combined with the use of the long form), economic characteristics have often been among the topics included in the long form. Owing to the complexities inherent in the concepts related to economic characteristics, more than one question is usually required to measure an economic characteristic. Moreover, for some topics, such as employment, income or informal sector, further probing may be required to ensure that all relevant situations have been covered. The decision to include economic characteristics in the census on a sample basis involves deciding (*a*) which, if any, of the characteristics should be in the short form, thus addressed the whole population; and (*b*) which of the characteristics should be included in the long form, which is administrated to only a sample of the population.

57.   The use of samples provides greater flexibility for investigating topics on economic activity. It can significantly reduce the response burden for respondents who are can-

vassed through the short form, as well as certain aspects of census processing costs. The possible impact of the use of sampling on the cost of census fieldwork will vary with such factors as the method of sampling used, population density and whether self-enumeration is used. Gains in terms of reduced total operational costs could be offset by the costs of complicating the interviewing instructions and procedures used in processing the census results, and of introducing sampling errors into the census results. Nevertheless, some countries have found that the gains from the use of sampling to expand the subject-matter scope and detail of the census outweighed the costs in complications, which need to be addressed if the application of sampling is to be successful.

58.   Census users and producers in some countries argue that, as the provision of small area (such as for localities or polling district level) and small group data on a nationally consistent basis is one of the main functions of a census, restricting data collection or processing to only a sample of units is inconsistent with the basic objective of a census. It may also be relevant that in most small countries the relative cost reductions from the use of sampling in the collection or processing of census data are much more limited than in larger countries.

59.   Sampling, if used, is usually carried out at the data collection phase and more rarely at the processing stage of census results. Sampling at the data collection stage is carried out in three broadly different ways, as follows: (*a*) a sample of persons (that is, only one in n persons in the census); (*b*) a sample of households (that is, all persons in one in m households); or (*c*) an area sample of some sort (that is, all households and unrelated individuals in one in k enumeration areas or census blocks). Although sample (*c*) is the simplest to administer in the field and may result in the greatest cost saving, it will usually yield estimates with large standard errors owing to the intensity of clustering. Whatever design is used, an expert in sampling should be involved with the choice, the actual sample design and its implementation, as well as in the preparation of the estimates based on the sample, including estimates of sampling errors.

60.   Sampling in data processing can be used to produce advance tabulations or to control the costs of the processing operation. Since the coding and processing of detailed occupation and industry information can be a time-consuming and costly part of census processing, occupation and industry items are often among the first topics to be proposed for the sample, if sampling is to be used. However, modern processing technology can help to reduce the need for the use of sampling in processing census data.

2.   Issues relating to counts of population present and usual resident population

61.   The issue of whether people are generally enumerated at the place of their usual residence (that is, a "usual resident population" approach) or at the place they are found at the time of the census (that is, a "population present" approach) gives rise to much discussion in many census offices around the world (United Nations, 2008b, paras. 1.461-1.468 and 2.71-2.77). However, there may be little practical difference between the two concepts if most of the population will be found at their usual places of residence on the census enumeration day. For that reason, countries often choose as the census date a day that maximizes, to the extent possible, the likelihood that persons will be found at their usual places of residence. Furthermore, most censuses actually use a combination of both approaches. For example, homeless persons are always enumerated on a population present basis even in a usual resident population census, while people away from their usual place of residence at midnight on the census reference night are usually enumerated on a usual resident population basis even in a strictly population present census if they are working or in travel status on enumeration night. In any event, it is recommended

that consistent approaches be used in labour force surveys (if any) and in the population censuses. Details may differ but should be minimized.

62.   Some employed people sleep in one location during their working period and in another during the days off work. Others commute daily across borders or over long distances, and those situations raise particular concerns for correctly identifying the place of residence.

63.   From the perspective of users of census statistics on economic characteristics the two main concerns should be the following: (*a*) which approach will minimize under-enumeration of the economically active population; and (*b*) which approach will yield the most reliable responses on the economic items. In both cases, there will be a trade-off between the coverage and response errors for visitors present compared to those for absent usual residents. Proxy respondents, which are still used in many censuses, could present special problems with either approach, if someone other than the individual is reporting on such characteristics as occupation, industry, hours worked, income and usual activity status.

64.   Objective assessment of the trade-offs can be derived only from empirical evidence from earlier census or survey experience or census tests. Similarly, although in theory the census coverage concept adopted should be in harmony with national accounting practices for the country,[7] the practical trade-off considerations relating to coverage and response errors should be determinative in deciding upon the concept to use in the census. Many national accounting concepts and distinctions are not possible to implement in the context of the census. In some countries where a substantial number of persons are absent for work-related reasons, information about absent household members can distort the results from the economic items under a usual resident population enumeration, unless those absent can be easily identified separately. For that reason, a time limit for absence is sometimes introduced. For example, to be included in the census the absent person has to have been at the household for a specified duration (for example, at least six months) in the last year and/or have no other private dwelling in the country that could be called the person's usual residence.

3.   Method of census data collection

65.   Principles and Recommendations for Population and Housing Censuses, Revision 2 (paras. 1.58 to 1.73) describes various methods for census enumeration: the traditional approach, the register-based approach, the rolling census approach and the traditional approach with yearly updates of characteristics.

66.   For censuses that are not register-based,[8] a major design issue in planning is whether the information will be obtained largely by interviews with each household or by self-administered questionnaires, which each household completes on its own without the use of census enumerators. The former approach is usually termed the "interview" method and the latter the "self-enumeration" method. Which method will be the most cost-effective in terms of the level of overall enumeration and the quality of data obtained depends on a variety of factors, including the level of literacy in the population, the pool of persons available to serve as census enumerators, the size of the country and population settlement patterns.[9] Linked to any decision on using the self-enumeration approach is how the census form is to be delivered to each household and, once completed, returned to the census organization. Countries with reliable postal systems and good address registers may rely primarily on a mail-out and mail-back approach, while other countries may use census enumerators to drop off and/or pick up the census forms. In other words, the operational and quality issues that go into the choice of methodology do not generally turn on issues related to any single subject-matter field.

[7] Commission of the European Communities and others, System of National Accounts, 1993 (United Nations publication, Sales No. E.94.XVII.4); see, in particular, paras. 14.7-14.21.

[8] Register-based censuses take advantage of data available in existing administrative records on households, persons and dwellings, and link them at the individual level with data on business, tax, education, employment and other relevant registers. The existence of a unique identification number for each individual, household and dwelling is of crucial importance.

[9] This question is covered more fully in part I of Handbook of Population and Housing Censuses (United Nations publication, 1992), paras. 307-315.

67.   As an extension to self-enumeration in some more developed countries, the Internet is being used in data collection for the census. The use of the Internet introduces special considerations for confidentiality and data security, but has advantages in that it facilitates question sequencing (allowing responses to one question to trigger the next logical question to be displayed on the screen for the respondent), and can be a less expensive way to conduct a census. In many countries, especially the less developed, access to the Internet is limited and therefore this approach would not be appropriate.

68.   The interview methodology is particularly suitable for the economic questions as it allows enumerators to probe, in a uniform manner, into the wide range of possible economic activities and their characteristics. However, the interview methodology places a heavy burden on the enumerator's understanding of the questions and how to ask them and record the answers, and thus on his or her selection and training, as well as on the questionnaire. By contrast, from the perspective of the economic questions, the self-enumeration methodology puts an even stronger dependence on a clear simple questionnaire and the few easily understandable notes, which a respondent with average literacy can read and follow. Design and testing of the questionnaire are important for both approaches but are particularly critical when the self-enumeration method is used.

69.   The approach to be adopted should be reconsidered very early before each census. With increasing literacy levels, some countries that formerly used only the enumerator approach switch to using a mixture of both approaches: for example, self-enumeration in large urban centres and the interview approach elsewhere. Such a combination of enumeration methods may lead to a smaller field staff and hence lower enumeration costs. However, it also requires procedures to obtain responses from those households in self-enumeration areas that are unable or unwilling to complete the census form on their own. Any new procedure will require thorough testing before it can be accepted.

4.   Forms used for data collection

70.   Traditionally, a number of developing countries designed their census enumeration form so that it would fit onto one page for most households (a second page being used for large households).[10] The arrangement had the advantage of being simple to use in the field and easy to print. The single-page format also permitted the individual sheets to be bound together so that all the forms for one enumeration area (EA) could generally be preserved together in a single packet with an enumeration area cover sheet and a stiff back. This approach is operationally simple and has many practical advantages.

71.   However, the one-page format also has a number of undesirable features. For example, the decision to include or exclude specific topics has depended, in part, on whether there was sufficient room to include them on a single page. As a result, in some countries, the decision has been to exclude items or topics from the census while in others additional items have been accommodated on a single-page form by enlarging the size of the page to such an extent that the census forms became expensive to produce and difficult for the enumerators to carry and work with. In other countries additional items were accommodated on a single-page questionnaire by so reducing either the length of some questions or the print size that the form was difficult to read or the space for recording responses was insufficient. It is particularly important that sufficient space be provided for recording responses to items such as industry and occupation, which should not be pre-coded. An overly simple or short question on activity status does not elicit an accurate response that captures the intended definitions of the concepts.

72.   With economic development and change, countries that in the past gathered little or no economic information in the census are giving serious consideration to expanding the number of topics to be included in the next census. The increasing education levels

10 In a number of English-speaking countries the term "questionnaire" has traditionally been used to refer to enumeration forms that contained the verbatim wording of the actual questions to be asked, while the term "schedule" was used to refer to enumeration forms that simply listed the topics being investigated at the tops of the columns. In the present Handbook, the terms questionnaire, enumeration form and schedule are used interchangeably.

experienced by most countries also means that the public and the census staff may have a better appreciation of the need for more information and are better able to cope with more questions. Several options for expanding the census form are possible. If a one-page form was used in the previous census, using the back of the form provides a simple way of making a minimal expansion. Some countries use a booklet for each household with a standard paper size (for example, A4). That allows for the use of larger and clearer fonts, provides more space in general for recording responses and facilitates data processing operations.

73.   Many countries separate the census enumeration form into two parts: (*a*) the household roster with basic information on one page, and (*b*) a separate page for each person above a specified minimum age, containing detailed items on economic and other topics. The two parts are usually printed as "household booklets", with provision for a single household roster and 8 to 15 individual (adult) person forms all stapled together. A similar approach is used in many household surveys.

74.   A variety of other design issues also arise. For example, a decision will have to be made as to whether the census form should be laid out in "portrait" or "landscape" format. (The terms "portrait" and "landscape", as used here, have the same meaning that they do in common word-processing software.) The landscape layout generally has the questions across the top of the page (in columns) and the individuals listed along the side (in rows). This format is best suited for a form containing relatively short questions that need simple pre-coded or numeric answers, such as age or number of children ever born. The portrait layout commonly has the questions listed along the rows, covering more than one page, and individuals listed across the columns on top of the pages. The portrait layout definitely is better for more complex questions, such as those related to economic topics. For countries that use a booklet for each household or individual, the choice between the two layouts may not be as critical, since space limitations are not so severe. In some cases a mixed methodology has been employed, for example, landscape layout for the basic household listing and portrait layout for each adult person.

75.   The collection of data on economic characteristics often involves the use of complex concepts that in turn sometimes call for complex skip patterns to ensure that each person is exposed only to those questions that fit their circumstances. Those concepts and the related skip patterns have often evolved on the basis of experience with labour force surveys. In the context of the census, however, certain simplifications must be introduced to allow census enumeration to proceed smoothly. Some of the strategies for reducing the complexity of the economic items for the census are considered in parts two and three. It is essential that any simplifications introduced to accord with the constraints imposed by the scale of the census be (*a*) carefully reviewed by experts in labour force statistics and analysis, and (*b*) thoroughly tested.

76.   Countries with two or more languages commonly spoken raise particular problems for census taking. In some cases questionnaires have been printed in different languages. In others, texts in two or more languages are put on the questionnaires. As previously noted, space is usually a major problem for census questionnaires and the inclusion of two language versions of each question can lead to a very cramped questionnaire (see also United Nations, 2008b, para. 1.187). The alternative is, however, to rely on the abilities of the enumerators to translate accurately on-the-spot both the census questions and the replies. That procedure obviously places extreme demands on the enumerators. In all cases, very special attention needs to be given to these matters in manuals and during the training of field staff.

77.   Decisions related to the design of census enumeration forms depend on a wide range of substantive and operational factors, so it is important for each country to review and

adequately test its design and layout well before the actual census. In addition to consideration of the census form in the light of issues at the data collection phase, data processing considerations also will need to be taken into account. Moreover, the appearance of census questionnaires should be appealing. In almost all countries, the population is now accustomed to an attractive layout of most things they see. Self-enumeration questionnaires particularly must be attractive, but it also applies to interview schedules. If the enumerator finds the questionnaire to be an attractive, well laid-out document, he or she is much more likely to fill it out neatly and correctly.[11] Coders and data entry staff will function more effectively if the form they are using is well designed. Questions on economic items (or those on any other topic) that are cramped into a small column heading in a truncated form usually give poor results since it is difficult to fully reflect the intricacies of the relevant concept in such a limited amount of space. In some cases, when the responses are not pre-coded, the space provided for responses is not adequate for recording the complete response (see also section H, "Issues relating to questionnaire design for data processing").

78.   Accordingly, census planners should, if at all possible, use or consult professional layout designers. At a minimum, advice can be obtained from professional printers on type styles, sizes and the use of colour, if that step has not already been taken. Excessive use of bright colours should be avoided as they can distract from the main objectives. They can also be wearing on the eyes of staff working on the forms, particularly data processing staff. However, the use of limited pastel soft colours and shading can make questionnaires attractive and more readable. In addition, the use of scanning and optical character recognition methods places particular constraints on colours and questionnaire designs. Employing different fonts and type sizes for different purposes can make the form easier to follow. For example, consistently assigning differing formats to the various elements (that is, questions, prompts, instructions and pre-coded answers) of the census enumeration form can greatly facilitate its readability. The cost of professional assistance in layout design is usually well repaid in better quality interviewing and responses. In addition, such services are often available free or at little cost from other government agencies.

F.   Selection, training and supervision of field staff

79.   It is of vital importance that field staff be adequately trained in the collection of data on economic activity and related characteristics since it is a major issue in the population census for most countries.

80.   The field staff (enumerators, and first- and second-level field supervisors) are an essential element of a successful census, regardless of the type of method employed, interview or self-enumeration, although the field staff play a more critical role when the interview method is used. Given the importance of the subject, both Principles and Recommendations for Population and Housing Censuses, Revision 2 (paras. 1.210-1.215) and the Handbook on Census Management for Population and Housing Censuses[12] (see chap. III, sect. C, and chap. II, sect. B) give it considerable attention.

81.   The relatively short duration of training for census enumerators cannot make up for major deficiencies in their basic education. Experience has shown that a relatively advanced level of formal education does not always guarantee that trainees will become the best census enumerators. For pre-tests, it is recommended that the staff used should have the same characteristics as the actual enumerators. Mid-level officers in close contact with the local community have been generally found to be most effective in asking economic questions in censuses. Schoolteachers working in their own areas where they are familiar with community activities have often been found to be particularly effective.

82.   The training of fieldworkers will need to include trainers, supervisors and enumerators. Trainers should, if at all possible, have some earlier training or teaching qualifica-

tions and experience. The training emphasis for the economic questions must be on asking the questions, understanding the major skip patterns, applying any prompts and probes as necessary and learning how to record the answers properly. The content of training on the economic items should leave the census field staff comfortable as to how they are to complete the census questionnaire for the major categories of persons who will be encountered during enumeration: those currently active (wage and salary workers, the self-employed and contributing family workers), those not now economically active but who were so formerly and those who have never been economically active. It is important that fieldworkers understand the full range of possible replies that can be given to each question and the type of information that those questions are designed to obtain. However, they do not need to understand the underlying concept and scope of economic activity. Thus enumerator training should give little or no attention to the economic theory or conceptual frameworks that underlie the economic items in the census. With the limited amount of training possible for the census enumerators, complex notes or long lectures need to be avoided, and the training should focus on practical and operational issues and on an understanding of the questions that have been formulated. As indicated earlier, training should ensure that gender, ethnic and cultural biases are minimized.

83.   Training materials and examples must be simple and aimed at bringing all enumerators up to a minimum level of competence. It is important to avoid complex examples and problem cases that may entertain the brighter enumerators but will almost certainly confuse most. Unusual cases will always occur and they should be written on blank spaces on the questionnaire with full details for later resolution by supervisors or office staff. As with any other type of statistical sources, uniform treatment of similar situations is a key to good economic statistics, so that good enumerator manuals and sets of examples covering all major common economic situations are essential. Training examples need to be selected with care so as not to foster gender-based or age-based biases or other stereotypes related to economic activity. In some cases, somewhat different examples may be prepared for enumerators working in distinctly different areas, for example, urban examples for urban enumerators and rural examples for those in rural areas. However, that approach should be used cautiously since enumerators in rural areas may encounter persons with urban work experience and urban enumerators, persons with rural work experience. Accordingly, census enumerators, wherever located, will need at least some preparation to deal with the typical situations in both urban and rural areas.

84.   Practical mock interviews in groups covering the main situations have proved an effective technique of training, but by far the most effective training is field practice in a real situation with a detailed review on the spot or immediately afterwards. At least one and preferably two practical field sessions should be included in any census training programme. It is often useful to include in the enumerator training sessions some audio-visual (videotape, DVD, film or slides) or audio (audio cassette or compact disk) materials that have been centrally prepared. Such materials, which illustrate the basic structure of major variants of common interview situations, can help to offset the damage arising from a poor trainer. This approach is particularly valuable for helping enumerators obtain a correct grasp of complex items with skips such as those that frequently occur in the economic questions.

G.   Field tests and the pilot census

85.   For the purposes of the present Handbook, pre-tests are taken to mean any test of census questionnaires or procedures. Many of the tests used may be quite narrow in scope, particularly in the early testing of alternate wording of census questions. Census pre-tests can be based on purposive samples designed to maximize the number of relevant

cases in terms of the content or procedures being tested. As census planning proceeds and options are narrowed, tests tend to become larger and involve more diverse settings. However, purposive samples are still adequate.

86.   It is essential that all items proposed for use in the census, including the economic questions, be adequately tested (United Nations, 1992, paras. 238-240). The results of those tests should play an important role in decisions about census content and procedures at every stage of census preparations. The fact that the same questions have been successfully used in other countries, in labour force or other sample surveys in the same country, or even in the previous national census, does not obviate the need for testing in the light of current national conditions, using members of the census staff, and adopting planned census procedures. Ideally, small separate tests should be held for different segments of the main census questionnaire, with at least one test focusing on the economic questions. Pre-testing is generally not expensive if test areas are carefully selected to include many different situations for the population and the test is well designed.

87.   It is important that a clear objective be defined for each census test. Too often census tests lack such well-defined objectives so that the results are difficult to assess and the process of developing a reliable and cost-effective census enumeration form is delayed or led astray. Census staff who plan and conduct tests need to be able to set out one or more clear test objective(s) prior to the test and then give a critical evaluation of the results in the light of those objectives. Very rarely do economic questions, particularly those for use in a census, work perfectly. Problems always arise with persons or circumstances that are difficult to enumerate or classify. The best design will function well in the common situations and minimize the errors arising in the rare situations. This sort of fine-tuning of the questions, forms and procedures can only be achieved through field tests.

88.   A pilot test is generally taken to mean a major test of both questionnaires and procedures. Often there is only one pilot test that serves as a complete test of all the questionnaires, forms and procedures. The final pilot test should be, in effect, a dress rehearsal of the complete census operation, including the tabulation phase. As such, the final pilot census should involve one or more relatively large administrative units so that census managers have some sense of the logistical problems that arise in a large-scale operation. In theory, no changes in census content or operations should occur between the final pilot and the actual census, although in practice some minor refinements are often introduced. Since the use of untested questions or procedures in a census involves considerable risk, any changes introduced after the final pilot would require strong justification and, even then, great restraint should be the guiding principle.

89.   Once census tests have been carried out, it is important that the results of those tests be preserved in written form. That will require a certain amount of discipline by census managers. With numerous demands on the time of those responsible for census preparations, there is a temptation to avoid or delay writing up the results of pre-tests and pilot tests. However, such documentation is a building block for future improvements: it should always be prepared and included in a handbook on the census that can be passed on to the census staff who work on the next census. Such documentation is an important element of the knowledge base needed to build experience and expertise in the country. Moreover, the results of census pre-tests and pilots, properly analysed, can also usefully contribute to the census evaluation programme (see part six).

H.   Issues relating to questionnaire design for data processing

90.   Data processing staff should be involved from the start of census preparations. Their participation in the development of census questionnaires from the earliest stages is critical. Such participation will help to ensure that data processing considerations are

adequately taken into account at the design stage (see also para. 77 concerning questionnaire layout and design). In addition, allowing the data processing staff to better understand the substantive and logistical priorities and concerns of others involved in census operations and planning will minimize the likelihood that subsequent decisions at the processing stage will introduce major problems.

91.   Where data processing staff have not been involved at the early stages of census planning, it has sometimes been necessary to completely redesign the proposed questionnaire on the basis of data processing considerations. For example, even first drafts of questionnaires should have some allowance for coding or optical mark reader (OMR) boxes. As with all aspects of questionnaire design, the final layout should be a cooperative effort. The data processing requirements should not, however, overshadow other important census design considerations, such as substantive content or the clarity and attractiveness of the forms for use in the field. For example, an unattractive form with a confusing skip pattern and full of boxes, with little space for clear questions or recording responses, is to be avoided regardless of its potential ease of use for data entry. As indicated in chapter IV, paragraph 199, the use of multiple responses may also be avoided. Compromises are usually necessary to produce a questionnaire that functions well for all purposes.

92.   **Use of pre-coded response categories.** One approach that facilitates data capture is the provision of pre-coded answer categories in the census schedules and questionnaires, wherever possible. Pre-coded answer categories have many advantages. Not only do they make processing easier, they also function very well as a guide to enumerators (or respondents) as to the type of responses needed. However, pre-coded answers do have limitations. They function best when there are only a relatively few relevant answers to a clear, precise question, or for numeric answers. The pre-coded list must cover all possible situations and it is usually best to allow for an "other, please specify …" category that can help capture misunderstandings of the pre-coded alternatives as well as situations that are rare or otherwise not anticipated. Written responses in the "other" category may be re-coded to their correct category in the classification, or a new category created if necessary. Testing should be used to help to establish the pre-coded list and identify problems with some of the categories, so that notes or examples can be supplied on the questionnaire or in the enumerators' manual.

93.   Enumerators often use such lists of pre-coded answers as a prompt list to respondents. Lists can be useful but should be planned and not left to the enumerator: for example, "Did you do any of the following … [list of activities] … during the [reference period]?" Experience shows that most lists used as prompts have some order bias; that is, the first items tend to have more answers recorded than correspond to reality. Such bias is most pronounced in attitudinal surveys but can still be a factor in the census for factual questions on economic activity. It is therefore best to try to get enumerators to read through all categories on the list before asking the respondent to identify a specific activity (or response category). There is no predetermined order in which questions or response categories are to be presented. However, many survey techniques, such as changing the order of the prompt list with different respondents or placing commonly forgotten items early in a list, might be tried in testing before deciding upon the order of the items in a final pre-coded list. It is often surprising to see what a difference a change in order can make in the frequency of responses. The order of categories in publications or in conceptual manuals definitely does not mean that those categories have to occur in the same order on the questionnaire.

94.   Pre-coded answers should be used only when the list of possible answers can be reduced to a short list from which an enumerator can easily choose on the basis of the respondent's reply, and for which the categories are all clearly distinguishable. For that reason, pre-coding should be avoided for "occupation" and "industry". Not only is it

almost impossible to make lists for those two questions simple and clearly understandable in any language, even with good training for the enumerators, but limiting the number of categories for the two variables (as required for pre-coding) results in statistics of very limited usefulness, especially if the pre-defined categories have been chosen to reflect the structures of the classifications rather than the economic structure of the country. Nevertheless, owing to the expense involved in coding the two items, some countries have chosen to pre-define the categories as a solution to the coding of occupation and industry, and the response alternatives have sometimes been a list based on the first digit of the respective classifications. However, pre-coding limits the kinds of basic economic analysis that could be carried out with detailed occupation and industry data. The unique opportunity the census provides to obtain detailed occupation and industry data should not be missed.

I.   Data processing and dissemination of census results

95.   During data capture, the information should be edited, including checks for valid ranges for codes and checks for consistency (for example, no answers to the questions on economic activity if the person's age is less than the level shown on the questionnaire).

96.   In many census operations there is little coordination between those responsible for census content and those responsible for the processing and dissemination of census results. Although both census producers and users are often open to discuss outputs, under the pressure of deadlines such discussions are frequently pushed aside by the need to consider census content, first for testing questions and procedures and then for printing the census forms. Furthermore, as the census date draws near, senior census staff tend to be increasingly involved with the management of field operations. As a result of those circumstances, many decisions on processing, with critical implications for the quality, cost and timing of census outputs, may be made solely by the census data processing staff without the benefit of inputs by users or subject-matter specialists. Such inputs to decisions on census outputs and certain aspects of the data processing timetable can be invaluable for both the user and the census office.

97.   The results from a population census may differ from those from a labour force survey held at much the same time. When this occurs, it is important for users to be given explanations of any differences and guidance in interpreting the census results as well as in making comparisons with alternative sources. In some countries, the same staff that prepare the analysis of the labour force surveys results also validate and analyse the population census results on economic activity and prepare the analytical text of that component of the census reports.

98.   Advance planning is required to ensure that user and subject-matter specialist inputs are obtained. In view of the tight timetable of operations in the pre-census period, users and subject-matter specialists must make time and staff available for such consultations as well as for those on census content. The agenda for the discussions would cover processing issues, such as how to deal with inadequate responses, as well as issues related to traditional census outputs, including the following:

(*a*)   The extent and timing of the basic tabulations, including those on economic characteristics, for the nation as a whole, for each province and for other geopolitical aggregates;

(*b*)   The extent and timing of the basic cross-tabulations of data across subject-matter fields and of detailed tabulations of topics with very extensive response categories such as detailed occupation and industry data, income distributions or place-of-birth data;

(*c*)   The content and timing of various published census reports and monographs.

99.   In addition to traditional census outputs, a range of newer intermediate and final census outputs should also be covered in the consultation process. Chief among them are the following:

(*a*)   Various types of databases developed to store census information to facilitate the production of additional tabulations once the initial round of census tabulations has been completed;

(*b*)   Public-use samples of anonymized census microdata that enable users to carry out special analytical studies, including multivariate analyses, on their own;

(*c*)   Other types of census outputs in machine-readable formats, and Internet dissemination of census results and/or public-use microdata files.

100.   The focus of the consultations should be the specification and clarification of user needs and the provision of advice from a user perspective on trade-offs that may arise in processing and dissemination. It is best if users avoid becoming enmeshed in providing advice on the actual processing work. In a country where, in some highly specialized area, such as occupational coding, the centre of technical expertise lies outside the census office, arrangements may be made to draw on that expertise on a regular basis outside the usual user consultation process. Users also need to be made aware that at times their interests in certain kinds of census outputs may be frustrated by a lack of resources for census processing, or by the need to avoid compromising laws or other restrictions designed to protect the confidentiality of individual responses. The former set of problems can sometimes be resolved by the user, particularly if it is a public agency, helping to underwrite the processing costs involved. Issues related to census confidentiality, which are closely related to the public confidence in the census and indeed to most statistical work in the entire country, are not always solved so simply. However, several approaches have been used by different countries to help meet user needs for detailed data while still providing the required protection for individual respondents (United Nations, 2008b, paras. 1.339 and 1.376).

Decisions related to topics for measuring economic characteristics

J.   Selecting topics to be included

101.   Table 1 in Principles and Recommendations, Revision 2 presents a list of topics related to economic characteristics that may be included in a census undertaking, including the following: (*a*) activity status (employed, unemployed, inactive); (*b*) status in employment; (*c*) occupation; (*d*) industry; (*e*) institutional sector; (*f*) informal sector; (*g*) informal employment; (*h*) type and geographical place of work; (*i*) working time; and (*j*) income.

102.   Most population censuses will cover only some of the topics. For too many countries the census is still the major, if not the only, source of information on several important population characteristics. Thus users of statistics and special interest groups often compete to have their topics of interest included in the census. The magnitude of the census exercise requires that the number of questions be limited so the enumeration can be completed within a reasonable period of time, even though the cost of finding and contacting all the relevant respondents is usually the main cost element for the census operation.

103.   Topics on economic characteristics are therefore assessed in competition with other important topics. In many countries, the space allocated for economic characteristics would permit four or five questions, and the most common topics investigated have tended to be activity status (employment and unemployment), status in employment,

occupation and industry.[13] Given that the preceding topics have been included more frequently than most others, there is more experience with questions that concern them and greater knowledge about how well those questions have worked. For most of the other characteristics—in particular the institutional sector and the informal sector—the limited experience with them in censuses and the requirements with respect to the number, format and content of the relevant questions to ensure a reasonable degree of accuracy and quality of responses suggest that more testing may be needed before it is decided whether it will be possible and worthwhile to include them in the basic census questionnaire.

104.   A general principle for selecting among the topics for census exercises is therefore to include only those for which experience and testing have demonstrated that a reasonable degree of accuracy can be assured with the resources available for census operations. The issue of whether other sources of statistics on the topic exist and the reliability of their estimates for other than national aggregates will also be relevant. For example, when the users' main concern is to get reliable statistics for national aggregates, countries with a regular household survey programme may do well to limit the topics covered in the population census and consider collecting data on the other topics through a labour force survey or other specialized survey such as an informal sector survey. In the context of the population census, another option when such alternative sources are not available could be to leave some of those questions to be investigated on a sample basis, if the approach employing both short and long forms is adopted in the census (see sect. E.1 above and part six for further information on the use of sampling in population censuses).

K.   Determining the location and sequence of the economic items on the census form

105.   The position of questions within a questionnaire and on a page should reflect their importance. Other things being equal, the later an item appears in a form or the lower it is on a page, the less attention it will be given; consequently, the completeness and quality of responses will suffer. In many cases, unfavourable positions for individual items cannot be avoided. Some items must appear later in the form than others. However, a variety of actions, ranging from the design of the census form (for example, arrows, different type faces, boldface, colour) to the training of enumerators and supervisors, as well as field supervision, can help to overcome the potential drawback of an unfavourable location on the census form.

106.   The usual sequence of census questions generally does not favour the economic questions. There is very often a progressive exclusion or filtering out of the younger population from successive sets of items, whether the census is carried out with a self-enumeration form or by a census enumerator. For example, sex, relationship and age are obtained for everyone; information related to education for all but young children; and information on training, fertility, and economic activity and related topics for the population above an even more restricted age cut-off. The sequence is to a certain extent logical but does tend to de-emphasize the questions on economic characteristics. Following the above approach, economic characteristics could be one of the last sets of topics in a census questionnaire unless there are questions on fertility and mortality, which are sometimes placed at the end of the questionnaire since they are potentially sensitive and because fertility questions are often for women only. Questions placed at the end present special challenges, especially if respondents find the questionnaire or interview to be too long or the nature of prior questions to be sensitive or personal.

107.   The first question or questions of the economic block of questions are almost always designed to identify those who are or sometimes have been "economically active". The subsequent economic questions on the census questionnaire apply primarily to that

[13] For a more detailed review of national practices with respect to the inclusion of topics in a census, see United Nations (1996) and ILO (1996a).

population, and thus the question(s) to identify them are the key to the whole block of questions on economic topics. For example, those who are not currently economically active do not have a current occupation, industry or status in employment, and those who are not currently employed cannot respond to questions related to current working hours or place of work. Thus the positioning of questions or topics should facilitate the use of skip patterns and thereby produce efficiencies, for example, by reducing the length of the interviews. While different ordering of questions is possible, it is useful to group questions relating to descriptive characteristics of the job separately from questions relating to descriptive characteristics of the establishment. One possible way of ordering the questions is to start with the economic activity status of the individual, followed by the selected topics on descriptive characteristics of the job for those that are employed; by those on the descriptive characteristics of the establishment; and by those on other characteristics of the individual (depending on which set of items is to be investigated). It is important for the ordering of the questions to be pre-tested to ascertain their effectiveness.

108.   Following the above principle, it is recommended that the "occupation" question be asked before the "place of work" and "industry" questions and after the "status in employment" question. The focus in the status in employment question is on the individual's work contract situation, and it is logical to follow it with occupation in which the focus is on what the individual does. For place of work and industry the focus then shifts to the place where the person works and the activity at that establishment. Following that order helps to prevent confusion between occupation and industry. Occupation may also be considered to be the most important topic and may therefore be given priority by being placed first (that the sequence is different in the well-tested questionnaires of Australia, Canada and the United States of America may, however, indicate that question formulation is as important as sequencing for good results.)

109.   When an income question is asked it is usually the last in the economic block of questions. This is done partly because it is logical to ask all details of employment and then to ask about the income from that employment. It is also often done to sequence all adults to this point, not just the employed persons, and to ask about total income including investment income, pensions, and the like. It would be illogical to sequence those not employed to a question part of the way through the economic block and then out again. It is also often thought that it is best to have a sensitive question last in the block of questions, as some persons may refuse not only that question but also all questions from that point onwards. However, as the last question (often on the questionnaire) the question does face the danger of being forgotten, missed or ignored. For example, in Australia the question on total cash income was placed before the economic block of questions and at the top of a page, presumably, to emphasize the importance given to the question. Countries should test the position and sensitivity of the question in their situation, but it does generally seem logical to have the question at the end of the economic block of questions.

L.   Formulating questions on selected topics

110.   As indicated previously, it is important that prior experience from labour force surveys (if conducted) be used in formulating questions on economic activity and related characteristics. To facilitate comparability, it is recommended that similar definitions, concepts, reference periods, and so on, be used and that, subject to the limitations of space on the census questionnaire, the same wording be used in the census questionnaire as in the labour force survey. In practice, identical wording cannot be used, since the census questionnaire usually has fewer questions than the labour force survey and in some countries the mode of collection differs (for example, paper and computer-assisted methods).

Census field staff are also less experienced than those used in the labour force surveys and have less training, and therefore the wording of the questions may need to be simpler (and will be perhaps less precise in measuring economic activity concepts as a result).

111. The act of measurement in a census in a face-to-face interview involves two persons, the enumerator and the respondent, with the questionnaire serving as a tool to facilitate communication between them and record the outcome. The level of understanding of the questions by the respondent will depend upon that individual's level of literacy, if not education, as well as on how questions are formulated on the questionnaire and communicated by the enumerator. In a census operation, enumerators are not expected to have a high level of understanding of concepts and definitions underlying the employment variable, primarily due to the limited training they undergo, a function of the large number of enumerators employed in a census and the cost of their training. Thus they have to be trained to communicate the meaning of the questions rather than the underlying concepts that led to their formulation. As indicated in sections B and C of the present chapter, the experience gained from labour force surveys, if any, is extremely useful in questionnaire design and in the training of field staff on economic activity and its characteristics. Comparable questions with the same reference periods are recommended, but it is accepted that it may not always be possible to use identical questions since there are fewer census questions on the topic and the level of training of field staff is lower. Some principles for formulating questions are outlined below.

112. **Avoid technical terms.** Simplicity is always the key to good questions. Technical, long or complex words should be avoided. The terms for concepts should not be confused with the words to be used in the questions. Such words as "economically active", "own account", "self-employed", "occupation", "industry" or "contributing family worker" should definitely be avoided in questions or in response categories given to respondents as they are very unlikely to know what they are intended to mean in that context. As described more fully in part two with respect to individual items, the questions should be put in language that operationalizes the technical concepts, using words and examples that make sense to a broad range of respondents. Such an approach, while always important, takes on added significance in cases where the questions will need to be translated into other languages by enumerators in the field. In practice, the translations of some apparently simple economic questions or pre-coded categories have yielded surprising results.[14] All translations should therefore be carefully checked during pre-testing and during the training of enumerators.

113. **Always keep questions short.** If necessary, a second part to the question or a second question may be used. The use of "or" in questions should be avoided. Such questions usually lead to the assumption that the enumerator will ask another question to find out to which part of the "or" question a "yes" answer applies. If follow-up questions are implied for enumerators, it is best to put them on the questionnaire to ensure they are asked, and asked uniformly. Also to be avoided are negative or vague formulations, such as "You did not do any of these activities in the last seven days", "What did you do?" or "What did you do when not working?". They can be confusing. Further, qualifying clauses at the beginning or end of the question and questions with negative formulations are all easily misunderstood, particularly in an oral interview.

114. **Pre-coded response categories should be exhaustive.** Pre-coded responses are commonly used for most economic characteristics; the usual exceptions are occupation and industry. Pre-coded response categories reduce the time and cost of data entry and processing, but require much testing and experimentation to generate the appropriate categories. If only one response category is to be checked as a response to a question, the categories should also be mutually exclusive and exhaustive. Categories of economic activity status (that is to say, worked, had a job but did not work, seeking work, avail-

14 See, for example, United Nations (1992), para. 227.

able for work and so on) have sometimes been combined with reasons for non-economic activity (such as schooling, home duties and retirement). Those situations are not all mutually exclusive and therefore present special problems for the wording of questions (for example, whether more than one category should be checked or whether respondents are to be directed to focus on the "main" activity). If that style of questioning is adopted, special instructions and training will need to be given to enumerators to ensure that the priority rule is correctly applied.

115.   **Minimize the use of explanatory notes.** Some explanatory notes are almost always necessary in connection with the economic items but they should be kept to a minimum. Only very important points should be put on questionnaires (or self-enumeration forms); otherwise they will clutter the forms, and the enumerator (or respondent, in the case of self-enumeration) will either ignore them or become confused. Notes on questionnaires that are instructions to the enumerator and those that are prompts to the respondent should be clearly identified as either one or the other. Both can be quite helpful. Differing type styles or sizes or colours are often effectively used to distinguish instructive notes from prompts.

116.   **Limit the use of skips in the questionnaire.** An issue related to the sequence of items on the census form is the nature and number of "skips" in the form. Skips (instructions to jump from one census item, n, to another one, not n+1, depending on the response given to item n or an earlier item) are used to direct the flow of the census interview, whether the census is based on the interview or on the self-enumeration approach. Although quite complex skip patterns, along with repeated probes for economic activity, are sometimes used effectively in labour force surveys, as a general rule skips and multiple probes have to be held to a minimum in a census, particularly when self-enumeration is used. To the extent that skips are used in the census, they should be as simple as possible, with clear guidance on the correct path to follow given in the form of arrows, instructions, or both. With respect to the economic items, it is difficult to avoid completely the use of skips, but every effort should be made to keep them as simple as possible. Common skips associated with economic items include age-related filters and the exclusion of those respondents from most or all of the subsequent items in the economic block, as well as respondents for whom no current (or past) economic activity is reported in the census.

117.   **Build on past experience.** Changes in the definition of concepts in the census and the ways in which a census gathers information related to those concepts may lead to major problems of comparability. Such problems may arise between one census and the next and in relation to other sources of data on economic characteristics. While changes in definitions or methods of data gathering can be introduced, those changes—whether intended to improve data quality or save costs—should be the result of a careful decision-making process. In the design of the questionnaire and the tabulations, steps should be taken, to the extent possible, to maintain some degree of comparability, for example, by making it possible to identify separately those elements that represent modifications to previous concepts or scopes.

118.   In most instances, such efforts to maintain comparability would necessitate additional questions on the census questionnaire, but the counter-balancing reward is that it is possible both to measure change over time since the last data collection and to try out improved methodologies. When using such an approach, however, it is important to recognize that it is likely to be most effective when only a limited number of core issues are treated in this way (see United Nations, 1988, chap. II, stage 5, and chap. III, sect. B). For example, given the changes in the production boundary and the implied changes in the definition of economic activity, a building-block approach would suggest keeping any past question(s) on activity status and introducing a question that focuses on the other activities in the household sector that tend to be overlooked by respondents or

enumerators. Tables on changes between the censuses should then be designed to show the differences in the measures, by presenting separately the statistics based on the old set of questions and those with the extended set. The additional question can, however, be costly. It should be examined in the pre-test and assessed in relation to other questions on economic characteristics that need to be included.

119.   **Be realistic and keep the limitations of the census in perspective.** Decisions to include or exclude a particular (economic) topic in a census and the degree of refinement that can be attempted in its measurement should be based on experience and careful testing. First, an item that has been difficult to measure in a sample household survey with well-trained enumerators and a sound sample design should not be included in a census. Second, the census should attempt to obtain only "rounded" or approximate measures of those variables that are measured more precisely by household surveys. It is not recommended in a census to attempt to measure with utmost precision a particular category if doing so requires complex questions or sequences of questions. Conversely, within its limitations, a census should be made to yield what a sample survey will not provide, that is, statistics at the lowest level of administrative, planning or political unit and on the relevant social groups of a country's population, for major economic topics (for example, the number of carpenters or nurses in a town or a local government area, by age group and sex and type of training). The census must not fail in its unique function of gathering data at the lowest level, for major economic characteristics. Attempts to provide statistics with high levels of conceptual precision for some of the more difficult-to-measure topics (underemployment, employment in the informal sector or many of the national accounts concepts and distinctions) can have negative effects on the overall quality of the basic census results. Such conceptually precise estimates should be obtained through specialized surveys. When required also for small areas and groups the "cruder" census results may provide a good starting point for targeting such surveys, as presented in part six.

Part Two

Measurement of economic characteristics in a population census

Chapter III
Measurement frameworks

A.   Understanding the concept of economic activity

120.   The single most important issue in obtaining reliable information on economic topics in a population census[1] is to accurately determine the activity status of the population: who is economically active and who is not? Statistics on the size and composition of the two groups are fundamental to formulating almost all economic and social policies and related planning and research. Moreover, most economic characteristics in the census are obtained only for the economically active; therefore the item(s) used to identify activity status are the key to the whole block of economic items. Omissions and misclassifications of persons can have a major impact on a broad range of census data pertaining to the socio-economic situation of the population.

121.   At present, the conceptual basis for distinguishing the economically active in the population is derived from the resolution concerning statistics of the economically active population, employment, unemployment and underemployment adopted in 1982 by the Thirteenth International Conference of Labour Statisticians.[2] That basis is tied to and fully compatible with the System of National Accounts (see Commission of the European Communities and others, 1993, chaps. VI and XVII). The resolution defines the "economically active population" as comprising "all persons of either sex who furnish the supply of labour for the production of economic goods and services as defined by the United Nations systems of national accounts and balances during a specified time-reference period".[3]

122.   The definition of the economically active population has evolved over the years. In 1966 it included "all persons of either sex who furnish the supply of labour for the production of economic goods and services" (see United Nations, 1967). The 1982 resolution (ILO, 1983) enlarged the concept and linked it to the SNA. As the scope of production within the SNA has changed, so has the scope of the concept of economic activity. The present situation, as of 2007, is reflected in an article published in the *Bulletin of Labour Statistics* (Hussmanns, 2007). A basic theme running through the revisions has been to obtain a more precise representation of the economic activity of the population, both in their contribution to production by way of their labour and in the creation of the product itself. However, changes in concepts have not automatically been translated into procedures that could improve the measurement or coverage of the economically active population correspondingly.

B.   The scope of production in the System of National Accounts

123.   The 1993 System of National Accounts includes all production of goods for own use within its production boundary, as goods can be switched between market and non-market use even after they have been produced. However, it excludes all[4] production of services for own final consumption within households, such as the production of personal and domestic services by members of households for their own final consumption (Commission of the European Communities and others, 1993, chap. I, sect. E) (see figure I). For consistency between the 1982 resolution on economic activity concepts and

Figure I

Economic production as defined in the current System of National Accounts
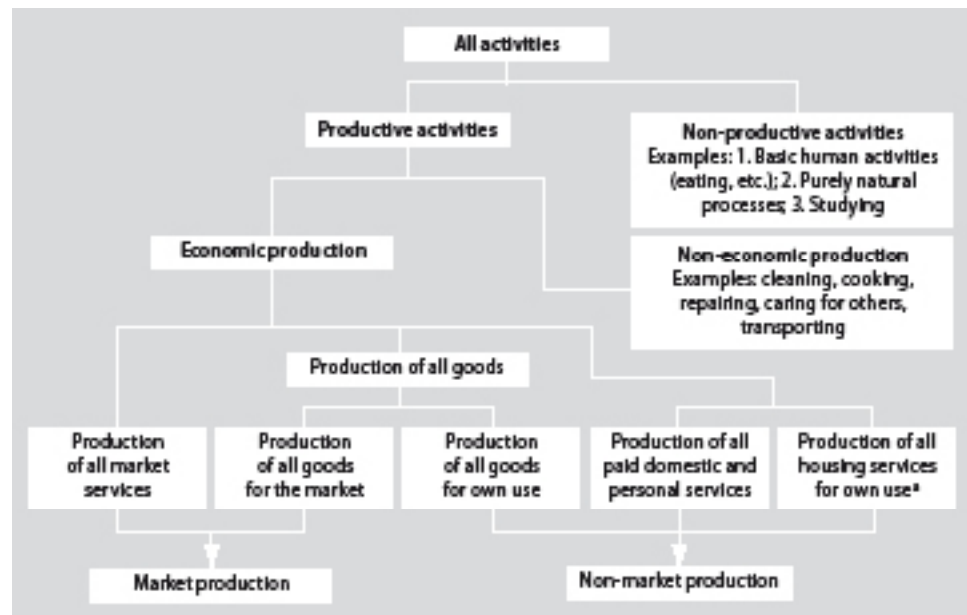
the SNA, the types of production that constitute the production boundary of the SNA should be clearly delineated. According to the 1993 SNA, the types of production such as those presented in table 1, which are for households' own consumption, are within the production boundary.

124.   Some experts object to the classification of certain activities, such as fetching water and gathering firewood, as economic activities. The contention is that those activities, like many others routinely carried out by women in particular, such as the processing of agricultural products for household consumption, or dressmaking for members of the household, would inflate the numbers of the economically active in a manner that would make the resulting statistics less useful for certain types of analyses or economic policies. Maintaining comparability between measures of economic activity that are consistent with the wider production boundary of the 1993 SNA and earlier data sets is also a concern. To address that concern, a building block approach is recommended. By using questions that make it possible to provide statistics according to the older formulation of the concept, it is possible to maintain consistency with the past. At the same time, new questions can be introduced, making it possible to cover the more comprehensive set of activities included with the newer concept.

125.   The treatment of volunteers (those working without pay or profit outside the household) has also been the subject of discussion (see for example, United Nations, 2003). Under the 1993 SNA, the production of goods by volunteers and the production of services by volunteers for market enterprises or for non-profit enterprises operating in the market (such as some hospitals and schools) are within the production boundary and are therefore treated as economic activity. The production of services by volunteers for households or for non-profit enterprises not operating in the market is not treated as economic activity. Therefore, volunteer work caring for neighbours' children or providing services for a religious group or a sporting body would not be considered as economic activity, whereas voluntary work to construct a bus shelter for the community or to provide services to a private (fee-paying) school would be treated as economic activity.

Table 1
Examples of non-market household production activities

Primary production	Processing primary products	Other production of goods	Fixed capital formation
Growing or gathering field crops, fruit and vegetables	Threshing and milling grain	Storing crops	Construction of dwellings
Producing eggs, milk and food	Making butter, ghee and cheese	Dressmaking and tailoring	Construction of farm buildings
Hunting animals and birds	Slaughtering livestock	Handicrafts made from non-primary products	Building boats and canoes
Catching fish, crabs and shellfish	Curing hides and skins		Clearing land for cultivation
Cutting firewood and building poles	Preserving meat and skins		Major repair and maintenance of dwelling and farm buildings
Collecting thatching and weaving materials	Making beer, wine and spirits		
Burning charcoal	Crushing oil seeds		
Mining salt	Weaving baskets and mats		
Cutting peat	Making clay pots and plates		
Fetching water	Weaving textiles		
Collecting firewood	Making furniture		

126.   For easy reference, table 2 lays out a summary of the major activities that are regarded as economic and non-economic. The emphasis in the table is on activities commonly found in many developing countries, although not every situation would be considered to be important in all of them, nor can a general table address every situation relevant for any individual country. On the one hand, the recommended treatments for some specific activities shown in the table may be inappropriate in a given census owing to operational considerations. On the other hand, rather than exclude some activities to satisfy certain user interests, it would be preferable to cover those activities and identify them separately, and exclude them from tabulations tailored specifically to those interests. Census planners should try to draw up a similar table specifically for common activities in their country to guide questionnaire development for the economic questions. In some countries, pictures or drawings depicting the various activities are also used. Discussions of the final list of economic activities for the census should include national accounts experts as well as other major users of census results related to the economic characteristics of the population. In those discussions it is important to remember (*a*) that inclusion of those activities as "economic" will lead to an increase in the number of employed persons only if the persons concerned do not carry out any other economic activities; and (*b*) that the census questionnaire can be designed to ensure that such persons can be identified separately and therefore be excluded for certain tabulations if that corresponds better to the needs of certain users.

127.   The proper identification of the activities that are economic is vital for all subsequent data that are collected with respect to economic activities. It is therefore important to pay particular attention to the design of the census form, the training of interviewers or the instructions given on the form, to ensure that the respondents correctly understand the questions translating the above notions. It is particularly important in economies in which a substantial proportion of the economically active population is

Table 2
Illustrative set of economic and non-economic activities

Economic activities	Non-economic activities
▪ Working in wage jobs	▪ Studying full-time
◆ Full-time or part-time	▪ All types of housework, including the following:
◆ Permanent or temporary	◆ Unpaid child minding own or other children
◆ Casual or piecework	◆ Education/training of own children at home
◆ Including paid child minding and other paid domestic work	◆ Housecleaning and decorating exclusively for own household
◆ Can be paid in cash or kind (e.g., food or accommodation)	◆ Cooking/preparing meals for own household
▪ Having business activities	◆ Caring for the sick and aged (unpaid)
◆ Large or small, agricultural or non-agricultural	◆ Repairs (minor) to own dwelling, etc.
◆ Small shop/kiosk/street stall	◆ Repair of own dwelling equipment and vehicles
◆ Preparation/selling of juice, soft drinks	▪ Begging
◆ Taxi operator	▪ Other types: persons doing no economic activity due to the following:
◆ Shoe cleaning/sewing business	◆ Retirement
▪ Any activities on own or family farms for the purpose of production for sale or for home consumption including the following:	◆ Sickness
◆ Planting crops	◆ Disability
◆ Harvesting crops	◆ Living off investment, rental or pension income (no current activity to earn it)
◆ Keeping birds and other pests off crops	
◆ Weeding	
▪ Transport of goods from the fields for storage or for sale	
▪ Fetching water and collecting firewood for domestic use	
▪ Fishing, collecting shells or seaweed for sale or home consumption	
▪ Processing goods for sale or home consumption	
◆ Mats, hats from natural or grown fibres	
◆ Furniture from natural timber	
◆ Butter/cheese and other products from milk	
◆ Oil from oil seeds/fruit	
◆ Preparation of charcoal	
◆ Dressmaking	
▪ House or farm building/construction	
◆ Fence/enclosure/storage construction	
◆ Road/irrigation construction	
◆ House construction/additions	

Source: International Labour Office; based on Commission of the European Communities and others, *System of National Accounts, 1993* (United Nations publication, Sales No. E.94.XVIII.4).

Note: A similar list of activities for each country should be made early in census preparations to assist in questionnaire design and testing and to be used later in manuals and training.

not engaged in full-time, full-year employment but in part-time employment, casual work, home-based work, work paid for largely in kind, unpaid family work and the production of goods for own consumption. In those instances, the use of an activity list could be very helpful.

C.   Correspondence between economic activity and production within the System of National Accounts production boundary

128.   The scope of economic activity and the production boundary as delineated by the System of National Accounts, though closely related, are not necessarily equal in the way they are measured. A component of production that is difficult to capture fully is that of the household sector, comprising, among others, subsistence agriculture, the informal sector[5] and the processing of primary products (see also list in table 1). Although the ILO concepts of economic activity seek in principle to cover all kinds of economic activity in the household sector, including part-time paid and unpaid work of children in small family businesses or farms as well as part-time paid work of students after school or during school holidays, there are practical limits to that ideal. For example, in providing guidance on household production, the SNA states "When the amount of a good produced within households is believed to be quantitatively important in relation to the total supply of that good in a country, its production should be recorded. Otherwise, it is not worthwhile to estimate it in practice."[6] However, in its resolution, the Thirteenth International Conference of Labour Statisticians stated that persons engaged in the production of economic goods and services for own and household consumption should be considered as in self-employment if such production comprises an important contribution to the total consumption of the household.[7]

129.   Similar considerations apply in deciding upon the value of attempting to measure the economic activity of children and early adolescents. Concerns with child labour may, however, mean that their activities would need to be recorded even if the criterion of their contribution being quantitatively important in relation to the supply of that good in the country is not satisfied.

130.   Difficulties encountered in translating the operational definitions into questions that would elicit the desired responses and in applying interviewing and other data collection techniques that would reach hard-to-measure categories of people may also cause the estimates of economic activity to deviate from the production boundary of the SNA.

D.   Decisions related to measuring economic characteristics

131.   The operational definition of the economically active population limits the coverage of production to those performed:

   (*a*)   Within a specified time frame or reference period;

   (*b*)   By individuals meeting an age-specific criterion.

132.   **The reference period.** As explained in *Principles and Recommendations, Revision 2*, two types of reference periods are recommended. The short reference period of a week or a day (though one month has been used in some cases) represents "current economic activity" and corresponds to the labour force concept; the long reference period of one year is referred to as "usual economic activity". The former is the focus of chapter IV, and issues related to measuring "usual activity status" are presented in chapter V.

133.   **Age cut-offs.** For such practical considerations as speeding up operations, limiting cost, and building rapport with respondents, age limits are frequently used to exclude segments of the population that are either too young or too old to be engaged in any economic activity or are at ages at which the level of economic activity is deemed to be

[5] See paras. 406-410 for the definition of informal sector.

[6] Commission of the European Communities (1993), para. 6.25; see also United Nations (2008b), para. 2.238.

[7] Subparagraph 9(6) of International Labour Office (1983). Available from http://www.ilo.org/public/english/bureau/stat/download/res/ecacpop.pdf.

very low. For example, in many countries it makes little sense to ask about the economic activity of infants or very young children in any population census. It is thus usual to specify at least a lower age limit, below which answers are not required even though some level of economic activity may take place among that group. Some countries also use an upper age limit, the maximum age beyond which economic characteristics should not be collected. *Principles and Recommendations, Revision 2*, however, cautions against the use of a maximum age (see United Nations, 2008b, para. 2.241), since valuable information, especially about the activity of the elderly or the ageing population, will then be lost. In many countries, in fact, the previous practice of a recognized retirement age for employees is now being abandoned.

1.   Setting the reference period

1.1.   *Period corresponding to current activity (labour force)*

134.   As previously indicated in respect of consistency with other data sources, countries may wish to arrange for the reference period used in the population census to be comparable with that used in the labour force survey programme (if any).

135.   The currently economically active population, referred to as the labour force concept, has been the basis of employment and unemployment measures in both censuses and household surveys for many years. The reference period for the current activity concept has in most national censuses ranged from a time frame as short as one day to as long as one month, with the most common being one week.[8]

136.   Some countries have been experimenting with one day (for example, the day before the enumeration day) as a reference period in surveys, but it is not known if any country has used it in a census. Such a measure would be the most "current" measure possible, and would minimize problems of recall by the respondent. A study by P. Visaria[9] has shown that, as expected, employment tends to be lower and unemployment higher with a one-day reference period compared to the one-week period, using the priority rule presented in chapter IV (see para. 169). In a population with a high incidence of intermittent work the chances of some individuals having one day without work is higher than their having a complete week without work. Temporary absences from work for the one day can cause coverage problems. Most countries have thought one day as a reference period is too short, and there is always a danger that something unusual may happen on that day, if not nationally, at least in some regions of the country. Avoiding Sundays or other traditional rest days for the country concerned may also present a problem when the one-day reference period is adopted.

137.   Some countries have tried a one-month reference period. However, recall lapse is likely to be present for some respondents, and there will be more problems when formulating questions about different activities during the month than during a shorter period. The usual approach is to determine a main activity if there is more than one employment activity. If both employment and unemployment occur during the month, it is necessary to decide on the approach: whether to use (*a*) "any work" during the month; or (*b*) "main activity" during the month, which is also used in the usual activity approach.

138.   On balance, the seven-day reference period, which has long been used in most countries for both censuses and surveys, seems the most useful one. The recommended reference period for "employment" is the seven days prior to the census enumeration day, which is the closest short period. However, some countries prefer to use the calendar week prior to enumeration day as they find respondents relate better to a fixed calendar period. When using the calendar week approach, the exact starting and ending day should be given, for example, "from Monday to Sunday last week". The use of either reference

8 As indicated in chapter IV, the measurement of unemployment has an additional criterion relating to job search, which includes the reference week but is longer ("last month", "last four weeks").

9 Paper presented at the Economic and Social Commission for Asia and the Pacific/International Labour Organization seminar on employment and unemployment statistics, held in Bangkok in January 1989.

period—the previous seven days or the calendar week—should be included in any pre-testing for the census. It is particularly useful to apply the same period for the census as for the national labour force surveys, for comparability purposes (within reasonable limits). However, if it is deemed necessary for whatever reason to use another reference period, it should definitely be subjected to pre-testing. An evaluation should be made of the effect of the change on comparability with the results from previous censuses, and the results of the evaluation should be published.

139.   The measures are termed "current" measures since the short reference period is close to the census date. Nevertheless, it is unlikely that they or any other census measures will give very current indicators of the employment situation at the time when census results are released, owing to the usual long period for census processing. As noted earlier, household surveys are used to obtain more current, that is to say, up-to-date, estimates for national aggregates and large groups. However, given the recent developments in the use of technology in census operations, such as scanning, in place of manual data entry, the census processing time has been drastically reduced in many instances, in some cases to within a year of the completion of census enumeration. It is also true the basic structure of the economically active population does not usually undergo major changes while census results are pending, and that those results are consequently very valuable indicators of the economic structure, particularly for small geographical areas.

1.2.   Period corresponding to usual activity

140.   The concept of usual activity defines participation in economic activity and relevant characteristics of the job and of the establishment or employer over a long reference period. The usual activity approach, in principle, represents the typical or usual employment patterns over the long period. The commonly used period is a year, although in some instances another period, such as the agricultural season, is used. The latter is particularly useful when associating employment statistics with production statistics. When a year is used, it is necessary to decide if it should be the previous 12 months or a calendar year. Unless census day is early in the year, the latter period has a greater risk of recall and memory losses. It has therefore been found best to use the 12 months ending at the month prior to the census day.

1.3.   The choice between current and usual activity reference periods

141.   Most countries have elected to measure current activity (that is, they have used a short reference period) in their population census. A relatively short reference period, close to the time of the interview, minimizes memory lapses among the respondents. It is also easier to measure the characteristics of economic activity for the short reference period, as fewer changes usually occur during that period than during a longer one. The disadvantage is that the short reference period can represent the activity situation only during the "current period", and it may not be representative or typical of the annual or average activity patterns. If the country or region has striking seasonal patterns of employment and unemployment and of economic activities, users might prefer the usual activity approach to get a reflection of activities over a long period, for example, 12 months. However, the retrospective data collected over the long reference period are subject to more recall errors than the data on the currently active population (Hussmans, Mehran and Verma, 1990, chap. IV, sect. 1).

142.   In addition, to truly represent typical employment situations over the entire reference period, all economic activities, including the main, secondary and any additional activities, should in principle be covered. Over a period as long as one year, the labour market and therefore the population could be subject to changes: the activity status of individuals may

change between active and not active and between employed and not employed. Thus, to obtain accurate information about usual activity, the approach would require several questions about those activities, far beyond the minimum set that would be required for the current activity approach. For example, at a minimum, respondents should be asked to state the amount of time they spent, measured either in weeks or months, in the employment, unemployment and non-economic activity states. Otherwise, respondents would be left with the task of accurately determining what their main situation or activity was during the reference period. Collecting those data would prove a major challenge for the design of the census instrument and would increase the cost of the census.

143.   Over a long period such as one year, changes could occur in occupations and other such characteristics of the job or jobs. Those changes would invoke special problems, such as the greater effort required to collect information on time spent in a possibly greater number of different jobs in order to objectively determine the main job for the period.

144.   The short reference period, usually seven days, has been used in most countries for censuses and surveys for a considerable period of time. Consequently, there is accumulated experience in handling the limited difficulties that occur. The relative precision of the current activity measure and general ease of interpretation for enumerators and respondents makes it highly preferable to any usual activity measure for questions on economic characteristics. Census planners therefore need to weigh the advantages of the current activity measure against the more difficult-to-measure item of usual activity, which nonetheless can reflect activities over the different seasons of a 12-month period. It is strongly recommended that any country starting to use usual activity for all their economic characteristic questions conduct a very extensive pre-testing programme to identify the difficulties and build experience in applying the concepts.

145.   As comparability with previous results is also of major value for census data, those considering adopting the usual activity approach for the first time should keep the current approach. They should carefully assess the extent to which questions applied to the usual approach should be included, and vice versa.

146.   Using both approaches has obvious advantages but also some drawbacks: an obvious disadvantage is that every additional question adds to the costs of the census and that the two approaches will generate different numbers. In addition, a number of difficult technical problems arise in using the longer reference period of the usual activity measures, which are explained below.

147.   Statistics on the economically active population should always be published with the reference period prominently noted so that users are fully aware of it: for example, "employed population in the week of 6-12 January 2008". Although that information has often been omitted in the past, it is important to include it if statistics in countries with pronounced seasonal variations in certain types of economic activity in certain regions, for example, agriculture or tourism, are to be understood in their proper contexts.

2.   Setting age limits

148.   Items under the various topics covered by the census will usually have different age cut-offs so that one of the first matters to be decided is the appropriate age limit(s) to apply for each population-specific item in the census. The rationale for setting an age cut-off is to define a more realistic target population to whom the economic activity questions will be addressed. It is usual to specify at least a lower age limit below which answers are not required for such questions. The same age cut-off should be applied to all items dealing with economic characteristics and should preferably also be the same age cut-off as that used in the labour force surveys, if any. For many countries, the cut-off used in asking questions on economic activity in a population census has been 14 or 15 years of age.

In countries where a large proportion of the population enters and completes secondary school, there is a tendency to adopt age 16 as the lower limit. In other countries, lower cut-offs, below age 10 and sometimes as low as age 6, are used (United Nations, 1996, table 4) for a variety of reasons, including inadequate funding for education and the consequent downward trend in school attendance, or an increased interest in measuring the extent and types of child labour. Developing countries, in particular, use such lower cut-offs.

149.   The following main factors should be considered in determining an appropriate lower age limit in asking economic questions in a census:

    (*a*)   The age at which significant numbers of the population begin to engage in or seek to be engaged in economic activity, often a function of income levels and school enrolment rates, and the distribution of school leaving ages;

    (*b*)   The objectives and intended uses of the data.

150.   In some countries consideration of the lower age limit is confused with the minimum legal age for paid employment. Such considerations may be avoided, as in principle there may be no connection between the minimum age for census questions on economic activity and the legal age for paid work. These questions should ideally record any type of economic activity, including part-time work and unpaid work in small family businesses or on farms, and interest in the extent and type of work activities of children has been considerable in recent years. In most countries some school-age children engage in part-time, paid or unpaid employment, for example part-time work after school or during school holidays. In rural areas children often help in agriculture at all ages whether going to school or not, and those activities are not usually covered by any minimum age legislation. In a population census, the aim should be to record the whole range of economic activities; setting a lower age limit that is too high could lead to the loss of valuable information.

151.   In some countries an upper age limit is also used, for example, 65 years of age. That procedure is even more likely to lead to the loss of valuable information. In almost every country some people work, either full or part-time, well beyond any legal or conventional retirement age. This is particularly so in rural societies but is not uncommon in urban areas, even if older persons often work fewer hours than when they were younger. Many Governments are now considering the need to encourage some form of work at higher ages, as active older people are often healthier than before. In almost all countries the population is living longer, which suggests that a greater number and range of economic activities may be missed if an upper age limit is imposed for such questions.

152.   The choice of a lower age limit to use for the economic activity items should be explicitly reviewed before each census, although usually the same cut-off will be used as in the previous census. In deciding whether to raise or lower the age cut-off used previously or to continue using the same cut-off, there is a trade-off: the benefits gained by asking a broader group of people the economic items must be balanced against the associated field and processing costs and added proxy respondent and enumerator burden. The lower the age cut-off the larger the population base to be interviewed (or of whom the information is to be sought). That increases the interview time and therefore the average cost per household. The lower age cut-off will generally give more realistic coverage of the economically active population in countries where there is either low enrolment in or a high dropout rate from primary school. However, if there are legal repercussions for children's non-attendance in primary school, the quality of responses may be poor, and therefore the benefits in coverage from increasing the population base will be diminished and will be outweighed by the cost. Another factor that could affect the outcome of the interview is the difficulty of eliciting accurate responses with the questions that have typically been used in censuses, for example, when the activities involved are informal in nature: without contracts, with irregular hours of work and so on. Therefore, if an objective is to measure child labour,

either the questions should be modified for the whole census population or it should be done on a sample basis in a census long form. Alternatively, it could be left for another more specialized survey specific to that objective, such as a time-use survey. Such modifications may also be warranted for a number of other reasons, some of which are considered below.

153.   In countries where child labour has become a sensitive issue, the overall effect on the quality of the census results obtained when using a low age limit should be taken into account and studied as part of the pre-census testing programme. In general, a low age limit for the economic questions will impose an extra burden on respondents, usually proxy respondents who have to report on several other members of the household, and on the data processing, but will provide valuable, otherwise unknown information. More relevant information about the form and circumstances of child labour and other difficult-to-measure production activities may, however, be obtained from specially designed surveys, including special purpose child labour surveys, labour force surveys and perhaps time-use surveys. A discussion of the merits and modalities of such methods is beyond the scope of the present *Handbook*.[10] In general, it would be an advantage if the use of a new age cut-off could be introduced first in ongoing sample surveys and subsequently in the census on the basis of the experience gained. The cost-effectiveness of reducing the age limit should in any case be assessed at the pre-testing stage.

154.   It is generally preferable to introduce age-based cut-offs, such as raising the lower age limit or setting an upper age limit, during tabulation and analysis, if the data will be used to meet particular policy, research, or administrative needs; for example, to produce tables comparable with those from previous censuses or to estimate the impact on certain totals of the use of a higher minimum age for asking the economic items.

E.   Organization of chapters IV to VIII

1.   Relevance of past experiences

155.   There is often much to benefit from by building on past experience, whether it is the country's own experience or that of others, as opposed to creating a new questionnaire formulation or design from scratch. The starting point for revising old questions or introducing new ones is therefore the lessons from previous censuses, and sometimes even from surveys. The next five chapters present examples of a range of questions that have been used in national censuses for each of the items on economic characteristics, as proposed in *Principles and Recommendations, Revision 2*. The strengths and limitations of the questions are identified, with some suggestions for improving their effectiveness.

156.   The examples have largely been drawn from both developed and developing regions around the world, except when no national experience was available to draw upon, in which case untested examples captioned as "hypothetical questions" are proposed. For each item, the explanations also address issues related to measurement, field testing, and data processing. The starting point for the material presented in the chapters is the set of broad topics related to the economic characteristics of the population proposed in the United Nations *Principles and Recommendations for Population and Housing Censuses, Revision 2*. Those topics are covered in the following chapters: current activity status (chap. IV), usual activity status (chap. V), descriptive characteristics of the job (chap. VI), descriptive characteristics of the establishment (chap. VII) and working time and income (chap. VIII).

157.   In general, topics and items for which national experience abounds include employment and unemployment, status in employment, occupation and industry; and, to a lesser degree, place of work and working time. National experience on the measurement of income and usual activity status in population censuses[11] is somewhat

10 For guidance on time-use surveys, see *Guide to Producing Statistics on Time Use: Measuring Paid and Unpaid Work* (United Nations publication, Sales No. E.04.XVII.7).

11 Based on a review of 176 national census questionnaires from the 2000 round of population censuses, undertaken by the United Nations Statistics Division in 2007. For a summary of national census experience covering the period 1965 to 1984, see United Nations (1996), tables 3 and 4. See also *Sources and Methods: Labour Statistics*, vol. 5: *Total and Economically Active Population, Employment and Unemployment (Population Censuses), Third Edition* (Geneva, International Labour Office, n.d.).

limited, and even more so with respect to the measurement of institutional sector, the informal sector and informal employment.

158.   The aim in the presentation of examples has been to select questions that illustrate particular points. For all topics on which limited national experience is available, as well as for any introduction of new or modified questions, an extensive programme of pre-tests and pilots needs to be considered, even if there is some household survey experience in gathering corresponding information.

159.   Several census questionnaires illustrating a range of national practices in formulating questions about economic characteristics in a population census are presented as an annex. Elements of those examples are referred to in the text wherever possible and relevant.

2.   Format of the presentation

160.   Chapters IV to VIII in parts two and three focus on preparing questions for inclusion on the census form. The principles for developing questions are generally the same regardless of the characteristic being examined. However, each characteristic has some specific considerations that influence the level of difficulty in obtaining accurate measures. Therefore, some issues are presented in greater detail than others, depending on the characteristic.

161.   The characteristics are discussed in five chapters corresponding to the unit to which the information sought pertains. They are as follows:

Part two

Chapter IV	Currently economically active population (labour force)
	• Employed population
	• Unemployed population
Chapter V	Usual economic activity status

Part three

Chapter VI	Descriptive characteristics for the main job
	• Status in employment
	• Occupation
	• Place of work
Chapter VII	Descriptive characteristics of the establishment
	• Industry
	• Institutional sector
	• Informal sector and informal employment
Chapter VIII	Working time and income
	• Working time
	• Income

162.   At the beginning of each broad topic, an introduction presents the general issues pertaining to the topic, following which the issues specific to each of the characteristics are covered in turn. The presentation follows the structure shown below and is explained in the remainder of this section:

- Operational definition
- Requisites for applying the definition
- Review of questions used in national censuses
- Suggestions for enhancing the effectiveness of the questions
- Issues of measurement

2.1. Operational definition

163.   The internationally recommended definition for the characteristic is presented. This is generally a restatement of the definition presented in *Principles and Recommendations, Revision 2*. Where necessary additional information is drawn from other sources, such as the ILO manual on surveys of the economically active population (Hussmanns, Mehran and Verma, 1990).

2.2. Requisites for applying the definition

164.   The specific elements of the definition and their implications for translating the concept into questions are examined. They may include the setting of criteria, the identification of keywords that require special attention and peculiarities that could affect the measurement of the characteristic.

2.3. Review of questions used in national censuses

165.   This section presents examples of questions that have been used in previous national population censuses and reviews them with respect to how well they could be expected to measure the characteristic. The names of the countries have, however, been suppressed since the examples are used purely for illustrative purposes. Wherever possible several solutions are presented and reviewed. The commentary explains aspects of the question that either enhance or limit the potential for achieving the intended result. Note that the remarks on the national examples will assume that the English version of the question(s) represents an accurate translation from the original. However, that may not necessarily be correct, and the translation may have introduced some limitations that the original formulation did not have.

2.4. Suggestions for enhancing the effectiveness of the questions

166.   In the light of the review in the previous subsection, suggestions may be made for improving the formulation. When reasonably good and well-tested questions are not available, in some cases owing to limited experience in the measurement of the characteristic in a census, the suggestion is in the form of a proposed question, often untested in a population census exercise, but based on experiences from a sample survey. It has been necessary to include such hypothetical questions in the manual to address topics that were included in *Principles and Recommendations, Revision 2* for the first time, such as informal employment, or emerging issues for which there is no prior census experience. The approach is used to make more options available to users of the *Handbook* and to encourage experimentation and pre-testing in census pilots.

2.5. Issues of measurement

167.   To complement the suggestions given for the formulation of relevant questions, methods for ensuring greater success in fielding the questions are also identified. Complementary tools such as prompt cards, areas to be emphasized in accompanying interviewer and supervisor manuals, training and supervision and the like are also presented, where they are deemed particularly critical.

Chapter IV

Currently economically active population (labour force)

A. Labour force framework

168.　The labour force comprises all persons who, during the reference period, were either employed or unemployed as described below. Figure II illustrates the relationship between the total census population (the total population net of any groups excluded from the scope of the census, such as foreign diplomats), the labour force, the employed population, the unemployed population and the population not in the labour force. Thus the questions in the census questionnaire for determining the labour force should be directed at identifying those categories in a way that is clear, unambiguous and mutually exclusive.

169.　To achieve the preceding objective, some priority rules of assignment are required. The basic rules for the economic activity questions are as follows:

　　(*a*)　That within the reference period, economic activities take precedence over non-economic activities;

　　(*b*)　That within that same period, within economic activities the status of being employed takes precedence over the status of being unemployed.

Thus priority is assigned to participation over non-participation and to employment over unemployment.

170.　It follows that in the measurement of economic characteristics a great deal of effort needs to be put into making the definition operational and adopting methods that will correctly classify those that are economically active. Provided the set criteria are met, a person should be classified as having been economically active during the reference period, whether or not they were also engaged in non-economic activities such as studying or home duties. It should be emphasized that school attendance is to be recorded in a census independently of other characteristics such as economic activity. Persons are considered for recording in the economically active population provided the age of the individual falls within the applicable age limits, irrespective of school attendance. Furthermore, if a person either worked for the set minimum number of hours or had a job (even though the person did not work) he or she should be classified as employed even if seeking another type of employment, for whatever reason.

171.　Students and all other persons who are within the ages for reporting economic activity should be subject to exactly the same economic questions as others. School attendance or attendance at other training institutions should be recorded in the education and training block of questions on the census questionnaire. It is important to remember that "students" can be employed either part-time or full-time, particularly during vacations. In some countries, they form a significant proportion of the employed population. Similarly, students can be actively seeking work and be currently available for work that is either full-time or part-time. Sometimes they are studying only because there is no work available, or they are seeking part-time work to help pay for their studies. However, some users may not want full-time students included, for example, in

tabulations of the unemployed, in which case the responses to the education questions can be used to screen out full-time students during tabulations.

172.   It should be noted that as the labour force comprises the employed and the unemployed, the approach to its measurement is to identify each of the two components. The population not currently economically active is then determined as the residual of the total population after separating out the population that is identified as either "employed" or "unemployed" in the reference period. The rest of the chapter follows that order, first addressing the measurement of the employed population, then that of the unemployed population and finally that of the population not currently economically active.

B.   Employed population

1.   Employed: operational definition

173.   An employed member of the economically active population is, broadly speaking, a person who, during the reference period, performs any activity to produce goods or services of the type that falls within the economic production boundary defined by the SNA, or who is temporarily absent from an activity of this type. The "employed population" comprises all persons above a specified age who, during a short reference period of either one week or one day (*Principles and Recommendations, Revision 2*, para. 2.253) met the following requirements:

(*a*)   They performed some work for pay, profit or family gain, in cash or in kind; or

(*b*)   They were temporarily absent from a job in which they had already worked and to which they maintained a formal attachment or from a self-employment activity such as a business enterprise, a farm or a service undertaking.

174.   The operational definition requires specification of the minimum number of hours that an individual should have worked during the reference period to be considered employed. A one-hour criterion is often used to ensure greater correspondence between total employment and measured production as well as coverage of all types of employment, in particular irregular employment, such as standby work, casual labour, short-term work and the like (Hussmanns, Mehran and Verma, 1990, chap. V, sect. 2). Other criteria that have been used include (*a*) the majority criterion—working most of the time during the reference period; (*b*) the one-day criterion—working at least one day during the reference period; and (*c*) a specified number of hours higher than one.

175.   The one-hour criterion is clear and unambiguous. There are several interrelated reasons (see Hussmanns, 2007) for the use of the one-hour criterion in the international definition of employment. One is to make the definition as broad as possible, to cover all types of employment that may exist in a given country, including short-term and part-time work, casual and temporary employment, standby work, employment in the informal sector and other types of informal employment. Another reason is to ensure that, at an aggregate level, total labour inputs correspond to total production. Such correspondence is needed to make possible a joint analysis of employment and production statistics. Last but not least, the one-hour criterion results from the priority rules of the labour force framework, which gives precedence to any employment activity over any other activity, and which defines unemployment as a situation of total lack of work (zero hours of work during the reference period). In that framework, situations of partial lack of work are covered by the definition of time-related underemployment,[1] that is, as a subcategory of employment. As the definitions of employment and unemployment are interrelated in the labour force framework, raising the minimum number of hours worked in the definition of employment would have the effect that unemployment would no longer refer only to situations of total lack of work.

[1] See "Resolution concerning the measurement of underemployment and inadequate employment situations", Sixteenth International Conference of Labour Statisticians, October 1998, in International Labour Office (2000).

Figure II
Labour force framework

Figure flow chart:

Total population
— Population above specified age
— — Worked during short reference period (1 week/day)
— — Did not work during short reference period (1 week/day)
— — — Had a job/enterprise from which temporarily absent
— — — Did not have a job/enterprise
— — — — Available to work
— — — — Not available to work
— — — — — Looking for work
— — — — — Not looking for work
— — — — — — (Relaxed definition)
— Employed population
— Unemployed population
— Labour force
— Population below specified age
— Not currently active population
— Population not in labour force

Source: ILO. Based on R. Hussmanns, F. Mehran and V. Verma, *Surveys of the Economically Active Populations, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods* (Geneva, International Labour Office, 1990), fig. 2.

176.   The one-hour criterion was reviewed by the Fourteenth International Conference of Labour Statisticians in 1987. While agreeing to retain the criterion, the Conference emphasized that the resulting employment data should be classified by the number of hours of work. Such additional information permits a distinction to be made between different sub-groups of employed persons and assists in making a sound interpretation of the statistics. For particular applications, some users may wish to exclude those working less than a certain number of hours.

2.   Employed: requisites for applying the definition

177.   A person is taken as employed if he or she does any work at all in the reference period of one day or week. Any work is commonly taken as lasting for a minimum of one hour during the reference period. Although some are concerned about the use of one hour, it is essential for (*a*) giving employment priority over all other activities; (*b*) consistency between employment statistics and national accounts production statistics; and (*c*) capturing part-time work as well as marginal forms of employment. Moreover, it has generally been found that only a small number of persons are working very few hours (for example, less than five).

178.   The aim is always to include as "employed" those persons who are temporarily absent from their jobs. The following persons who were temporarily not at work during the reference period should be included as employed: paid employees who continue to have a "formal job attachment" and persons with an enterprise that continues to exist. However, the distinctions are not very clear-cut. The details have been the subject of much discussion,[2] and a fine distinction has been made, particularly in developed countries. In developing countries and particularly for censuses, trying to draw such fine distinctions could be confusing, and it is suggested that only the following rules be applied when defining those who are "temporarily absent" from employment:

2 See http://www.ilo.org/ public/english/bureau/ stat/download/guidelines/ exleave.pdf for the relevant international guidelines.

(*a*)   Paid employees who are:

   (i)   Absent on paid leave (recreation/annual/sick/maternity leave) or paid training;

   (ii) Absent from work (with or without pay) as a result of industrial disputes, a reduction in economic activities or a temporary disruption of work due to machinery breakdowns, a shortage of materials and the like, but continue to have a formal relationship to their job or feel that they definitely have a job to go back to;

(*b*)   Persons with an enterprise who respond that they are temporarily absent from work but that their enterprise continues in their absence since, for example, other members of the family are continuing to work on the family farm or business or orders are still being received, or the activity is functioning in some other way that indicates that it is still operational.

179.   For persons engaged in seasonal activities and contributing family workers, the correct treatment of the concept of "temporary absence" is as follows:

3 In some countries, persons who are seasonally away from work but receiving less than 50 per cent of their wages or salary are considered to be not employed, even though they are receiving that income.

(*a*)   During the off-season, seasonal employees, especially in agriculture, should not be included as employed unless the employer continues to pay their wages or salary;[3]

(*b*)   Self-employed persons engaged in seasonal activities should not be included as employed during the off-season unless their enterprises continue to exist and their owners continue to work in them;

(*c*)   Contributing family workers do not have an enterprise and so cannot be included as employed when temporarily absent from work.

180.   In general, a common-sense understanding of temporary absence has usually managed to cover most cases successfully. Detailed discussions of unusual cases have often confused interviewers and produced excessive numbers of persons being temporarily absent. That situation should be carefully examined in pre-tests and re-interviews to ensure proper recording.

3.   Identification of the employed: special cases

181.   There are numerous categories of persons in all countries who present special measurement issues: for example, persons engaged in several different types of activities during the reference period, persons who have a job but are temporarily laid off or persons seeking employment for the first time. Some categories of employment present special problems of identification. People in those groups frequently may not, without prompting of one kind or another, report themselves or be reported as having undertaken an activity that will lead them to be classified as being employed or having a job. The following are some of the major groups that, without enumerators' probing or taking special care in formulating the questions, will tend to be under-enumerated.

(*a*)   **Persons in part-time or casual work.** Part-time work, casual work, piece work or similar activities are often underreported in many censuses for several reasons, including the misconception that the Government is interested in recording only full-time jobs for the census and suspicions about the purpose of the question, particularly when no taxes are being paid on income. When such work is performed by persons who are not the principal breadwinner, the jobs are not considered to be "real";

(*b*)   **Persons with small businesses, particularly informal businesses.** Small-scale informal activities such as street selling, shoe cleaning, beer brewing and the wide range of informal businesses common in some countries are often not regarded as real jobs or real work and are not systematically recorded;

(*c*)   **Unpaid work in a family business (non agricultural).** Persons, particularly women, working without pay in family shops, small family manufacturing activities or similar family-based enterprises are one of the most difficult groups to record as employed in censuses (and even in sample surveys), since those activities are seen to be part of family responsibilities or obligations rather than as "economic";

(*d*)   **All types of agricultural work, not only small-scale farms.** There is usually little difficulty in obtaining information about employment of those working on commercial or formally registered farms either as owners or as regular employees. However, work on a small common village farm, on a peasant farm or as a casual farm labourer is often more difficult to identify in a census. A deliberate effort would need to be made to improve the coverage of those activities in a census, such as including a module for screening farming households (see chap. XV for an explanation of the use of the population census to develop agricultural frames). Moreover, respondents do not always regard some types of farm work, particularly subsistence farming activities carried out by women, as a real job or real work;

(*e*)   **Persons in collective dwellings who are working.** This group includes prisoners who are remunerated for any work that they do, as well as mine workers and other workers who live in dormitory situations. Care should be taken not to overlook them even though reporting may be difficult;

(*f*)   **Economic activities not yet widely acknowledged.** The following are examples of activities that are considered as economic because they contribute to production as conceptualized by the SNA. They are, however, yet to be accepted by many of the planners of censuses and may not be recognized at all by the respondents or even enumerators:

(i)   **Agricultural work for home consumption.** In many countries, the biggest problem besides identifying unpaid work in family businesses is to identify agricultural work mainly for own consumption. In most of those countries, such work is very significant and is usually done by women who do not participate in any other type of economic activity. They will generally report themselves as homemakers unless the questions or prompts used in the census probe specifically for this type of activity;

(ii)   **Processing primary products for home consumption.** Activities include making mats, pots and furniture, processing milk for butter or oil seeds for oil and weaving textiles, mainly for home use. The SNA states that such activities should be included within the production boundary if they represent a significant proportion of the total production of the relevant item for that country. When households sell part of the products, the activities are sometimes recorded in special surveys, but they will usually be missed;

(iii)   **House and farm building work.** Construction work, not simple repairs or maintenance, that adds to the household's capital is regarded as economic work by the SNA, and the persons engaged are to be recorded as employed when doing this work. Activities include fence making, construction of animal enclosures (for example, for cattle, pigs or poultry), which are sometimes agricultural off-season activities, road building or irrigation ditch making, and home building or major additions to the housing structure;

(iv)   **Voluntary production of goods and voluntary production of services for market-oriented enterprises.** As explained in paragraph 125 of chapter III, such activities should be classified as economic but may be

easily overlooked. In developing countries, voluntary self-help community work to construct schools, bus shelters, bridges and the like is common but may not be reported since there is no payment for the work undertaken. Even in countries with more advanced economies, the efforts by volunteers in fee-paying hospitals and schools may be overlooked. Note, however, that volunteers producing services for households and non-profit institutions should not be included in this group;

(v)   **Collecting firewood and fetching water.** These two activities have only recently (since 1993) been included within the boundary of the SNA and are therefore economic activities, even though there is some reluctance to accept them as such. In many villages and some urban areas, fetching water is a daily activity of women and often children. Apart from the difficulty in perceiving the activities as economic, the effect of including them would in some countries make almost all women economically active and employed, with a corresponding reduction in estimates of the not economically active population and in many cases in unemployment estimates. That outcome can be worrying for those wishing to use the census results as a basis for formulating employment policies, if persons who are classified as employed solely on the preceding basis cannot be separately identified. However, their identification will be very important for a better understanding of the economic and social circumstances of particular groups and localities.

182.   In general, labour force surveys give special attention to the identification and measurement of the economic activity of such special cases, particularly where they are significant in a country's economy. However, countries may wish to assess whether their inclusion is necessary in their population census.

4.   Employed: review of questions used in national censuses

183.   Questions aiming to establish activity status, that is, whether a person should be classified as employed, unemployed or not economically active, have been formulated in a variety of ways. In some cases the questions have addressed each status separately; in others the statuses have been combined as alternative responses to one question. Focusing on the employed, the examples presented below are classified under the following four headings:

(*a*)   One simple question used to determine if an individual was employed during the reference period (examples A.1, A.2 and A.3);

(*b*)   Two questions used: one to determine if an individual did any work during the reference period, and the second to determine if those who were not at work had a job or business to which they would return (examples A.4 and A.5);

(*c*)   One composite question that includes elements to determine two or more economic characteristics of the individual (example A.6);

(*d*)   Two or more questions, one of which is a basic question on activity status (or employed status), and any additional ones to cover specific categories of workers who are not easy to enumerate (examples A.7 to A.11).

184.   The single question in example A.1 is posed to determine whether the individual was employed during the reference period. The question incorporates the main element of the concept—worked for a minimum time or attached to a job during the reference period. The keyword "job" can be understood in a narrower sense than is intended

Example A.1

A.1.1	⌊_⌋	A.1.1	**Do you have job for which you worked 4 hours or more in the past week (or would have worked if you had not been away due to vacation, illness, pregnancy, or a labor dispute, etc.)? (Fill in space A.1.1)**
			1. Yes → **Go to question X.X**
			2. No → **Go to question Y.Y**

(in other words, as formal, standard or regular wage employment), and it may not be understood that it should also include informal-type employment, unpaid work in a family business or in some cases own-account work. The two parts of the question, the main and the parenthetical, would be better separated. That would make it easier for respondents to know to which question they have answered "yes" or "no". In addition, some ambiguity is introduced with "had not been away", which should be qualified.

185.   It has been noted that there is a potential problem with questions such as "did you have any job or business of any kind in the last …", since they are likely to produce an immediate "no" from respondents who are farmers, petty traders or the like. They may hear the word "job" and assume the question is about wage jobs. Such questions alone without probes or follow-up questions will almost always produce an underestimate of the employed population. Depending on country circumstances (especially in countries with a large informal sector), the word "job" might be avoided in the basic question on this topic. It is recommended that countries use a variety of terms or phrases to cover all types of employment. Some of the major problem groups for identification of the employed were presented in chapter II. Persons engaged in only those kinds of activities may not regard themselves as employed or as having a job without some prompting by the questionnaire and/or the enumerator.

186.   The question in example A.2 is too simple and can easily be responded to with a "yes" or "no" by those familiar with the concept of employed. That group may exclude people who judge their economic activity to be insignificant or of little value. Pertinent information is sacrificed when the question is so short. It would be more effective for the reference period to be stated in the question, which would ensure more consistent use or mention of it. The way the questionnaire has been designed, the reference period is stated in the instructions to the enumerator and may not be systematically used when the question is introduced.

Example A.2

To be asked for person 15 years old and over about work in the last week before census (March 25–April 1)

A.2.1 Are you employed?

X	Yes	→	**Question X.X**
X	No		

A.2.2 Are you unemployed?

X	Yes	→	Do you have a status of unemployed?
			⌷X⌷ Yes
			⌷X⌷ No
X	No	→	**Question Z.Z**

Example A.3

A.3.1   Last week, did the person have a full-time or part-time job of any kind?

- Mark one box only
- A "job" means any type of work including casual, temporary, or part-time work, if it was for one hour or more.
- See page 12 of the census guide for more information.
- Remember to mark box like this: ──

☐ Yes worked for payment or profit
☐ Yes, but absent on holidays, on paid leave, on strike, or temporarily stood down
☐ Yes, unpaid work in a family business → **Go to X.X**
☐ Yes, other unpaid work → **Go to Y.Y**
☐ No, did not have a job → **Go to Y.Y**

187.　In example A.3, respondents are given clear instructions and guidance in completing the questionnaire. The inclusion of the fourth response category ("other unpaid work") will ensure that volunteers are given an obvious choice, but the question sequencing will treat them as not employed even if they are working as volunteers for profit-making enterprises.[4]

188.　The basic economic activity question in A.4.1 is used to determine if any work was done during the reference period. However, when taken together with the clarifications given on what is included or excluded from work, the question appears to be too long. In addition, for an interview-based approach, it should be made clear to whom the explanatory notes are directed (specifically to the interviewer or to the respondent).

189.　The notes serve as prompts on the major problem categories, and the response categories can also be useful prompts. However, if those features are included, the question tends to become more complex, in particular when basic economic activity, subsistence

[4] A number of countries have specific separate questions to identify different categories of volunteer work, but no country seems to treat such persons as employed, even if working for profit-making enterprises or producing goods.

Example A.4

A.4.1   Did … work at any time LAST WEEK, either full-time or part-time?
Work includes part-time work such as delivering papers, or helping without pay in a family business or farm. It also includes active duty in the Armed Forces. Work does NOT include own housework, schoolwork or volunteer work. Subsistence activity includes fishing, growing crops, etc. NOT primarily for commercial purposes.

Read each category and mark (X) the ONE box that applies.

1　☐　Yes, worked full-time or part-time at a job or business AND did NO subsistence activity

2　☐　Yes, worked full-time or part-time at a job or business AND did subsistence activity

3　☐　Yes, did subsistence activity only

4　☐　No (did not work OR did only own housework, schoolwork, or volunteer work)

Skip to X.X

A.4.2   Was … on layoff from a job or business LAST WEEK?

If "No", ask "Was … temporarily absent or on vacation from a job or business last week?"

1　☐　Yes, on layoff

2　☐　Yes, on vacation, temporary illness, labour dispute, etc.

3　☐　No

activity and other activities are combined in one question. Furthermore, the use of the term "subsistence" further complicates the question, as the scope of subsistence activities may not be fully grasped.

190.   One way to improve upon question A.4.1 is to give separate categories for subsistence and non-subsistence work, either in the response categories to be checked in addition to the basic ones or in a separate question.

191.   The second question, A.4.2, ascertains if there is an attachment to a job, and complements the identification of the employed. Without that question the count based solely on A.4.1 (that is, only those who did any work in the last week) would incompletely measure the employed by excluding (most of) those temporarily absent. Both questions have the same reference period. The second question is followed with one on unemployment, so that those responding "no" to both questions are asked questions to establish their unemployed status.

192.   The two questions from the interview schedule in example A.5 are slightly different from those in example A.4. The basic question used to determine if an individual did any work during the reference period is directed at measuring "number of hours worked" (issues specific to that characteristic will be covered below in the appropriate section). In both examples, the need for clarifying what comprises work is demonstrated. The exclusions are inserted as part of the question, and examples of what should be considered work are presented as prompts. In example A.5, the major types of work are ordered to present first the ones least likely to be reported and the standard types at the end of the list, to establish some degree of importance for the former activities. While the list may be adequate for more advanced labour markets, further clarification with activities specific to less organized labour markets will be required.

193.   Combining a basic activity question with the collection of "hours worked" data may have influenced data quality on one of the two topics. Such approaches need to be tested if there is interest in trying them out.

Example A.5

A.5.1   **Last week, how many hours did you work (not including volunteer work, housework, maintenance or repairs for your own family)?**

Include as work:

- *Working without pay in a family farm or business (e.g., assisting in seeding, doing accounts);*
- *Working in your own business, farm or professional practice, alone or in partnership*
- *Working for wages, salary, tips or commissions.*

Number of hours (to the nearest hour)

○   OR

○   None

A.5.2   **Last week were you on temporary layoff or absent from your job or business?**

Mark one circle only

○   No

○   Yes, on temporary layoff from a job to which expect to return

○   Yes, on vacation, ill, on strike, or locked out or absent for other reasons

Example A.6

Type of activity during past 30 days	

A.6.1 During the past 30 days did … work for cash?
 	
1. Yes, for someone else → **Go to X.X**
2. Yes for self
3. No
 	
        } **Go to A.6.2**
	A.6.2 Then what did … do during the past 30 days?
	
	1. Family business
	2. Work at lands/farms/cattle post
	3. Actively seeking work
	4. Housework
	5. Student
	6. Retiree
	     (Other specify)

194.   The questions used here are fairly short because they were used with a landscape layout questionnaire format, common in some parts of Africa, which allocates one page to collecting details for all family members.

195.   The combination of the basic activity status question with other topics, in this case status in employment, as presented in A.6.1 should be avoided if at all possible. In this particular case, it is unclear what the effect of having just two categories—working for someone else or for self—will be on the responses of those in unpaid work, producer cooperatives, home-based workers, etc. The answer categories for question A.6.2 are useful in that they prompt for unpaid work, particularly in agriculture and family businesses. However, strong interviewing skills would be required to ensure that the priority rule is strictly applied if only one category is to be checked for each eligible person. The principles for good interviewing would suggest that all possible responses should first be read out before a response is accepted. However, unless it is clear to the respondent that work done in the first two categories takes precedence over the non-economic activity categories, it is likely that many of the responses to question A.6.2 for women will be "housework", especially among those for whom that is a time-consuming activity. Since the categories are not mutually exclusive, the problem can be addressed by having interviewers check all the applicable categories, as in example A.8 below.

196.   With the landscape layout there is always very little room for prompts or explanations on the questionnaire, and in this case the manual and training would have to have thoroughly covered the meaning of "work" as well as the other categories of "work". Prompt cards could be used to provide further clarification with a list of country-specific types of activities. There is no question to determine whether those who did no work during the period and were not looking for work were attached to a job or business, and were therefore economically active.

197.   It has been common to determine the economically active population through a question with response alternatives that allow the derivation of particular categories, for example, for "status in employment" the categories for status in employment would be used as prompts. For example, asking the question "In the last seven days did you do any work as (*a*) a wage earner/paid employee, or (*b*) an own-account worker with employees …" may save space on the questionnaire, but it would almost certainly lead to an underestimation of the economically active population. That problem will be accentuated by the use of such technical terms as "own-account worker" that are difficult for non-statisticians to understand correctly. If at all possible, the initial question on employment should be left to stand alone, and questions for other topics should be asked separately. The census training of enumerators must emphasize the basic importance of the key question(s). The importance of doing so has been shown in quality investigations of both censuses and surveys.

Example A.7

A.7.1 ¿Durante la semana pasada, trabajó aunque sea por pocas horas?	**A.7.1 During last week, did you work even for a few hours?**
Sí **Pase a X.X**	Yes **Skip to X.X**
No	No
Ignorado	Don't know
A.7.2 ¿Hizo algo en su casa para afuera o ayudó a alguien en un negocio, chacra o trabajo?	**A.7.2 Did you do something in the house for the outside or did you help someone in a business or job?**
Sí **Pase a X.X**	Yes **Skip to X.X**
No	No
Ignorado	Don't know
A.7.3 Tenía trabajo pero, ¿estuvo de licencia por enfermedad, vacaciones, etc.?	**A.7.3 Did you have a job, but were absent from work due to illness, vacations, etc.?**
Sí **Pase a X.X**	Yes **Skip to X.X**
No	No
Ignorado	Don't know

198.   The first of the three questions asked in example A.7 is a simple formulation that includes both the reference period and the idea of "any" work. The second serves as a prompt for some types of work that tend to be overlooked by the respondent or the interviewer. Those categories could be extended to cover specific activities that should be included as economic. The third draws in those who hold jobs even though they were away from work.

199.   The questions in example A.8 were included on the short form and intended for self-enumeration. The first question lists broad categories in the range of activities that could define an individual's economic activity and employment status in the reference

Example A.8

A.8.1 Last week were you doing any of the following? (*X all the boxes that apply*)

  ☐ In paid work

  ☐ Retired from paid work

  ☐ In full-time education or a student

  ☐ Working on a government sponsored training scheme

  ☐ Looking after family or home

  ☐ Permanently sick or disabled

  ☐ Temporarily sick or injured

  ☐ Looking for work and available to start within two weeks

  ☐ Looking for work but not available to start within two weeks

  ☐ Waiting to start a new job

  ☐ Doing unpaid work in own/family business

  ☐ Doing unpaid voluntary work

A.8.2 Are you currently in paid work?

  ☐ Yes

  ☐ No

week. It combines paid work (at the top of the list) and unpaid work (economic and non-economic, at the end of the list). Unemployment categories and non-economic activity categories are also included (some of the categories may not be suitable for other national situations). The technique of providing multiple answers is an interesting approach that census planners in other countries might consider, although some countries prefer to avoid multiple responses based upon their national experience. The identification of the status of employed, unemployed and not currently economically active is then done during data processing, making use of the priority rules of the labour force framework. The second question validates the response to the first category in question A.8.1 and, for some, might also show attachment to a job (but only if it is paid). In countries where unpaid work in family enterprise or self-employment is highly prevalent, the question may not yield the desired results for those groups of workers unless it can be modified accordingly.

200.   In example A.9, the first question (A.9.1) presents many of the categories presented in the previous set of questions (example A.8). They are limited to the main economically

Example A.9

A.9.1 Condición de actividad

¿La semana pasada (NOMBRE)

(Marque con "X" un solo círculo)

trabajó? . . . . ◯ 1 }
no trabajó, pero   } Pase
sí tenía trabajo? . . . . ◯ 2 } a
                                              } X.X
buscó trabajo? . . . . ◯ 3

¿Es estudiante? . . . . ◯ 4

¿Se dedica a los
quehaceres de su hogar? . . ◯ 5

¿Es jubilado(a)
o pensionado(a)? . . . . ◯ 6

¿Está incapacitado(a)   Pase
permanentemente   a
                            Y.Y
para trabajar? . . . . ◯ 7 →

¿Realiza otras actividades? . . ◯ 8

A.9.2 Verificación de actividad

Además de (CONDICIÓN DE A.9.1),
¿la semana pasada (NOMBRE)

(Marque con "X" un solo círculo)

ayudó a trabajar en
un negocio familiar? . . . . ◯ 1

vendió algunos productos
(ropa, cosméticos u otros)? . . . . . ◯ 2

hizo algún producto para vender
(alimentos, artesanías u otros)? . ◯ 3

ayudó a trabajar en
actividades agrícolas o
en la cría de animales? . . . . ◯ 4

a cambio de un pago realizó
otro tipo de trabajo   Pase
(lavó, planchó o cosió)? . . . . ◯ 5   a
                                            Y.Y
¿No trabaja? . . . . ◯ 6 →

A.9.1 Activity status

Last week (NAME)

(Mark with "X" one circle only)

Worked? . . . . ◯ 1 }
Did not work,   } Go
but had a job? . . . . ◯ 2 } to
                                              } X.X
Looked for work? . . . . ◯ 3

Is a student? . . . . ◯ 4

Is a
homemaker? . . . . ◯ 5

Is retired
or receiving pension? . . . . . ◯ 6

Is incapacitated   Go
permanently   to
                            Y.Y
to work? . . . . ◯ 7 →

Did other activities? . . . . . . ◯ 8

A.9.2 Verification of activity

Besides (THE SITUATION IN A.9.1),
last week (NAME)

(Mark with "X" one circle only)

Helped work in
a family business? . . . . ◯ 1

Sold products (clothes,
cosmetics or others)? . . . . ◯ 2

Made something to sell (like
food, handicrafts and others)? . . ◯ 3

Helped work in agricultural
activities or in animal
husbandry? . . . . ◯ 4

In exchange for payment, did
other type of work (washed,
ironed or sewed clothes)? . . . . . ◯ 5   Go to
                                            Y.Y
Did not work? . . . . ◯ 6 →

and non-economically active categories. In this case only one category is to be checked. Presumably once a category is checked the interviewer is instructed to proceed to the next applicable question. Technically these can be seen as representing eight questions (though fewer than eight are posed as the sequence may stop after the first "yes" in the priority order). Question A.9.2 is more elaborate than the corresponding questions in example A.7. Question A.9.2 serves to prompt those who do not report themselves as employed (that is to say, did not get registered as either "worked" or "had a job"), and who are not incapable of working owing to a disability, about non-standard economic activities (income-generating and unpaid economic activities, such as in a family enterprise or a helper on a farm). The preceding categories are important for capturing women's work more completely, for they include activities done by many women who would, however, be reported as homemakers in the first question. Other categories that will be missed are those for which the only economic activities are the production of goods for own household consumption, fetching water and collecting firewood, as well as volunteer work for profit-making institutions.

201.   In example A.10, the first three questions in the set of four used to determine activity status are simple questions presenting one clear idea, requiring a "yes" or "no" response. The first, A.10.1, is the basic and commonly used question, especially when the landscape format is adopted for the questionnaire. It includes the main elements of the concept of employment, minimum hours worked and a reference period specifying the date. The next two questions focus on different types of activities that are likely to be poorly reported. In the third question (A.10.3) a local term is used, which is bound to increase the recognition and understanding of what is entailed in the term "work" used in A.10.1. The fourth question in the set (A.10.4) is similar to the first question in exam-

Example A.10[a]

A.10.1	**Did you/did …… work for at least one hour during the first week of September 2001?**
	☐ Yes → **Go to V.V** ☐ No ☐ Not stated
A.10.2	**Did you/did …… do anything like farming, buying and selling during the first week of September 2001?**
	☐ Yes → **Go to V.V** ☐ No ☐ Not stated
A.10.3	**Did you/did …… do any type of odd job or hustling during the first week of September 2001?**
	☐ Yes → **Go to V.V** ☐ No ☐ Not stated
A.10.4	**What were you/was …… doing for most of the time during the first week of September 2001?** (READ CATEGORIES)
	☐ Working in agriculture or any other business without pay
	☐ With job not working → **Go to W.W**
	☐ Seeking first job → **Go to Y.Y**
	☐ Seeking a job which was not the first → **Go to X.X**
	☐ Did not seek work but wanted work and was available → **Go to X.X**
	☐ Student → **Go to Z.Z**
	☐ Did home duties → **Go to Z.Z**
	☐ Retired did not work → **Go to Z.Z**
	☐ Disabled unable to work → **Go to Z.Z**
	☐ Not interested in work → **Go to Z.Z**
	☐ Other → **Go to Z.Z**
	☐ Not stated

a Used in a long form on a 10 per cent sample.

ples A.8 and A.9 and combines non-economic activities with unemployment categories as well as the residual from the economic activities addressed in the first three questions. Like the questions in example A.9, set A.10 covers all the major activities except production of goods for own household consumption, fetching water and collecting firewood, and it seems likely that persons who do only such activities will be excluded from the employed population.

202. Each of the first four questions in example A.11 (A.11.1 to A.11.4) deals with a specific aspect of the work that qualifies as economic activity. The first concerns the ownership of land, and though a positive response to this question does not necessarily mean that the individual is employed, the response forms a basis for probing further if at the end of the series of questions the individual is to be classified as non-economically active. The second (A.11.2) is the standard question that many countries use to determine if the individual worked during the reference week. The third question (A.11.3) probes for specific activities that tend to be missed, especially in countries with a large informal sector or with subsistence activities. It is important to note that in A.11.2 and A.11.3 both the reference week and the principle of "any work" are reflected, as the respondent cannot be expected to assume that the two questions are connected. The fourth question (A.11.4) represents the group of people who had a job or enterprise but were not at work during the reference week.

5. Employed: suggestions for enhancing the effectiveness of the questions

203. One important consideration in any decision to adopt the extended scope of economic activities is how comparability with results from previous censuses and with other sources can be maintained, while setting a new base year for the new concept. Nonetheless, if the previous questions were inefficient in capturing specific groups of workers, it may be more critical to improve coverage than to maintain comparability. It is, however, still important that the approach adopted permit separation of the household economic activities from the other activities, especially for those who have concerns about how including

Example A.11

A.11.1 Do you/Does .... own any land for farming?

   ① Yes

   ② No

   ⑨ Don't know/Not stated

A.11.2 Did you/ .... do any work for pay, profit, or family gain for at least one hour during the week ending 12 May? This includes helping in a family business or farm, street vending, or work at home.

   ① Yes → **Skip to A.11.5**

   ② No

   ⑨ Don't know/Not stated

A.11.3 Did you/ .... engage in any economic activity on the following list, for pay, profit or family gain, for at least one hour, during the week ending 12 May?

   ① Yes → **Skip to A.11.5**

   ② No

   ⑨ Don't know/Not stated

A.11.3 (*continued*): **INTERVIEWER: READ THE FOLLOWING LIST**

Babysitting	Nurse's aid for pay
Laundry, ironing for pay	Subsistence farming
Cleaning yard/cutting yard	Sewing for pay
Selling food or snack at market/bus stop/school	Cleaning of offices
	Car washing
Bicycle/cart deliveries	Taxi driver
Selling food from home	Lottery vendors
Selling sweets from home (fudge, etc.)	Any other similar activity

A.11.4 Did you/ .... have a job during that week, from which you were temporarily absent?

   ① Yes

   ② No → **Skip to X.X**

   ⑨ Don't know/Not stated → **Skip to X.X**

A.11.5 How many hours did you/ .... work during that week?

   [ ] ⓪ 1 2 3 4 5 6 7 8 9

   [ ] ⓪ 1 2 3 4 5 6 7 8 9

   99 Don't know/Not stated → **Skip to Y.Y**

some types of household production, such as fetching water and collecting firewood, might dilute the policy relevance of the measures of employed and unemployed.

204.   To address those concerns, the following steps can be adopted:

(*a*)   Keep the basic questions used in the previous census (with improvements as deemed necessary, based on the census experience);

(*b*)   Add one question that prompts for the main activities comprising the difficult-to-measure household economic activities, including fetching water and collecting firewood.

205.   Several points to note about introducing the question described in (*b*) above:

(*a*)   Countries can make fewer adjustments to previously used questions:

(i)   If they used two or more questions in their census;

(ii)   If the questionnaire format allows more space for writing questions: for example, one or more pages of questions per individual, or several columns for each person with ample space in the rows to enter the questions, as opposed to one page for all or almost all members of the household;

(iii)   If they have regular labour force or household surveys that can be used to determine the effect on employment and unemployment figures of adopting the broader scope of economic activities. The census could then be limited to the main set of questions.

(*b*)   It is preferable that the question described in (*b*) above be confined to probes on specific economic activities (relating to work done) and not be mixed with categories of job attachment;

(*c*)   The probes can be presented either as a set of categories in a question, and the response(s) entered against the appropriate category(ies), or as a question with "yes" and "no" response categories, and the activity list entered on the questionnaire or on a prompt card that will be read by the enumerator. An example of the prompt card is presented in figure III. It should be modified to suit national circumstances;

(*d*)   A decision is required on whether there is to be only one response per question (as in example A.9) or whether there can be multiple responses (as in example A.8). It can be useful to allow multiple responses (although some countries prefer to avoid such approach), so that different types of aggregation can be possible at the data processing and tabulation stage. In that context, the implications for coding and data processing should be carefully considered.

206.   The final example (A.12) is an untested hypothetical set of questions to show how to ensure coverage of the new activities to be regarded as employment. Some other questionnaire techniques are also suggested.

207.   The major points to note are:

(*a*)   Prompts are included on the major employment activities. It is preferable to make them specific and include the most important activities likely to be missed;

(*b*)   Separate questions emphasize the one-hour limit in order to include short-time work, unpaid work and temporary absences. Those are the situations commonly missed when only one question is asked on whether the individual did any work during the reference week.

208.   A possible extension would be to make each of the prompts in A.12.3 a separate question with its own "yes/no" answer, or a multiple response question (which some countries prefer to avoid). One drawback of that approach is that it would extend the

Figure III
Current economic activity prompt card

Economic activity

Did you do any of the following activities in the last seven days?

Include any temporary absences for leave, sickness or family/village responsibilities

- **Any activities on your own or family farm, garden, cattle post/kraal of any kind** (products can be for sale or family use).

  Examples:
  - Planting crops
  - Weeding
  - Harvesting crops
  - Chasing birds or other pests off crops
  - Looking after cattle or other animals
  - Moving crops to storage or sale

- **Fetch any water or collect any firewood**

- **Any kind of business activity, big or small**

  Examples:
  - Small shop/kiosk/tuck shop
  - Street or market selling
  - Shoe cleaning/repair
  - Beer brewing
  - Making/selling hats, mats, baskets, caps

- **Catch, collect, cultivate any fish, prawns, seaweed, shells or other food from the sea, river or fish/prawn farm** (fish etc. can be for sale or family food)

- **Make anything from farm or natural products for sale or own use**

  Examples:
  - Making handicrafts, mats, hats from straw or timber
  - Making charcoal
  - Making butter from milk

- **Do any construction or major repair work on your house or farm or for anyone else**

  Examples:
  - Build/repair fences/kraals
  - Make/repair irrigation channels

- **Have any type of wage or salary job, paid in money or by food/clothing/accommodation**

  Examples:
  - Full- or part-time jobs
  - Casual or piecework jobs
  - Permanent or temporary jobs

- **Do any volunteer work to produce any building or other physical good**

  Example:
  - Self-help community construction of bus shelter

- **Do any volunteer work to provide a service for a market enterprise**

  Example:
  - Volunteer work at private hospital

Example A.12

<table>
<tr><td colspan="2">Hypothetical questions, recommended for testing and use in countries</td></tr>
<tr><td>A.12.1</td><td>Did you/… do any work for one hour or more in the last seven days?</td></tr>
<tr><td>A.12.2</td><td>Did you/… do any unpaid work in a family business or farm in the last seven days?</td></tr>
<tr><td>A.12.3</td><td>Did you/… do any work at all in the last seven days?</td></tr>
<tr><td></td><td>Prompts: Any work on your farm or kraal?
Any fishing or seafood collecting?
Make anything for sale or your own use from farm or natural products?
Fetch any water or collect any firewood?
Any work in a business of any type?
Any type of wage job (full-time or part-time)?</td></tr>
<tr><td>A.12.4</td><td>Do you/Does … usually work but happened to be absent last week because of leave, sickness, bad weather, industrial troubles or other reasons?</td></tr>
<tr><td></td><td>Interviewer: If "yes" to any of A.12.1 to A.12.4, skip to next section.</td></tr>
</table>

sequence of questions. It would, however, supply additional information and could be used for creating comparability with the results from previous censuses. If compatibility with historical data is a priority objective, the building-block approach may be adopted. That would involve keeping the questions that have been used to enumerate the employed and adding a combination of A.12.1, A.12.2 and A.12.3 to capture the other groups of employed that may have been missed, as in example A.9. The temptation to incorporate the status in employment variable here should be resisted. An alternate approach would be to use a prompt card in place of the prompts to question A.12.3. A possible prompt card is given as figure III. Note that the examples of activities will differ for each country, therefore the examples on the prompt card need to be varied and country specific. In addition, relevant activities can be added.

209.   Prompt cards have proved to be very effective in surveys, and they certainly should be examined and tested for censuses. The list and ordering of categories and the specific activities covered need to be tested with respect to (*a*) the ease with which the respondent can identify with some of the activities; (*b*) the spontaneity with which the respondent reacts to the categories as they are read out; (*c*) the effect of the order of the questions and of the list of activities; and (*d*) the significance of the numbers added to the employed when using a multiple-category question as compared with using a prompt card to probe for important activities most likely to be missed.

6.   Employed: issues of measurement

210.   The following aspects may need pre-testing on current activity:

(*a*)   Any change in length or type of reference period from that used in previous censuses;

(*b*)   Success or otherwise in recording difficult groups, for example, part-time and unpaid work;

(*c*)   Effects of the inclusion of all additional groups from the extended production boundary on all classifications, for example, the effect of the inclusion of fetching water and collecting firewood on activity status, industry and occupation groups;

(*d*)   Effect of prompts, additional questions or prompt cards in relation to subparagraph (*c*);

(*e*)   Possible additional categories/wording/examples for prompts or prompt cards based on persons being wrongly recorded;

(*f*)   Temporary absences: what is being taken into account?

(*g*)   Evaluation of "seeking work" compared to "available for work" for the unemployed;

(*h*)   Categories being recorded in "other" for the "not economically active";

(*i*)   If a simple layout is employed, its success in recording both employed and unemployed.

C.   Unemployed population

211.   In many developing countries, unemployment as defined with three criteria (without work, available for work and seeking work) may not encompass all the imbalances in the labour market. A low general unemployment rate certainly does not mean that the developing countries or regions do not have employment problems. Lacking the unemployment benefits and other social services common in most developed countries, most people in developing countries cannot be totally unemployed or they would not survive. Few families can support totally unemployed youngsters or adults. Many people in developing countries generate incomes from a variety of informal activities or survive by growing their own food, and those activities represent employment. In developing countries unemployment in the strict sense is therefore mostly found in urban or peri-urban areas where a person can be totally unemployed while obtaining support from other employed family members or relatives. Among young persons unemployment may also tend to be concentrated among those who are better educated, as they tend to come from better-off families that can afford to support periods of unemployment for them. Thus, in those countries, the unemployment rate, as conventionally defined, is a useful indicator of labour market problems only for groups such as those mentioned above.

212.   The problems in developing countries are more complex than the simple rate can show. That is why better and broader statistics on economic characteristics are needed, so that the situation of both the economically active population and the population not currently economically active can be examined in some detail. However, the census may not be the best source to obtain those more refined measures of the labour force, and they should be left to labour force and other household surveys. The major value of the census measure is that it supplies statistics for small areas and small population groups.

1.   Unemployed: the operational definition

213.   The category of "unemployed" comprises all persons above a specified age who during the reference period were:

(*a*)   Without work, that is, were in neither paid employment nor self-employment;

(*b*)   Currently available for work, whether paid employment or self-employment;

(*c*)   Seeking work, that is, had taken specific steps in a specified recent period to seek paid employment or self-employment.

2.   Unemployed: requisites for applying the definition

214.   The first major requirement to be counted as an unemployed member of the economically active population is that the person should not have been employed during the reference period. The second major requirement is that the person could have worked in the relevant reference period, or could have started working almost immediately afterwards if a job had been offered during that period. The third requirement is that, in the recent past, the person should have been actively seeking work.

215.	The requirement to be "not working" applies to the reference week in such cases, following the priority rules. For the "available" criterion the international recommendations state that the period should be the current reference week (although some countries use the current reference week and the following two weeks). The extension of the availability period into the future allows a person to make the arrangements necessary to cope with other commitments after having received a job offer, for example, women engaged in childcare or housework. The international recommendations do not specify a precise reference period for actively seeking work, but the previous month or four weeks has been used most commonly. Lengthening either of the two reference periods will give broader measures of unemployment and add to the numbers of unemployed persons. It is important to be aware of and to try to avoid including any public holidays and other major events of more than a few days duration in reference periods used for the categories "seeking work" or "available for work".

216.	The means to "actively seek work", as understood in developed countries, include the following: registration at a public or private employment exchange; application to employers in writing or in person; checking at work sites, farms, factory gates, market or other assembly places and placing or answering newspaper advertisements. Steps towards self-employment include looking for land, buildings, machinery or equipment to establish own enterprise; arranging for financial resources; applying for permits and licences and the like, or taking other steps to start a business or similar steps to start an agricultural activity. Such "active" steps are often difficult to identify in most developing countries, particularly in rural areas. Employment offices often do not exist, or are known to be ineffective and few bother to contact them. Often, there are also few businesses or government offices in the area at which job seekers may apply for work. Funding or suitable land for businesses or agriculture is frequently very limited. However, a person who has no work might in fact be available for work if any type of economic activity were open for him/her in the area. Women in particular may keep themselves busy with housework and take no active steps to find employment, but if the opportunity presented itself they could easily be available for that work. In addition, there are other categories of persons who did not actively seek work but were available for work in the reference period who may be considered to be unemployed under the international standards. They include those who were expecting to start work shortly, those who did not believe work was available, those who did not know how or where to look for work, and so on (see Hussmanns, Mehran and Verma, 1990).

217.	The traditional strict measure of unemployment that requires persons to have been actively seeking work will therefore generally mean that the measured unemployment rates for many developing countries will be low. The international recommendations recognize the problem and allow countries to relax the actively seeking work criterion if necessary in their situation. However, very few countries seem to have done this in their censuses, and that may partly explain why very low unemployment rates are reported from most developing and transition countries.[5] It is not only the requirement that a person should have been actively seeking work that might act to lower the unemployment rate, but also the broad definition of employment if it is effectively applied.

218.	Ideally, questions on both "actively seeking" and "available" should be asked to make it possible to give statistics for both "strict" and "relaxed" unemployment measures. It may, for example, be useful to ask both questions in one census to provide both strict and relaxed estimates of unemployment, particularly for comparison of the results, and then possibly continue in future censuses with one concept only. It would also be useful to distinguish first-time job seekers from other job seekers in the classification of the unemployed.

219.	The priority rules of the labour force framework give precedence to unemployment over economic inactivity. Thus even persons who during the reference period were mainly not economically active should be classified as unemployed if they satisfy the three con-

[5] The possibility and relevance of using only the "currently available for work" criterion therefore should be explored.

ditions in the above definition of unemployment. Such persons could include students, homemakers and pensioners. Note, however, that on the basis of the same priority rules, persons who performed an economic activity during the reference period should be classified as employed even if they were mainly without work, available for work and looking for work during the period.

3.   Unemployed: review of questions used in national censuses

220.   Questions on unemployment have been formulated in a variety of ways, and they are differentiated (*a*) by the length of the reference period, that is, they may use either the same reference period as the employed (for example, one week) or a reference period longer than but including the reference period for the employed (for example one month, including the week for the employed); (*b*) by whether only one of the two criteria of seeking work and available for work or both are used; (*c*) if both criteria are adopted, by whether they are applied in one question, in a two-part question or in two separate questions; and (*d*) by whether to use the same reference period for all three (or two) criteria.

221.   The present section explores the following types of questions, with emphasis on the use of the two criteria of seeking work and available for work:

(*a*)   One question that uses one criterion only, applying either the short or the long reference period (examples B.1 to B.6);

(*b*)   One question with both criteria asked together (example B.7);

(*c*)   Two or more questions or parts to a question addressing each criterion in turn, applying either (i) the same reference periods, or (ii) a different period for each criterion (examples B.8 to B.13).

Example B.1

B.1.1 Are you:	Employers, the self-employed, farmers, as well as workers in the family business are considered to be employed if they work 12 hours a week or more.
a) Employed:	
Full-time work (33 and more hours per week)	
Part-time work (12 to 32 hours per week)	**Please answer questions 11 to 16!**
b) Not economically active, but:	
Unemployed	**Please answer questions 11 to 13 on your last occupation!**
Parental leave, maternity leave	**If you have never worked before and are looking for a job, answer question 12 with "no profession yet"!**
Were you unemployed at the time you entered maternity leave:	● Yes ● No
Military service, replacement service	**Please answer questions 15 and 16 for your journey to the barracks or to your place of service.**
Retirement benefits from own employment	**Please answer questions 11 and 12 on your last occupation!**
Widow's pension	
Homemaker	
Pupil, student	**Please answer questions 14 to 16 on the school you are presently attending!**
Child presently not attending school	
Other livelihood	**E.g., rent, alimony, social aid, financial support through relatives, etc.**

Example B.2

B.2.1 **Condición de actividad**	B.2.1 **Activity status**
¿La semana pasada (NOMBRE)	Last week (NAME)
(Marque con "X" un solo círculo)	*(Mark with "X" one circle only)*

trabajó?	1	Worked?	1
no trabajó, pero sí tenía trabajo?	2	Did not work, but had a job?	2
buscó trabajo?	3	Looked for work?	3
¿Es estudiante?	4	Is a student?	4
¿Se dedica a los quehaceres de su hogar?	5	Is a homemaker?	5
¿Es jubilado(a) o pensionado(a)?	6	Is retired or receiving pension?	6
¿Está incapacitado(a) permanentemente para trabajar?	7	Is incapacitated permanently to work?	7
¿Realiza otras actividades?	8	Did other activities?	8

(Pase a X.X for 1–3; Pase a Y.Y / Go to Y.Y for 7. Go to X.X for 1–3.)

222. In example B.1, the meaning of "unemployed" is left for the interpretation of the respondent. It could therefore reflect individuals who fit either one of the criteria or both. In addition, the structure of the question suggests that the unemployed are considered among the "not economically active", which is conceptually not the case, according to the international recommendations.

223. In example B.2, the basic question seeks to determine the activity status of individuals and uses the reference period of the past week, which is applied to all categories presented, one of which refers to whether the individual looked for work. The question that follows (shown in example A.9) acts as a filter to exclude from the groups reporting that they did not work in the reference week those that performed some kind of economic activity. Together those questions give an estimate of those who may have been employed. The reference period is probably too short, since some people who may be looking for work do not keep at it continuously. The criterion of "being available" is not applied. The two elements can produce opposite effects on the total number of persons classified as unemployed, but do not necessarily cancel out.

224. Example B.3 also uses the same basic question for some types of activities of the employed, the unemployed and the not economically active groups. The 30-day reference

Example B.3

Type of activity during past 30 days

B.3.1 **During the past 30 days did ... work for cash?**	B.3.2 **Then what did ... do during the past 30 days?**
1. Yes, for someone else — Go to X.X	1. Family business
2. Yes, for self	2. Work at lands/farms/cattle post
3. No → Go to B.3.2	3. Actively seeking work
	4. Housework
	5. Student
	6. Retiree
	7. Other, specify _____

period applies to all activity statuses but is a reasonably long period to consider measuring the unemployed. It is the duration typically used for "actively seeking" even for those that apply the one-week reference to the employed. With the landscape layout used in this country's questionnaire little room is available for prompts or explanations on the questionnaire, and in this case the manual and training would have to have thoroughly covered the other categories of "work" and the meaning of "actively seeking work" in particular. Note also that in practice, the "actively seeking work" activity will be combined with one or more of the others mentioned; therefore, allowing for multiple replies would have been an advantage.

225.   Question B.4.1 is addressed to persons who neither worked nor had a job from which they were temporarily absent during the reference week. The main criterion applied in the question is "available for work"; it therefore relates to the relaxed definition. The question is complicated by the fact that the reasons why the individual would not have taken a job offered in the reference period are sought without first establishing whether the individual was available for or interested in taking a job. Only one category gives the clear idea that the person is not seeking work. Moreover, the response category "nothing" would seem to suggest that it is the group that can be taken as the unemployed, but it presents some ambiguity as to what "nothing" really represents. People in that category may still not take a job that is offered to them, although nothing prevents them from taking it. The reference period of one week is also short, because some people with home duties may need more time to arrange for help, even if they want to accept the job offered.

226.   Question B.5.1, posed to those who neither worked nor were on leave or temporary layoff from their job during the past week, is straightforward and should be understood by most people. It does not state what is entailed in "look for work", but presumably the respondents could take the smallest form of effort as seeking work. The duration is the four weeks preceding the census, as recommended in the international standards.

227.   In this particular census questionaire, question B.6.1 is preceded by two others on employment that refer to "last week" as the reference period; the question on seeking work refers to the last four weeks (incorporating the last week as well). A single question is used to determine the unemployed status, and only the actively seeking criterion is applied. The question includes some clarification of what actively seeking entails. How-

Example B.4

B.4.1   **If you/ ... had been offered a job that week, what would have prevented you/ ... from taking up that job?**

1. Nothing   4. Retirement

2. School   5. Not interested in working

3. Home duties   6. Other (specify _____)

                               9. Don't know/Not stated

Example B.5

B.5.1 **Durante las últimas cuatro semanas, ¿buscó trabajo?**	B.5.1 **During the last four weeks did you** (OR NAME) **look for work?**
Sí ○	Yes ○
No ○	No ○
Ignorado ○	Don't know ○

ever, not all those who were seeking employment may be available to take a job offered to them on short notice. The measure from question B.6.1 may therefore overestimate the number of people who fit the strict definition of unemployed, that is to say, those who are seeking and are available for work.

228.   In example B.7 "marginal" and "non-worker" are defined on the basis of previous questions. The concepts "seeking" and "available" need to be defined for the question to be effective, but such definitions are not presented. It cannot be deduced from the question whether respondents, when they indicated "yes", had sought work, were available for work, or both. The brevity of the question stems from the very limited space that the landscape format of presentation offers.

229.   Question B.8.1 combines several elements of both activity status and status in employment in the list of possible response categories. Only one response is expected, given the keyword "mainly" in the basic question. The reference period of one week applies to the definitions of both employed and unemployed, and therefore some of the unemployed may be lost to other categories such as "homemakers", for example, those who had searched for a job a bit earlier and/or would be available for work a bit later. However, the use of "mainly" might inflate the number that could be considered as unemployed if, for example, more time was spent in the reference week seeking work than working (for example, if the person was working one day and searching for work the other four to six days). Thus, it is the respondent that decides the order or priority to assign to the different activities that he or she was engaged in during the reference period. The two

Example B.6

B.6.1	**Did you actively look for work over the last four weeks?**
	For instance: through employment bureaux, wanted ads, contacting potential working places, etc.
	☐ Yes **1** ☐ No **2**

Example B.7

B.7.1	**If marginal worker or non-worker, is the person seeking/available for work?**
	☐ Yes **1** ☐ No **2**

Example B.8

B.8.1	**What were you mainly doing in the last seven days?**	
	Worked—Paid non-seasonal	[1]
	Worked—Unpaid non-seasonal	[2]
	Worked—Paid seasonal	[3]
	Worked—Unpaid seasonal	[4]
	On leave	[5]
	Unpaid work on household holding or business	[6]
	Unemployed and seeking work	[7]
	Not seeking work but available for work	[8]
	Full-time housewife/homemaker	[9]
	Full-time student	[10]
	Not available for work for other reasons	[11]

criteria are reflected in the categories (B.8.1, responses 7 and 8). However, the response category "Not seeking work but available for work" is not a direct response to a question on what one was mainly doing, since there is overtly no action involved.

230.   In example B.9 only the strict concept of unemployment was used. Response 7 of question B.9.2 should preferably distinguish those who have never worked before from those who have worked only in subsistence agriculture or fishing.

231.   The three questions in example B.10 are related to seeking and being available for work, but the second and third are more specific to measuring unemployment. The first, B.10.1, targets those who are waiting to start work within the next four weeks, whereas the other two apply to those who did not work the previous week and have been actively

Example B.9

B.9.1 (*a*) **Has this person been looking for work to earn money during the last four weeks?**

    1 ☐   Yes → **Go to X.X**

    2 ☐   No → **Skip to B.9.2**

B.9.1 (*b*) **Last week could this person have started a job if offered one, or returned to work if recalled?**

    1 ☐   Yes, could have gone to work

    2 ☐   No, because of temporary illness

    3 ☐   No, because of all other reasons (in school, etc.)

B.9.2   **When did this person last work even for a few days? Do not include subsistence activity.**

    1 ☐   2000

    2 ☐   1999

    3 ☐   1998

    4 ☐   1995 to 1997

    5 ☐   1990 to 1994 → **Skip to Y.Y**

    6 ☐   1989 or earlier → **Skip to Y.Y**

    7 ☐   Never worked or did subsistence only → **Skip to Y.Y**

Example B.10

B.10.1   **Last week did you have definite arrangements to start a new job within the next four weeks?**

    ☐   No

    ☐   Yes

B.10.2   **Did you look for work during the past four weeks?** For example, did you contact a National Employment Centre, check with employers, place or answer newspaper ads?

    Mark one box only

    ☐   No

    ☐   Yes, looked for full-time work

    ☐   Yes, looked for part-time work (less than 30 hours per week)

B.10.3   **Could you have started work last week had a job been available?**

    Mark one box only

    ☐   Yes, could have started work

    ☐   No, already had a job

    ☐   No, temporary illness or disability

    ☐   No, personal or family responsibilities

    ☐   No, going to school

    ☐   No, other reason

looking for a job. The reference period for the search (B.10.2) is reasonably long, and the time within which the job seeker could start work, should a job be available, is just one week, as recommended in the international standards.

232.   In question B.11.1 (see also example A.8), not all categories used are applicable or relevant for other countries, and testing may show that some adjustments may be necessary if this approach is adopted. Each of the two categories dealing with the unemployed has two parts that are differentiated with conjunctions ("and" and "but"). The formulation could work out reasonably well with a self-administered questionnaire (as in this particular case), but it may not work well in the interview mode of data collection. The two categories may sound the same if the respondent is not very attentive. Highlighting—or for the interview mode, stressing—the words "and" in the first and "but not" in the second may help. Since the technique allows multiple answers,[6] interviewers will need to be made aware that the respondent may check only one of the two categories relating to unemployment, although any others may be checked if they apply. Such a restriction in a multiple-choice question would need to be tested out. Also, there is no category for "available for work, but not looking for work", which would permit using the relaxed definition of unemployment. Finally for example B.11, the reference period for availability for work (namely "available to start within two weeks") is not standard. It should refer to "available to start last week" or, in some countries, to "last week or the next two weeks".

233.   The reference period for looking for work in question B.12.1 follows the international standard of four preceding weeks, but the reference period for availability for work in question B.12.2 is for the following two weeks only. It should be "last week" or "last week and the next two weeks" to be consistent with international practice. The third and fourth questions are used to distinguish first-time job seekers from those who have recently lost a job or from re-entrants into the labour market.

234.   Two questions are used in example B.13 to determine unemployed status. The first uses the past four weeks as the reference period, as in most of the previous examples, and gives examples of what constitutes taking active steps. The "yes/no" response categories help to simplify the question. Availability to start work is assessed in the second question. Instead of applying a specific reference period, four time intervals are used, three of which correspond to reference periods generally used with this question. That approach provides the flexibility to define and subsequently adjust the reference period through

[6] As indicated in paragraph 199, some countries avoid the use of multiple responses.

Example B.11[a]

B.11.1   Last week were you doing any of the following?

X all the boxes that apply

☐   In paid work

☐   Retired from paid work

☐   In full-time education or a student

☐   Working on a government-sponsored training scheme

☐   Looking after family or home

☐   Permanently sick or disabled

☐   Temporarily sick or injured

☐   Looking for work and available to start within two weeks

☐   Looking for work but not available to start within two weeks

☐   Waiting to start a new job

☐   Doing unpaid work in own/family business

☐   Doing unpaid voluntary work

a Included in a short form questionnaire.

Example B.12

B.12.1	**Indicate whether, during the last four weeks, the person has actively tried to find a job or attempted to open own activity.**
	1 ☐ Yes
	2 ☐ No
B.12.2	**Indicate whether, should the opportunity arise, the person would be willing to start work within 2 weeks.**
	1 ☐ Yes
	2 ☐ No
B.12.3	**Indicate whether the person has ever held a paid job or worked as a contributing family worker during their lifetime.**
	1 ☐ Yes → **B.12.4 Indicate the year when terminated** _____ → **Go to question X.X**
	2 ☐ No → **Go to question X.X**

Example B.13

B.13.1a Active steps	B.13.1 Availability
If NO to X.X	**If NO to X.X**
In the PAST FOUR WEEKS before 10 October has (the person) taken active steps to find employment?	**If offered to work, how soon could (the person) start?**
Y = Yes	1 = Within 1 week
N = No	2 = More than 1 week, up to 2 weeks
For example, (the person) went to visit factories or other employment places, placed or answered advertisements, looked for land or a building or equipment to start own business or farm.	3 = More than 2 weeks, up to 4 weeks
	4 = Some time after 4 weeks
	5 = Does not choose to work
	Go to Y.Y

aggregation of the responses. It also permits analyses and some policy questions related to unemployment that could be explored. If such a formulation is to be adopted in countries that have not previously used it, it is important that the question be tested. In particular, it is important to ascertain the effect that presenting the range of options might have on people's willingness to start work offered within a short period. It is, for example, possible that many people who could start work within two weeks might have the higher time range, if the option were presented. If a reference period was then imposed at the tabulation stage, the results could be misleading.

4.   Unemployed: suggestions for enhancing the effectiveness of the questions

235.   To obtain a better insight into the magnitude of unemployment, one might explore the possibility of relaxing the actively seeking work criterion. Household surveys often include questions on actively looking for work and on available for work. Including both questions provides a possibility for investigating the effects of relaxing one of the criteria when defining unemployment. Testing should be done to establish the effect before the relaxed definition is introduced into the census questionnaire. The available for work criterion is better as a separate question than as a simple response category. However, it is probably best to aim for a simple, crude unemployment measure from the census and

leave it to household surveys to apply more precisely defined concepts. While the topic of unemployment is often thought to be simple, census planners should definitely investigate it carefully to establish exactly what they are measuring.

236.   Example B.14 follows up on example A.12 in the section on employment. Questions B.14.1 and B.14.2 deal with unemployment. Both the strict and the more relaxed definitions of unemployed can be used on the basis of the responses given to these two questions.

237.   The responses shown in question B.14.1 are perhaps too detailed and might be simplified. Those who indicate family responsibilities and illness in question B.14.1 are not available for work and could skip question B.14.2.

238.   Care is needed in phrasing the question on availability for work (question B.14.2). In some countries, such a question has raised respondent's expectations that a job was being offered to them. The question may also be inappropriate where proxy responses are provided on behalf of other household members.

5.   Unemployed: issues of measurement

239.   The possible variations on the unemployment question(s) should be thoroughly tested before being introduced. Pre-testing should also focus on the respondents' interpretation of the word "available", particularly in countries where translation in the field by the enumerator is common.

Example B.14

Hypothetical questions, recommended for testing and use in countries

B.14.1   Did you/ ... actively look for work sometime during the last four weeks?

☐   Yes

☐   No—seasonal worker awaiting busy season

☐   No—believe no work available

☐   No—tired of looking

☐   No—don't know how or where to look

☐   No—awaiting appropriate work

☐   No—bad weather

☐   No—family responsibilities

☐   No—own illness

☐   No—other reasons (specify _____)

B.14.2   If offered a work opportunity could you have started work last week?

☐   Yes

☐   No

B.14.3   What were you doing/what was your situation last week?

Studying	1
Performing housework	2
Not working and disabled	3
Sick	4
Retired/aged	5
Pension, rental or other income recipient	6
Other	7

D.   Population not currently active

240.   As noted earlier, the "population not currently active" consists of all persons in the population who, during the reference period, are neither employed nor unemployed. It is useful to determine the reasons why those persons are in this group in terms of the alternative activities they were engaged in that are assigned less priority in the labour force framework than being employed or being unemployed. Question B.14.3 elicits responses from those who were not employed during the reference period as to the nature of their actual activities during the period. Responses to that question could permit the analysis of the reasons why some persons are not currently economically active. The categories used in the question should be checked in pre-tests to ensure they are being understood and are sufficient and useful for the prevailing situation in each country.

Chapter V
Usual economic activity status

A.   Introduction

241.   Statistics on economic activity over a 12-month period, rather than in respect of a recent short period as described in chapter IV, are important to users who have an interest in a summary picture of activity covering all seasons. Such data are used in economic and manpower planning and can be particularly valuable as a basis for national accounting estimates. They are less dependent on the timing of the census date (see chap. II, sect. C) and fit in with other statistics determined on an annual basis, such as household income, or statistics reflecting the normal situation for household members, such as enrolment in school.

242.   Statistics on economic activity over a 12-month period may be collected using a "predominant activity" approach or a "usual status" approach, as follows:

(*a*)   **Predominant activity.** The predominant activity approach relates to the respondent's view of his or her predominant or main activity (among employed, unemployed and inactive) over the 12-month period;

(*b*)   **Usual status.** The usual status approach is more complex and requires the separate measurement of time spent employed, unemployed and inactive (more details are given below).

243.   The predominant activity approach differs from the usual status approach. A person who is inactive for 22 weeks, employed for 18 weeks and unemployed for 12 weeks would probably describe themselves as predominantly inactive, but under the usual status approach described in more detail below, they would be classified as usually active (employed).

244.   In the 2000 round of population censuses, fewer than 10 countries collected data on economic activity and economic characteristics exclusively with the "predominant activity" or "usual status" approach, typically using a small number of questions (ranging from three to six). However, many countries that use the current activity approach described in chapter IV supplement their census questions with a question or set of questions on predominant or usual activity (see also *Principles and Recommendations, Revision 2*, paragraph 2.246).

245.   During a 12-month period many persons change between economic and non-economic activities, depending on the season as well as on economic and demographic factors and events, and they can do different types of economic activities at different times of the year. It therefore is very difficult to obtain meaningful responses by simply asking persons, "What was your usual activity in the last 12 months?" or "Were you usually employed?" The questions that are required need to probe for all types of economic activity and preferably aim to quantify the answers in terms of time spent on the various forms of activity. As a result, it may be necessary to use several questions. There may also be some difficulty in identifying a main job that would serve as the reference for such economic characteristics as occupation and industry. The relative precision of

the current status measure and its general ease of interpretation for interviewers and respondents make it preferable to the usual status measure for the questions on other economic characteristics.

246.   It has been recommended in *Principles and Recommendations, Revision 2*, that countries consider measuring both the usually economically active population and the currently economically active population. It is also strongly recommended that any country starting to use usual activity for all their economic characteristics questions conduct a very extensive pre-testing programme to identify all difficulties and build experience in applying the concept.[1] The broad aim of measuring usual activity is to obtain a "total" or "average" picture of economic activity over all seasons.

247.   There is less need to measure usual activity in a population census, if a country has the following:

(*a*)   A reasonably even pattern of economic activity throughout the year;

(*b*)   A very high proportion of employment in wage-earning or formal business activities, which by their nature are not strongly seasonal;

(*c*)   A very good programme of household surveys to measure economic activity at different times of the year, thus covering the various seasons.

248.   Most of the rest of chapter V deals with the measurement of usual status but examples are also shown for predominant activity.

B.   Usually active: operational definition

249.   The "usually active population" comprises all persons above a specified age whose activity status, as determined in terms of the total number of weeks or days during a long specified period (such as the preceding 12 months or the preceding calendar year) was "employed" and/or "unemployed" as defined within the labour force framework.

250.   The "usually employed population" comprises all persons in the usually active population who, during the period of usual activity, had a total number of weeks or days of employment that was at least half of this period. Conversely, the "usually unemployed population" comprises all persons in the usually active population who, during the period of usual activity, experienced a total number of weeks or days of unemployment greater than half the period.

251.   The "population not usually active" consists of all persons in the population who were not usually active during the long specified reference period, including those below the specified age.

252.   Figure IV illustrates the above-mentioned groups and the relationship between them.

C.   Usually active: requisites for applying the definition

253.   One fundamental issue in the choice of a reference period is whether to use the last 12 calendar months or a calendar year, for example, 2010. Recall problems or memory loss are a major problem with such long periods, in particular concerning the timing of events and duration of activities, and it has been found best to use the 12 calendar months ending at the month prior to the census. However, some of the countries collecting predominant activity have used the calendar year, as it is believed to give a more definite reference period to which respondents can relate. It also gives a reference period for the results, which can more easily be combined with other types of statistics for the year. If the census is early in the calendar year then there may be little difference, and the calendar year could be preferred; however, that should be established by testing. A

[1] Surveys in countries that have measured both current activity and usual activity have shown considerable differences in the results, even in the basic activity classification (as employed, unemployed and so on). Thus, any change from a current activity approach to a complete usual activity approach is likely to result in lack of comparability with previous census results.

Figure IV
Usually active population
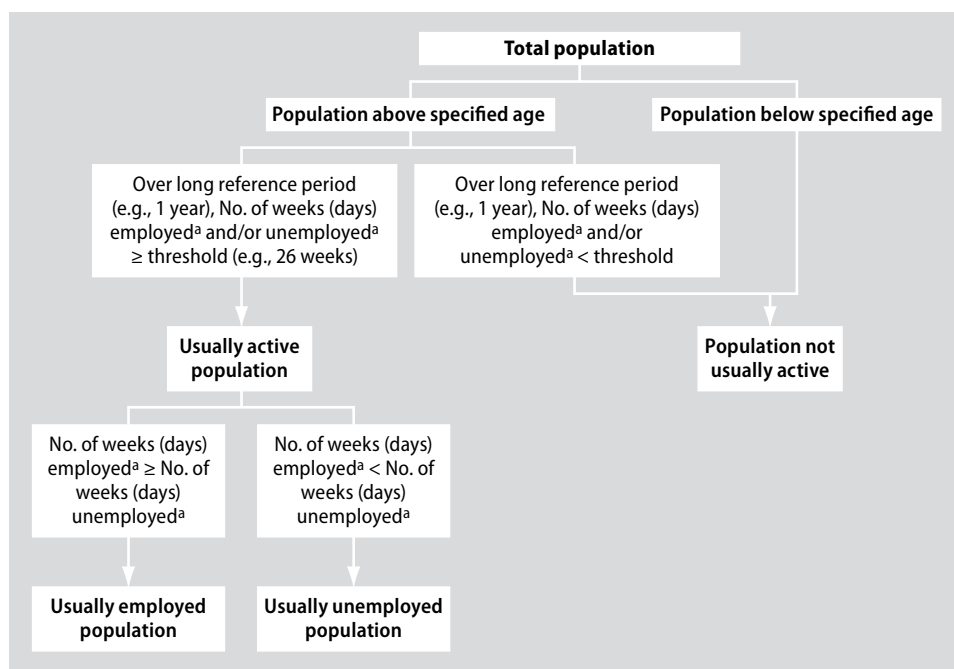
a As defined for current activity.

possibility for making recall easier is to fix the reference period by referring to a national event that took place just before the census enumeration (if there is one).

254. The basic priority rules regarding economic and non-economic activities, as presented in the section on current activity (see para. 169), state that within the reference period (*a*) economic activities take precedence over non-economic activities, and (*b*) within economic activities, employed activities take precedence over unemployed activities. Some people are both employed and unemployed during the same day, week or month as they are looking for other work. The above rules determine the order of the questions. Questions to determine the length of all types of employment activity should always be asked first, followed by questions to determine the length of experiences of unemployment and finally by questions about the length of any non-economic activities. For usual activity the rules also determine the way in which individuals are to be classified to the main groups:

(*a*) If a person spends more time as economically active (employment and/or unemployment) than as not economically active over the 12 months, the person is considered to be "usually economically active", and the converse holds for the person to be "not usually economically active";

(*b*) If an economically active person spends more of the active time as employed than as unemployed, then the person is taken as "usually employed", and the converse holds for the person to be "usually unemployed".

255. In applying the definition of usually active, two procedures may be followed at the time of enumeration to determine the usual activity status of each person. One is to interpret it as the status that prevailed over most of the 52 weeks or most of the 365 days during the specified reference year. Another is to set a specific number of weeks or days as the cut-off point and classify anyone with at least that many weeks or days of economic activity as the "usually active population" (Hussmanns, Mehran and Verma, 1990, chap. IV, sect. 3).

256.   Another decision in the measurement of usual activity concerns the unit of counting to be used, that is, days, weeks or months, to record a person's activity and ultimately derive the major usual activity categories. A measure based on days would be very good if it could be obtained, but it would be very unlikely to achieve a reasonable degree of accuracy. Respondents have great problems remembering the number of days they spend on any activity in a reference month and certainly will not be able to give accurate estimates for the last 12 months. In principle, the choice of unit is an issue of precision of measurement (just as "kilometre" is a less precise measurement of distance than "metre"), as follows:

(*a*)   Weeks may represent a higher degree of precision, which may be possible and would warrant testing;

(*b*)   Months are the easiest but would be a very approximate measure as activities may change significantly during a month.

However, depending in part on question formulations, reference to a week or a month may be interpreted as a whole or full week or month by the respondent. Thus it may be better to get an approximate (and possibly rounded) estimate of days than an estimate for a type of period, which one would not want. It is important to note that the main activity status of a person could change substantially depending on the choice of measurement unit.

257.   Moreover, in countries where employment is mostly of a regular and continuing nature and where a week of employment means by and large a week of full-time employment or, at any rate, employment for a major part of the working time, the question about main activity status may well refer to weeks of employment or unemployment. However, in countries where employment is largely of an irregular nature and where a week of employment does not generally mean a week of full-time employment or even employment for a major part of the working time, the question about main activity would better refer to days of employment or unemployment.[2]

258.   In determining the length of periods of employment, unemployment and inactivity, respondents may be asked (*a*) to give values for the entire long reference period, or (*b*) to give those values on a month-by-month (or season-by-season) basis. The latter approach (*b*) is likely to have less risk of memory recall and telescoping errors. However, it is more complex and takes up more space on the census questionnaire, which could be a major concern for one-page landscape formats. Another approach is to formulate the questions in terms of the jobs held and their durations.

2 Further information on the topic can be found in Hussmans, Mehran and Verma (1990), chap. IV, sect. 3.

D.   Usually active: review of questions used in national censuses

259.   The following review shows examples for activity over a long reference period and shows both predominant activity and usual status approaches:

(*a*)   Example C.1 is a question on any employment activity over the previous calendar year and is taken from a questionnaire in which most of the questions on economic activity were on a current basis (last week);

(*b*)   Examples C.2 and C.3 are questions on predominant activity and are taken from questionnaires in which some of the questions on economic activity were on a current basis (last week);

(*c*)   Example C.4 is an example that has questions on both predominant and usual activity. There are no questions on a current activity basis;

(*d*)   Example C.5 is an example of questions only on usual activity status.

Example C.1

C.1.1 (*a*) **Last year (1999), did ... work, even for a few days, at a paid job or in a business or farm?**

    1 ☐ Yes   **2** ☐ No → **Skip to X.X**

(*b*) **How many weeks did ... work in 1999? Count paid vacation, paid sick leave and military service.**   **Number of weeks →** ☐☐

(*c*) **During the weeks WORKED in 1999, how many hours did ... usually work each week?**   **Number of weeks →** ☐☐

260.   The question in example C.1 came after the main set of economic characteristics questions, which referred to current economic activity and which included reasonable detail on what to include as "work". Thus respondents had already been exposed to those concepts. The question in C.1.1 is a simple one that reflects the recommendation that "any work" should be measured for those who are employed; the number of weeks worked is useful but not sufficient for determining the usual status. The reference to "even for a few days" in question C.1.1 (*a*) may avoid the interpretation "full weeks" in question C.1.1 (*b*). The set of questions in C.1.1 is, however, limited, as it does not include the number of weeks of unemployment, which is necessary to identify those who are usually active and usually unemployed.

261.   Example C.1.1 also includes a simple measure of the number of weeks worked with an additional question on the usual working time, without repetition of the prompts

Example C.2

Section 12. Economic activity

C.2.1   **What did ... do most during the past 12 months—for example, did you/he/she work, look for a job, keep house or carry on some other activity?**

    ☐ **1** Worked → **Go to Q.X.X**

    ☐ **2** Had a job but did not work → **Go to Q.X.X**

    ☐ **3** Looked for work

    ☐ **4** Wanted work and available

    ☐ **5** Home duties

    ☐ **6** Attended school

    ☐ **7** Retired

    ☐ **8** Disabled, unable to work

    ☐ **9** Other (please specify _____)

    ☐ **10** Not stated

C.2.2   **Did you/he/she do any work at all in the past 12 months? Include work at home, for example, piece work, decorative stitching, handicraft, sewing, etc.**

    ☐ **1** Yes → **Go to Q.X.X**

    ☐ **2** No

    ☐ **3** Don't know

C.2.3   **Have you/he/she ever worked or had a job?**

    ☐ **1** Yes → **Go to Q.Y.Y**

    ☐ **2** No → **Go to Q.Y.Y**

C.2.4   **How many months did you/he/she work in the past 12 months?**

    Number of months

    1 2 3 4 5 6 7 8 9 10 11 12

    ☐☐☐☐☐☐☐☐☐☐☐☐

Example C.3

C.3.1 What were you mainly doing in the last seven days?		**C.3.2 What have you mainly been doing for the past 12 months?**

C.3.1		C.3.2	
Worked—Paid non-seasonal	[1]	Worked—Paid non-seasonal	[1]
Worked—Unpaid non-seasonal	[2]	Worked—Unpaid non-seasonal	[2]
Worked—Paid seasonal	[3]	Worked—Paid seasonal	[3]
Worked—Unpaid seasonal	[4]	Worked—Unpaid seasonal	[4]
On leave	[5]	On leave	[5]
Unpaid work on household holding or business	[6]	Unpaid work on household holding or business	[6]
Unemployed and seeking work	[7]	Unemployed and seeking work	[7]
Not seeking work but available for work	[8]	Not seeking work but available for work	[8]
Full-time housewife/homemaker	[9]	Full-time housewife/homemaker	[9]
Full-time student	[10]	Full-time student	[10]
Not available for work for other reasons	[11]	Not available for work for other reasons	[11]

Go to X.X (bracket spanning items [7] through [11] of C.3.2)

C.3.3 Employment status

Since August 1999 what have you been mainly:

☐ An employer
☐ An employee
☐ Self-employed
☐ An unpaid family worker

C.3.4 Occupation

What was your main occupation for the last 12 months?

Write name of occupation and enter code

(0)	(0)	(0)
(1)	(1)	(1)
(2)	(2)	(2)
(3)	(3)	(3)
(4)	(4)	(4)
(5)	(5)	(5)
(6)	(6)	(6)
(7)	(7)	(7)
(8)	(8)	(8)
(9)	(9)	(9)

C.3.5 Industry

What kind of main product or service is/was produced where you work?

Write name of industry and enter code

(0)	(0)	(0)
(1)	(1)	(1)
(2)	(2)	(2)
(3)	(3)	(3)
(4)	(4)	(4)
(5)	(5)	(5)
(6)	(6)	(6)
(7)	(7)	(7)
(8)	(8)	(8)
(9)	(9)	(9)

on what to include. However, there is no attempt to measure the number of weeks of unemployment, and there is therefore no possibility of measuring usual economic status.

262.   In the census questionnaire from which example C.2 was drawn, the questions on activities in the past 12 months were followed by more detailed questions in respect of the current economic activity in the last week (not shown). The example provides few probes or prompts as to what to include as "work", and the use of months worked is a less precise measure than weeks worked.

263.   The set of questions in example C.3 was drawn from a landscape layout. One question on current activity was asked, followed by the same question with the long reference period (the past 12 months). Questions C.3.1 and C.3.2 ask "What was … mainly doing …?" during the two different reference periods. The long reference period was then used for the status in employment, occupation and industry questions, but the question on status in employment (C.3.3) has a different description ("since August

Example C.4

Questions are asked of persons 10 years and above		
C.4.1 What work was ... usually doing during the last 12 months? 1. Agriculture 2. Salary/wage 3. Own eco. enterprises 4. Extended eco. 5. Job seeker 6. Household work 7. Student 8. No work	**C.4.2 How many months did ... work during the last 12 months?** *(Specify approximate number of months spent in each category)* 1. Work done for . . . . (months) 2. Extended eco. . . . . (months) 3. Work seeking . . . . (months) 4. No work . . . . (months) **Total 12 months**	**C.4.3 What work did ... usually do?** **(Occupation)** *(Describe the actual work done)*
C.4.4 Where did ... do the work? **(Industry)** *(Name the establishment or the organization where worked)*	**C.4.5 What was ...'s employment status?** 1. Employer 2. Employee 3. Own-account worker 4. Unpaid family worker	**C.4.6 What was ...'s reason for usually not working during the last 12 months?** 1. Student 2. Housework 3. Aged 4. Pension 5. Physically and mentally handicapped 6. Sickness or chronic illness 7. Others . . . .

1999" rather than "for the past 12 months") although the periods are the same. It is recommended that the same wording be used in all questions to avoid any confusion for the respondent.

264.   For current activity the "mainly doing" approach is not recommended, since the respondents do not know the priority rule and "mainly" would be interpreted differently by different groups of people. It is assumed that the manuals and training gave details of definitions for "work", "leave", "unemployed", "available" and so on. Details concerning the success of the "mainly doing" approach are not available. However, problems of comparability with respect to the statistics on the other economic characteristics would arise if the approach illustrated in example C.3 were adopted and the previous census used the more common current activity questions. Therefore, when considering this type of approach, it will be necessary to pre-test it well and be aware of likely problems.

265.   In examples C.4 and C.5, sufficient details are collected on the number of weeks or months in each of the three statuses (employed, unemployed and inactive) to be able to determine whether the person is "usually active, usually employed"; "usually active, usually unemployed"; or "usually inactive".

266.   In both C.4 and C.5, interviewers and respondents are expected to apply fairly complex rules to determine whether the person was employed, unemployed or inactive within each week or month and then aggregate those periods. That task is not trivial, and there may be problems with the quality of reporting that would need testing as well as post-enumeration checks.

267.   In example C.4, questions are asked to measure both the predominant activity and the comprehensive "usual" status. In respect of the usual status, questions are asked

Example C.5

C.5.1	**Activities in the last 12 months and their durations (in weeks):**	**Activity**	**Weeks**

C.5.1   Activities in the last 12 months and their durations (in weeks):

   A. Paid employment

   B. Employer

   C. Own-account worker

   D. Contributing to family enterprises (unpaid family worker)

   E. Available/Seeking work

   F. Non-economic activities (student, housekeeping, income recipient/pensioner, unable/too old to work, child not attending school, other non-economic activity)

If not engaged in any activity write '00' for that activity.

The sum of all durations should be 52 weeks

	Activity	**Weeks**
A.	Paid	☐☐
B.	Emp.	☐☐
C.	O.A.W.	☐☐
D.	Family	☐☐
E.	Available	☐☐
F.	N.E.A.	☐☐

C.5.2   Principal occupation or kind of work done in the last 12 months:

If C.5.1 A to D '00' go to P X.X

State clearly exact occupation or kind of work done during the last 12 months.

E.g., Paddy cultivator, tea plucker, accounts clerk, mathematics teacher, vegetable seller (retail), domestic helper, etc. If the person had more than one occupation/work, give the principal occupation/work on which most time was spent.

If C.5.1 A to D '00' go to P X.X

_____

_____

_____

☐☐☐☐

C.5.3   Industry, business or service:

Describe the kind of products made or services rendered at the work place of the principal occupation/work.

Write the name of the work place.

E.g., Paddy cultivation, tea cultivation, education department, textile shop (retail), garment factory, etc.

_____

_____

_____

☐☐☐☐

C.5.4   Employment status and sector:

   1. Government employee

   2. Semi government employee

   3. Private sector employee

   4. Employer

   5. Own-account worker

   6. Unpaid family worker

C.5.6   Principal non-economic activity in the last 12 months:

If C.5.1 F is '00' go to P Y.Y

   1. Student

   2. Household work

   3. Income recipient/Pensioner

   4. Unable/Too old to work

   5. Child not attending school

   6. Other non-economic activity

---

on the approximate number of months employed (two categories called "work done for" and "extended eco." that are presumably described in more detail in the interviewers' manual), unemployed ("work seeking") and the inactive ("no work"). The time spent in those four categories adds up to 12 months. As already stated, the use of approximate number of months is a relatively imprecise measure, but reflects the reality of memory recall over the 12-month period.

268.   Although it could be assumed that the questions on status in employment, occupation and industry relate to the main job in the last 12 months, that is not stated and should have been included in the questions.

269.   The basic economic activity question in example C.5 combines status in employment with unemployment and non-economic activities. However, it enquires about the number of weeks spent in the last 12 months on each of those statuses so that there is no confusion on the part of the respondent as to what order of priority to assign to a specific category. The responses on "weeks" can then be used as a basis for determining the main activity status and on that basis to determine the other characteristics. With respect to the category representing unemployment, the criterion being applied—available for work, seeking work or both (see paras. 213 and 216)—will need to be clarified. Although it could easily be assumed that all the questions have a 12-month reference period, the questions on industry and status in employment do not state the reference period and could be interpreted as the industry for the present job rather than for the main job.

Example C.6

Hypothetical questions, recommended for testing and use in countries

C.6.1 During the last 12 months up to the end of last month did you:

Interviewer: Prompt on each group with examples and tick if yes.

☐   Do any work of any kind on your own or the family farm, cattle post/kraal, or other agricultural holding, e.g., chicken house, vegetable plot, etc.?

Prompts: The work can be for cash income or for your own food. Include any farm or home construction or major maintenance

☐   Conduct any type of business big or small, e.g., brew beer, sell cakes, make hats/clothes, etc.?

☐   Provide unpaid help in a family business of any type?

☐   Catch/collect fish, shells or other sea or river products? Collect any water or firewood?

☐   Make anything from your farm or natural products for sale or for your own use?

☐   Do any type of wage or salary job—full-time or part-time, temporary, casual, piecework or permanent?

C.6.2 How many weeks in the last 12 months were you doing any of these activities in total? (*AA*) ☐☐

Interviewer: Make sure that respondent includes paid leave or paid training as work. Also include short absences when respondent is sick or temporarily absent from work. Record as zero if no work done.

If AA = 52, skip to next section.

C.6.3 When not working were you available for work if given the opportunity?

Yes ☐   No ☐

C.6.4 How many weeks were you available for work but not working in the last 12 months? (*BB*) ☐☐

Total weeks *AA + BB*: ☐☐

Interviewer: If AA + BB = 52, skip to the next section.

C.6.5 What were you doing when not working?

Student	1 ☐
Housework	2 ☐
Not working and disabled	3 ☐
Sick	4 ☐
Retired/aged	5 ☐
Income recipient	6 ☐
Other	7 ☐

E.   Usually active: suggestions for enhancing the effectiveness of the questions

270.   Example C.6 is a hypothetical example that illustrates many of the points that have to be considered in designing questions on usual activity.

271.   The sequence of questions in example C.6 is complex but would give labour policy analysts and national accountants quantitative measures of the number of weeks in each usual economic activity category as well as the allocation of the population to the overall usual activity categories. The number of weeks unemployed is a particularly useful measure in countries with a seasonal unemployment problem. The three categories (usually employed, usually unemployed) and usually not economically active can be derived later during tabulations, from the following algorithms; the decision-making criteria are as follows:

(*a*)   Usually economically active and usually employed = AA + BB equal to or greater than 26 weeks and AA > BB;

(*b*)   Usually economically active and usually unemployed = AA + BB equal to or greater than 26 weeks and AA < BB;

(*c*)   Not usually economically active = AA + BB less than 26 weeks,

where AA refers to the number of weeks employed and BB refers to the number of weeks when available for work (that is, unemployed). This sequence also makes it possible to classify the population according to different patterns of activities during the year, by mixtures of employment, unemployment and inactive. For many users, statistics describing such patterns may be even more interesting than statistics by usual activity.

272.   A series of prompts is used for the first question (C.6.1). This is in line with the earlier recommendations and covers forms of employment that are usually difficult to identify. Such prompts would need to be varied to suit the situation in a particular country. An alternative would be to use a prompt card, similar to the one suggested in figure III for current activity, but referring to the last 12 months.

273.   The questions have been worded "Did you …" or "Were you …" to encourage interviewing the respondents in person. However, it is recognized that interviewers often will have to change the question to "Did … " or "Was …" when interviewing through another informant. It is better generally not to assume the use of informants and to word all questions with "you". Interviewers should be able to manage the change necessary for the informant situation.

274.   The unemployment question (C.6.3) uses the availability concept, but it should be subject to testing, and "actively seeking work" can be substituted if required. A question of this type would need thorough pre-testing. One major problem will be space on the questionnaire, but that concern needs to be weighed against the value of the information that would be obtained.

275.   The basic categories of usual economic activity should not be combined with those of other variables, such as status in employment, as those categories will most likely change over the course of 12 months for some groups of people with changing jobs. Respondents would then need to make a decision on the main status (or type of work or type of activity) without the necessary guidance. If it is thought essential to have the status in employment or other variables for the usual activity, it is suggested that the necessary information be collected with separate questions, as in example C.5.

276.   There is a possibility of introducing more complex questions. Once questions on separate activities are posed, as in C.6.1 above, there will often be demands, from national accountants in particular, for recording the number of weeks spent on each activity. In addition to taking time and space, obtaining such data can be very difficult as often sev-

eral activities are done in any one week or month and it would take considerable time and effort to discuss and work out a major activity in terms of time spent each week on each activity. However, it may be possible to allow a "yes/no" answer in a separate column for each of the prompts in C.6.1. This could supply a measure of the effect of each prompt and assist in achieving comparability if a briefer question had been asked in a previous census.

277.   There are other possible questions that could be asked and recorded to extend the basic usual activity questions, for example: "Which is the major work activity in terms of time spent?"; "For all agricultural activities, were they mostly for cash or home consumption?" However, it is strongly recommended that more detailed measurements such as those, which would be very useful for many types of analyses at the national level as well as for smaller regions, be collected by household surveys or by detailed case studies, since in such surveys more time is usually available with better-trained interviewers, and detailed checks can also be made. Trying to collect too much information could jeopardize the overall accuracy of the census. Some additional detail on current activity, where the reference period is shorter and there is less of a problem of multiple activities, could be obtained.

278.   It is strongly recommended that initial questions on usual activity be relatively restricted until experience is gained with the concepts.

279.   Example C.7 presents a hypothetical approach towards the collection of data on predominant activity status, which also provides an assessment of usual activity status. In a situation in which the emphasis is on current activity but a rough measure of usual activity is required for cross-analysis, the question could be placed after the questions on current activity. Its advantage is that, without requiring the respondent to give a specific length of time during the long reference period for the different activity statuses, one can arrive at classifying the population into the usually active/inactive statuses. However, it assumes that respondents are already familiar with the notions of employment

Example C.7

Hypothetical question, recommended for testing and use in countries		
	What was each person's predominant activity status during the last year?	

P E R S O N N U M B E R	**Predominantly active**	(*a*) **All the time**
		    1. Employed
		    2. Unemployed
		(*b*) **Most of the time (26 weeks or more)**
		    3. Employed
		    4. Unemployed
		    5. Employed or unemployed

	Not predominantly active	(*c*) **Not active all/most of the time**
		    6. Student
		    7. Home worker
		    8. Disabled/sick
		    9. Retired/income recipient/aged
		    10. Other reasons

1.	Code _____	
2.	Code _____	

and unemployment from the earlier questions on current activity or that the interviewer could explain those concepts to them. It may be necessary to do some prompting for employment activities that are difficult to recognize. Note, however, that the information collected is not sufficient to determine if the usually active person is usually employed or usually unemployed.

280.   Again, it is very difficult to ask a simple question to get good measures of usual activity. Given the pressure on space on census questionnaires, some countries might accept such a question or something similar but they should be aware of its serious limitations. In any case, it would need to undergo thorough testing before trying to use it.

F.   Usually active: issues of measurement

281.   Only two countries are known to have asked usual activity questions in censuses in the past, and only a few countries have tried them in surveys. Thus experience with the topic is limited. It is strongly recommended that the question(s) be introduced only after thorough pre-tests and that current activity be the major target for the economic questions. This is also important for reasons of comparability with other censuses and surveys in the country. Most characteristics of the economically active population (for example, occupation, industry, status in employment) should thus be collected only for current activity. However, in some countries, where usual activity has already been used as the major concept in prior censuses and where the users of the statistics are satisfied with the results, they should continue with that measure.

282.   All countries should at least consider the possibility of including a usual activity question. Census planners should work closely with their national accountants and workforce or labour policy colleagues to establish a minimum desirable need for measures of usual activity. It is strongly recommended that any question on usual activity be put separately, either before or after questions on current activity. Respondents will be very confused if they are asked to switch from a week or month reference period to a year period and then back to a week or month. It would, however, help if such questions came after a block of current economic activity questions that probed on the categories and explained them.

283.   Landscape layout works best with relatively simple questions and, as noted above, asking a simple question on usual activity is very difficult. It may nevertheless be possible to design a question about "main usual activity" with interviewer training and manuals stressing all the problem categories and the rules to apply. Some notes on the questionnaire would definitely help. It may also be possible to ask for the number of weeks employed and unemployed in the last 12 months. The number of weeks not economically active should be calculated as a residual. The usual activity prompt card should differ from the current activity prompt card, for example when activities are seasonal and certain activities dominate at census time and others during other months of the year. Some countries may wish to try such questions to obtain a rough measure of usual activity.

284.   Pre-testing will be particularly important for this topic to validate the following:

    (*a*)   The use of the calendar year or 12 months as the reference period;

    (*b*)   The use of weeks or months (or even days) as a counting unit;

    (*c*)   The use of a prompt card;

    (*d*)   The use of actively seeking work or available for work as unemployment questions;

    (*e*)   General coverage of difficult activities, such as unpaid work, part-time work and the like.

285.   Pre-tests must be well designed to be effective. A quantitative method of testing would be to use one question for half a large sample and an alternate question for the other half and then analyse the differences in the results for each half. Two matching samples plus a thorough knowledge of anything likely to cause other differences would be necessary. Another approach would be to try the alternative questions on matching representative groups of the population and then subject all the respondents to a more in-depth re-interview on their activities over the reference period and on their problems with the question(s) as initially posed. That activity would also be followed by quantitative analysis. The objective of any test must be clear and the results must be well documented for future planning.

Part Three

Measurement of characteristics of jobs, establishments and persons

Chapter VI

Descriptive characteristics
for the main job

286.   Once the currently economically active persons have been identified, the remaining economic topics refer only to that subset of the total in-scope census population. Two major issues must be decided: which job or jobs associated with the person should be the focus of the questions, and whether to put those questions to only the currently employed or to both the currently employed and the unemployed. The relevant characteristics of a job include status in employment, occupation, and place of work.

287.   **The reference job.** Employed persons can and do sometimes have more than one job in any reference period. In general, the reference job should be chosen as the job on which most time is spent during the reference period, although other criteria such as the job that provides the most income or with the greatest number of regular hours are used in some countries. The decision taken should be consistently applied. Selecting the reference job is usually not difficult for the short reference period for current employment activity. However, problems may arise with persons temporarily absent from a main job in the reference period and with selecting a previous job for the unemployed. *Principles and Recommendations, Revision 2*, recommends that jobs from which the person is temporarily absent should not be considered as the main job if the person is employed and at work in another job during the reference week. It is not stated, but implied, that if the person does not have any other employment in the reference period, then the job from which he or she is absent should be the reference job. It has usually been found that most economically active persons do report about any job from which they are temporarily absent and expect it to be recorded. That applies particularly to wage earners on leave, and the like, but is also very common for the self-employed. For those identified as unemployed, the most common practice is to make reference to the last job or to the main type of job over a recent, longer period, for example, last year.

288.   Census planners need to review the issue of the reference job for their own country and decide what procedure to follow. If it is decided to use the job from which the respondent is temporarily absent as the reference job in the cases where the person is at work in another job during the reference period, then a note to that effect will be needed on the questionnaires and/or manuals. If secondary activities are also being recorded, interviewers should be instructed to include any other current activity as a secondary activity. That will also affect the way working time for the main job is recorded, and clear guidance needs to be provided. Systematic exclusion of temporarily absent jobs could cause major problems in respect of comparisons with other employment data, such as those from establishments, as well as some anomalies. Such problems can arise particularly when a census is conducted during school holidays, as is the case in some countries, for example, when teachers are given the jobs temporarily as enumerators and supervisors for the census.

289.   It is essential that once the main job is identified all characteristics should refer to the same main job in all subsequent questions. The questionnaire design must be such that it will not confuse respondents or interviewers on that point. When any secondary

activities are identified, they must be clearly separated and placed at the end of the set of questions relating to the main job, to avoid any confusion. This is usually not difficult with current activity, but can be quite difficult with usual activity.

290.   **Employed persons only, or employed plus unemployed.** *Principles and Recommendations, Revision 2*, recommends the collection of employment characteristics for the unemployed as well as for the employed. For clarity, such questions for the unemployed should be a separate block of questions from those for the employed. However, owing to space considerations on census questionnaires, in some cases only a simple note has been inserted to tell the enumerators to ask the same questions for the unemployed as for the employed, but for the unemployed to refer to the last job they held. That practice has led to some confusion and relies on uncertain interviewing skills to change the question for differing situations.

291.   The number of unemployed recorded in censuses is often relatively low in developing countries and is concentrated in the younger age groups in urban and peri-urban areas (see chap. IV, sect. C). Very few such persons have had previous job experience, and what little experience they have had is usually of limited relevance for the longer-term economic or labour-market policy planning and implementation for which census results are most useful. In some countries, a "job they are seeking" approach has been adopted for the unemployed. However, it has been shown that few of the unemployed have a clear idea about the exact type of job they are realistically seeking owing to their limited experience. Thus very often the most appropriate answer to a question on the characteristics of the job they are seeking would be "any job", and that response is also of limited use for the formulation or implementation of economic or labour market policies. For the above reasons it is generally thought that in censuses, questions on the characteristics of the last job of unemployed persons should be severely limited. Such information is better collected in specialized household surveys where more in-depth questioning is possible.

A.   Status in employment

292.   Status in employment is one of the more common economic topics in population censuses and one of the most important. Knowing the number of persons working for others ("employees") or for themselves ("self-employed") is very important for the understanding of the structure and functioning of any labour market. A question on status in employment was asked by as many as seven in eight countries in the 2000 round of censuses, although in some of those countries the information was collected only partially with mixed questions. The most recent and thorough discussion of the topic took place at the Fifteenth International Conference of Labour Statisticians in 1993. The resolution concerning the International Classification of Status in Employment (ISCE-93) (ILO, 1993a), adopted at that Conference, provides a much stronger conceptual base for the classification and deals with many of the problem situations that arise when using it. In broad terms, the resolution classifies jobs according to the type of economic risk the job entails and the control that the person in that job has over the establishment and other workers.

293.   The Classification is generally applied to the status of the jobs of employed persons only. If an employed person has more than one job and only one job is to be recorded and classified (as is the case for most censuses), then it should be the main job, which is generally the one in which the most time was spent during the reference period.

1.   Status in employment: operational definition

294.   "Status in employment" refers to the status of an economically active person with respect to his or her employment: that is, the type of explicit or implicit contract of employment with other persons or organizations that the person has in his or her job.

The basic criteria used to define the groups of the classification are the type of economic risk, an element of which is the strength of the attachment between the person and the job, and the type of authority over establishments and other workers that the person has or will have in the job.

2.   Status in employment: requisites for applying the definition

295.   A major distinction exists between paid jobs and self-employment jobs. The key points are that (*a*) paid jobs have some form of supervision and an agreement of some sort regarding the amount of payment in cash or kind; while (*b*) income from self-employment jobs depends solely on the profits (or expectation of profits), where own consumption is considered part of the profits from the activity undertaken or the enterprise controlled.

296.   The International Classification of Status in Employment specifies six broad groups:

(*a*)   Employees (including apprentices);

(*b*)   Employers;

(*c*)   Own-account workers;

(*d*)   Members of producers' cooperatives;

(*e*)   Contributing family workers;

(*f*)   Workers not classifiable by status.

297.   The groups are virtually the same as those that were used before, except that behind the change of terminology from "unpaid family workers" to "contributing family workers" there is a modification of the definition of that group.[1] The first two sub-groups of the self-employed jobs are differentiated according to whether the enterprise that the person singly or jointly controls has paid employees on a continuous basis (employers) or not (own-account workers). The distinction is based on the presence of employees on a continuous basis: that is, an employer must be employing and paying someone on a continuous basis, not only on an ad hoc basis and for a short period as, for example, to solve a particular problem or address a short-term increase in the workload. Payment may be in cash or in kind (such as food, clothing or accommodation, or, in some cases, training only).

298.   Contributing family workers are by definition unpaid. They usually live in the same household and are related to other self-employed family members who actually control the enterprise as either employers or own-account workers in the enterprise. However, the restriction on living in the same household can be relaxed (for example, relatives' households).

299.   The last group of self-employed comprises "members of producers' cooperatives" where the members jointly determine the organization of work and the distribution of proceeds of the enterprise. Members of producers' cooperatives may be rare in many countries, and the category is not used in censuses in countries where they are thought to be unimportant numerically. If used, care must be taken to ensure that a person merely selling or buying goods through a sales or retail cooperative is not mistakenly categorized as a member of a producers' cooperative. The person must actually work as a member of the cooperative producing goods or services. Employees of any form of cooperative should be excluded from this group and be classified as employees.

300.   The ICSE-93 also provides for a number of particular groups, which are additional, more detailed, status in employment subcategories. Some of them may be important in the country, and it may be possible to identify them through separate response categories on the census questionnaire. It is strongly recommended that during the preparations for

1 When the ICSE-93 was adopted, the criterion of "work without pay" previously used in respect of "unpaid family worker" was dropped and replaced by the criterion of "degree of commitment to the operation of the establishment" (which needs to be lower than that of the head of the establishment) for contributing family workers. Accordingly, in its resolution on the measurement of employment-related income (paras. 75-77), the Sixteenth International Conference of Labour Statisticians recommended the imputation of an income for contributing family workers (ILO, 2000).

the census, the need to include such subcategories be discussed with potential users of the census results and that their feasibility be tested carefully. Among the possible particular groups one may expect users to be especially concerned about are: owner-managers of incorporated enterprises, who in many respects resemble employers; casual workers and seasonal workers; and outworkers (home workers).

3.   Status in employment: review of questions used in national censuses

301.   Two broad types of questions on status in employment have been included on census forms. That characteristic is either combined in a question with other characteristics, such as activity status or institutional sector, or presented on its own. Five such examples are provided below, as follows:

(*a*)   Status in employment combined with the determination of activity status (example D.1);

(*b*)   Status in employment combined with institutional sector (examples D.2 to D.4);

(*c*)   A question on status in employment on its own (example D.5).

302.   Example D.1 comes from a landscape format questionnaire and exhibits the common problem of that format, limited space and cryptic instructions. The question does suffer considerably from the combination of the determination of current activity and status in employment in one question. It was stressed in the chapters on economic activity status that it is undesirable to combine topics in the same question. However, it has been relatively common to combine ICSE categories with the determination of economic activity status, whereby the status in employment categories are used as prompts. That practice has usually detracted from the accuracy of the data on both topics since, as noted in the chapter on measuring employment, a single question does not provide a satisfactory measure. For questions that ask "Did you do any of the following in the last seven days?" it is better not to focus attention first on paid employment. Several tests have shown that if paid employment is the first option some respondents then think the whole series of questions is about paid employment. This leads them to answer "no", not considering whether their situation can be described by other status in employment groups. Such respondents are recorded as not employed and are therefore missing from all the following questions on occupation, industry, and so forth.

303.   The major point of interest in the question is, however, the first response category in example D.1 called "mlimi". This word means "working on own farm" in the

Example D.1

D.1.1   Type of activity last week
What was … doing last seven days? Enter code.

Active		Inactive	
Mlimi	= 01	Non-worker (never worked before and not seeking work) = 09 →	**If female skip to X.X; otherwise go to the next person or to part Y.**
Employee	= 02		
Family business worker	= 03	Home worker = 10 →	
Self-employed	= 04	Student = 11 →	
Employer	= 05	Other = 12 →	
Unemployed			
Worked before, seeking work	= 06		
Worked before, not seeking work	= 07		
Never worked before, seeking work	= 08		

Note: Mlimi = Working on own farm.

local language. Similar categories have been used in a few other countries, and this technique can have major advantages. Respondents in countries with large traditional agricultural sectors rarely recognize themselves as being self-employed, employers, own-account workers or in business in either the English wording or in local translations. Nevertheless, they usually have no problem recognizing that they were working on their own farm or ranch, or the like, but it will be necessary to ask another question to determine if they regularly employ anyone or not. Without that follow up, the employer category is bound to be underestimated in favour of the own-account worker category with this kind of question.

304. No attempt is made to split the household members into own-account worker, employer and contributing (unpaid) family helpers in two of the main categories, that is, mlimi and family business worker. An advantage of having separate categories for those forms of employment is that one can then get a relatively clear count of those who are self-employed outside agriculture and of the number of contributing family workers in such non-agricultural business activities. In the past, trying to make that distinction within a family often caused problems with lengthy discussions about whose farm it was, who owned the land, what to do if the head of household or farm owner was absent, and so on. In reality such farms are family farms with all members contributing and sharing the risks and rewards of their labour. In those cases the separation into such categories as own-account worker and contributing family worker is not correct from a conceptual point of view, and the results from previous censuses have often been difficult to interpret owing to the different practices in the field.

305. In many countries the common meaning of "business" does not include traditional agricultural activities. However, commercial agriculture is understood to be a business, as is the buying and selling of agricultural products. Those interpretations are also suitable as the basis for economic policies. Thus, in example D.1, business activities "with employees" and "without employees" would be a better formulation on the questionnaire than using ICSE-93 terminology. That illustrates the need to adjust to local situations and use local terminology. Examples should be given if necessary.

306. Combining status in employment with institutional sector in D.2.1 is an approach that seems to work well for a census. The question is a standard status in employment question with paid employees split into two government categories and private for a sector classification. "Unpaid worker in family business or farm" is listed as part of working for someone else, even though the corresponding ICSE-93 category is part of the self-employed. It may have been shown in testing that respondents recognized the category more readily under the "worked for someone else" heading. For tabulation and publication of the results, however, they may be easily, and should be, regrouped. The question relies on ICSE terminology; and use of local wording might have been more effective, as mentioned in D.1.

Example D.2

D.2.1 Were you self-employed or working for someone else in your (main) job last week?

Self-employed

☐ With paid help (employer)
☐ Without paid help

Worked for someone else

☐ As Government employee
☐ As employee of a foreign Government
☐ As employee of private company/person
☐ As unpaid worker in family business/farm
☐ Not stated

307.   The question provides more of the sector categories than usual. It assumes that respondents know whether their workplace is a "private, for profit" company or a "private, not for profit" organization. It is strongly recommended that respondents' understanding of any such categories be thoroughly tested before being introduced. Government employees as a group are separately identified; but "self-employed" is not differentiated by "with" or "without" employees. If that distinction were made, it would be possible to separate employers from own-account workers, which has policy applications in enterprise promotion and in equal opportunity and gender issues. Instead, the self-employed are categorized by whether their business is incorporated or not. Such categories have other uses (including as a component in the identification of informal sector enterprises—see chap. VII, sect. C, on the informal sector), but the full ICSE-93 classification could not be produced from the response categories presented in example D.3.1.

308.   In question D.4.1, the institutional sector of employment is also combined with status in employment, but with a simple breakdown for employees only. Three categories of institutional sector are used. Proxy respondents will most likely know if the person about whom they are reporting is in government or in the private sector. It may, however, be difficult for them to determine if the individual is in "semi-government". Terms like "own-account worker" will generally have to be explained so that the distinction from employer is clear, and the reference to "unpaid family worker" could be updated to reflect current terminology.

309.   Example D.5 comes from a portrait format questionnaire with separate columns for each person. "Directors" have been identified separately, illustrating the above-mentioned point on owner/managers of incorporated companies (see subsection A.2). The other interesting variation is the separate identification of "doing home handicraft", which is presumably significant enough to warrant separate identification. That category would cover both the production of such goods for sale and the production for home use included within the SNA production boundary. To avoid any mix-ups, the category "family worker" should be qualified as unpaid. It could be made equivalent to the "contributing family worker" category of ICSE-93; otherwise it should be clearly explained in the field manual. To revert to the standard categories in the ICSE-93, the answers could be combined, at the data

Example D.3

D.3.1   **Was …** *Read list. Mark (X) one box.*

- ☐ Employee of PRIVATE-FOR-PROFIT company or business or of an individual for wages, salary or commissions
- ☐ Employee of a PRIVATE NOT-FOR-PROFIT, tax-exempt or charitable organization
- ☐ Local or territorial GOVERNMENT employee (territorial/commonwealth, etc.)
- ☐ Federal GOVERNMENT employee
- ☐ SELF-EMPLOYED in own NOT INCORPORATED business, professional practice or farm
- ☐ SELF-EMPLOYED in own INCORPORATED business, professional practice or farm
- ☐ Working WITHOUT PAY in family business or farm

Example D.4

D.4.1   **Employment status and sector**

1. Government employee	4. Employer
2. Semi-government employee	5. Own-account worker
3. Private sector employee	6. Unpaid family worker

Example D.5

D.5.1 Employment status

Part-time workers are classified as employee.

Self-employed includes a proprietor of unincorporated business or a person working on his/her own account.

☐ Employee

☐ Director of a firm or corporation

☐ Self-employed, employing others

☐ Self-employed, not employing others

☐ Family worker

☐ Doing home handicraft

---

processing stage, with the appropriate categories: the handicraft workers who do not employ anyone on a regular basis to be classified as own-account workers, and the directors classified as employees or employers, with the assumption that, for the director to be classified as employer, the directors' enterprises employ at least one person on a continuous basis.

4.   Status in employment: suggestions for enhancing the effectiveness of the questions

310.   The major points to note about a question on status in employment are that the ICSE categories as worded do not have to be given in the census and that the order of the ICSE categories does not have to be strictly followed. The aim is to be able to derive these categories from the replies to the question or questions. Otherwise, considerable effort may be wasted on training on the meaning of such terms as self-employment, own-account worker and so on. Interviewers in the field have often used those words as prompts to respondents who did not understand them. In those instances, time has been wasted with lengthy explanations, and the replies from respondents were still confused. For example, errors could ensue such as classifying a respondent as self-employed in government. It is far better to use locally recognized terminology.

5.   Status in employment: issues of measurement

311.   The topic of status in employment has been included frequently in the past, but it is strongly recommended that tests be carried out in the light of the recommended wider interpretation of economic activity and of some of the suggestions above. There are several groups of jobs on the margin between employee and self-employed, which pose problems for correct classification. Some of them are covered below. Census planners should identify those that are common in their country and decide whether they should be identified separately and how they should be classified. Some research is often necessary, and field manuals and the training of field staff should cover the commonly found situations. In some countries, the distinction is determined by how the taxation authorities treat those groups of workers.

312.   **Owner-managers of incorporated companies.** Owner-managers are usually paid a salary by the enterprise in which they or members of their family also own a controlling part of the shares. They may also receive a part of the profits through the shares they own or may purchase shares at a favourable price (for example, through stock options). They act as the top managers of their enterprise and thus exercise functions similar to those of employers. Since they receive wages in remuneration for their work and not for the capital investment, they are for national accounts purposes defined as employees. The International Classification of Status in Employment (1993) recommends that they be identified separately, if possible, to allow users of the statistics to choose whether to group them as self-employed or as employees, since other users of the statistics may tend to group them with the self-employed for the purpose of labour market analysis.

313.   **Outworkers.** "Outworker" is a common designation for a significant number of persons, particularly women. In many countries, outworkers usually work from their homes or at another location away from the factory, with some agreement with a factory or entrepreneur that supplies the materials and/or tools used for the job. The outworker agrees to complete the set task with the materials supplied by a set time for a set payment. When such written or oral contracts exist, the outworker is in effect a paid employee similar to a factory-based pieceworker. They have little control over their work and do not share either the risks or rewards of the entrepreneurial part of the activity. They should not be confused with other home-based workers who buy the materials to work with and bear the risk of selling the product to others.

314.   **Home-based workers.** Home-based workers are in effect self-employed, even though they may have a continuing and long-term relationship with both the suppliers of raw materials and the buyers of their end product. With no formal agreement or guarantee on the supply of raw materials and the sale of the product these workers are bearing the risk and gaining the reward of self-employment. Often other family members become involved in the activity either as partners or as unpaid helpers. The subcontracting of part of the work to others is also quite common. Research needs to be done in each country to determine the situation(s) of any home-based workers, the ways to identify such activities and, if necessary, the ways to differentiate paid employee work from self-employment work. That information must be stressed in manuals to ensure not only that those activities are recorded as economic activities, but also that all field staff classify the activities in the same manner as for status in employment.

315.   **Business partners.** Workers may enter into a partnership with another person either long term or for just a short-term contract for a particular job. Those arrangements should pose no problem. Since the persons in such an arrangement are sharing the risks and rewards of the business activity, each should be recorded as self-employed and classified as employers or own-account workers depending on whether they have paid employees on a continuous basis.

316.   **Commission workers.** Commission workers are often paid a small basic retainer and then a commission based on their sales. Usually the basic retainer in such cases is very small and the major component of the person's remuneration is derived from commissions on sales. Such persons should usually be recorded as self-employed. They are bearing most of the risks and gaining the rewards of the job. Another group of workers in a situation similar to that of commission-based workers are workers who rent such factors of production as land, a piece of equipment or a place of operation for operating a business (for example, taxi drivers, street hawkers, hairdressers, shoeshine boys). Those workers often see themselves as employed by the owners of the assets they rent. However, since they are bearing the risk of production and generating income, they should be considered as self-employed and not as employees. That distinction should be stressed in manuals, and illustrative examples should be used in training in countries where such work arrangements are common.

317.   In each country it is likely that there will be other categories that are either important or problematic, or both. Their classification should be based on the degree of risk to which the person is exposed for the major part of his/her remuneration from the job. If the person bears most the risk (and gets the reward) then the person is in self-employment. If the enterprise takes most of the risk and pays most of the remuneration then the person is a paid employee. Some additional marginal cases are reviewed in International Labour Office sources (see ILO, 1988 and 1993a; and Hussmanns, Mehran and Verma, 1990).

318.   It is not unusual for data users to want the census to provide them with subclassifications under the major general groups of ICSE-93. In addition to those already mentioned, some of the more common requests include the following:

(a)   **Paid employees.** They are classified by whether "permanent/regular" or "temporary/casual/piece work". Those types of subclassifications of employees are also specifically mentioned in ICSE-93. The Classification states that employees with stable contracts (regular workers) could be separately identified. Census planners will have to decide whether there is sufficient demand for such sub-groups, and which groups;

(b)   **Self-employed.** They are classified by whether in agricultural activities or not, and those in agriculture by whether the agriculture is mostly for cash or for home consumption. Those subclassifications of self-employed are of particular interest to national accountants; in some countries even further subdivisions have been made (see Bain, 1996). However, it is generally recommended that such detailed subcategorizations not be collected in censuses but through special surveys or case studies.

In the case of the self-employed, cross-tabulation of the status in employment variable with the appropriate agricultural categories of the industry variable (see chap. VII) can give an approximate distribution. Since a complex set of questions is normally needed to establish the other distinctions, those issues are better left for labour force or other household surveys.

319.   Aspects to be investigated include the following:

(a)   Recording for status in employment for those engaged in the broad range of agricultural economic activities now included in production, using either the suggestions above or other formats;

(b)   Examples of problem groups (mentioned above) in the local situation;

(c)   Local interpretation of or terminology for the categories "employees on a continuous basis", "own business" and "unpaid help in family business";

(d)   Any sub-groups of paid employees or self-employed, if it is decided to introduce them;

(e)   Producers' cooperatives, if they exist, and how are they currently recorded, and whether a separate category is needed for this group;

(f)   Any problems with persons temporarily absent from their main activities.

320.   A good critical review of the past and proposed use of statistics on status in employment groups by national accountants and labour market analysts is strongly recommended before the current round of censuses to guide the development and testing on the topic. Care should be taken to ensure that an economically active person is classified by status in employment on the basis of the same job(s) as used for classifying the person by occupation, industry and sector.

B.   Occupation

321.   "Occupation" is a topic that almost all countries have included in their censuses. In fact, it is the most common economic topic asked in censuses, except for basic activity status (employed/unemployed). Since in many countries the census will be the only source for statistics on the occupational profile of the entire employed population, the occupational distribution from the most recent population census is a major component of a country's labour market information (LMI). Changes in occupational distribution from one census to the next reflect the changes taking place in the economic and social structure of the country. Such statistics are of great value for many forms of labour-market description and analysis and are needed as a basis for policy formulation and implementation. The major target of occupational questions is the employed population, but for the unemployed population it is also common to ask about the occupation of the last job. It is recommended that the question to be asked of employed persons be separated from

that of the unemployed. The task of changing the wording will fall to the interviewers if only one question is applied to both groups. However, making such changes is usually confusing, strains the capability of many census interviewers and leads to poor results.

1.   Occupation: operational definition

322.   Occupation refers to the type of work done during the time reference period by the person employed (or the type of work done previously, if unemployed), irrespective of where (the industry) or under what economic status (the status in employment) the work is being done. An "occupation" is defined as a "set of jobs whose main tasks and duties are characterized by a high degree of similarity" (ILO, 1990a). A person may be associated with an occupation through the main job currently held, a second or other job or a job previously held.

2.   Occupation: requisites for applying the definition

323.   The International Standard Classification of Occupations (ISCO-88, updated in 2008 as ISCO-08)[2] and national occupational classifications, developed with the same principles, are structured so that jobs requiring similar skills are grouped together. That does not mean, however, that the response to be written on the census questionnaire should provide information about those skills. The information to be recorded on the questionnaire should reflect the type of tasks and duties undertaken in the work, as that is the information needed to determine the correct occupation code for the respondent's job. As explained in part five, which addresses the coding of occupation and industry, this information is usually most effective if it consists of a job title, supplemented with a few words to indicate the main tasks and duties of the job. With properly constructed coding tools, this type of information will allow detailed, reliable and effective coding.

3.   Occupation: review of questions used in national censuses

324.   Three types of approaches to the question(s) on occupation have typically been used in national censuses.

(*a*)   One or two pre-coded questions on the jobs that the individuals have (or had) (example E.1);

(*b*)   One question to obtain occupation-relevant information about an individual's job (examples E.2 and E.3);

(*c*)   Two or more questions, a basic question on the title of the position held with follow-up on main duties of the individual in the job (example E.4).

325.   Example E.1 shows the application of two pre-coded questions. Given that very broad categories are presented in the questions, considerable training would be necessary to ensure that interviewers are sufficiently aware of the contents of all the groups to guide respondents. It is likely that more detailed occupational categories could be derived by combining the answers to the two questions. The International Standard Classification of Occupations, 2008 (ISCO-08)[3] has 10 major groups, which are subdivided into 43 sub-major groups, with 131 minor groups and a total of 425 unit groups.

326.   It is also important to emphasize the point made earlier, that answers pre-coded for occupation will not produce the type of statistics needed by the majority of data users. However, there may be censuses where that solution is unavoidable owing to limited funding, in which case the pre-coded categories should be carefully thought out. Occupations for which information is definitely required should be given separate codes, and others may be grouped together. There will be very severe limitations on the number of

2 The new classification will be available for use in 2008 so that it can be adapted or used in censuses from 2010 onwards. The updated ISCO does not change the nature of the information required for coding by occupation. More information can be found on the ILO website: http://www.ilo.org/public/english/bureau/stat/isco/index.htm.

3 Available from http://www.ilo.org/public/english/bureau/stat/isco/index.htm.

Example E.1

E.1.1.   Indicate your professional position for your actual employment

- Manual worker, specialized worker _____ 1

- Qualified worker or highly qualified _____ 2

- Master agent supervising workers _____ 3

- Master agent supervising technicians _____ 4

- Technician, designer _____ 5

- Teachers, social workers, nurses and staff
  of B category in public services _____ 6

- Engineer, professionals (persons who do not have
  professional skills are not considered even if they
  contribute for social security under professionals) _____ 7

- Professors and all persons of A category of the
  public service _____ 8

- Employed in office, employed in commerce, agent
  of service, assistant nurses, nursery staff, persons
  of C or D category of public service _____ 9

- Other cases (give details) _____ 0

E.1.2   What is your main function?

- Production, manufacturing _____ 1

- Installation, maintenance, repair _____ 2

- Cleaning, caretaking, housing services (servant) _____ 3

- Lifting, storage, transport _____ 4

- Secretary, data entry, teller _____ 5

- Management, accounts, administration _____ 6

- Commerce, sales, technical business _____ 7

- Research, study, method, information technology _____ 8

- Managing directors or direct deputy _____ 9

- Other cases (give details) _____ 0

---

separate codes possible, and even with extensive training of interviewers only rough estimates of the numbers in the groups will result. The resulting response categories normally will not be represented by the broad groups presented in the occupational classification in use in the country. One should definitely not try to simply use the first digit groups of ISCO-08 or of the national standard classification. The ISCO-08 was not constructed for use in that way, nor was the national classification. It is also very unlikely that the meaning behind the titles of some of the groups in the classifications, for example, "technicians and associate professionals", will be recognizable to interviewers.

327.   In question E.2.1, sufficient space was provided to record the description given by the respondent. The interviewer will need to probe to get the level of detail required since the basic question does not indicate what that level is. If respondents are able to give their job names or titles, that data should be recorded as it may be of benefit when coding the answers. However, the use of the word "occupation" could present some difficulty for some categories of worker (for example labourers and some unskilled workers), who may not be accustomed to referring to the work they do as an occupation.

328.   Example E.3 is from a questionnaire that was presented in portrait format. The suggested question was asked, with explanations and examples given to show not only the type but also the desired level of specificity. Space was available for the answer. Although the use of the word "occupation" in the actual questions may not be well understood in all circumstances, the explanation and examples given should alleviate that problem. Very similar questions were also asked in a separate block for the unemployed, referring to their last job, which is good practice.

329.   In example E.4, the two questions asked are as suggested earlier. The questionnaire was in landscape format, which usually has restricted space for both questions and answers. This is, however, in the case of occupation and of industry, less of a problem

Example E.2

E.2.1   What occupation did you perform in your principal job during the week of 23 to 29 June 2000?   (The criteria for determining principal job were: the highest number of regular hours; the most time spent during the reference week; and the job that provides the most income).

Example E.3

E.3.1   Occupation   **If YES to X.X**   **What is the main occupation of (the person) in this workplace?**   Occupation refers to the type of work (the person) performed in the seven days before 10 October.   Use two or more words.   For example, street trader, cattle farmer, primary schoolteacher, domestic worker, fruit vendor, truck driver, warehouse manager, filing clerk, etc.

Example E.4

To be asked of all persons aged 10 years and over	
E.4.1 Occupation (*a*) What kind of work did ... do? (*b*) What are ...'s main duties at this job? (*a*) _____ (*b*) _____ _____	**E.4.2 Name and address of workplace** What is the name and address of ...'s workplace? _____ _____

for the questions than for the answers. The questions are standard, although in the present example the blank space in the question area could have been used for more explanation and examples. The very limited space for the two answers is, however, the major problem, since it severely limits the amount of information (or description of activities) that can be provided and is bound to affect the quality of data. The data are affected by interviewers' need to summarize the responses and then being precluded from probing for additional information.

330.   The use of questions worded "What kind of work did … do?" is quite common in labour force surveys. Some respondents, however, will tend to reply with words, such as "farm work" or "hospitality", that indicate tasks performed or the kind of industry worked in, rather than with an occupation title. Since the information required is both an occupation or job title and a statement of the main tasks or duties performed, it is sometimes necessary for interviewers to probe further or convert the information provided into an occupation title. Although it is possible to perform those tasks in a labour force survey with highly trained interviewers, it is more problematical in a census where it is less likely that interviewers could be trained when and how to probe. For similar reasons, it may be best to avoid such wording in self-enumerated questionnaires.

4.   Occupation: suggestions for enhancing the effectiveness of the questions

331.   Ideally the questions on occupation should seek full details, including (*a*) the title of the job and (*b*) a statement about the main tasks and duties performed. Considerable research has been carried out on how to best formulate the questions and code the responses. Most of the research has been done in English-speaking developed countries, but much is still likely to be relevant for developing countries and other languages. Some of the major findings are as follows:

(*a*)   The word "occupation" can be misleading, in some circumstances, and may best be either left out of actual questions on the topic or supplemented with a more easily understood word. For developing countries where translation in the field is very common, the terminology being used by interviewers in the local language should be carefully checked during testing and training periods. As always, the aim is to get the meaning of what is being asked across to the respondent, consistently, rather than insisting on the terminology in the language used in the questionnaire;

(*b*)   Space allowed for recording has been shown to be a major factor in poor descriptions of jobs and thus poor coding. In some countries good questions have been asked, but the space allowed for recording the response has been very limited;

(*c*)   It is most effective if two separate questions are asked. The first should elicit a description of the job in the form of the job title or name. The second should be a question on the tasks or duties usually performed. Seeking two separate pieces of information significantly increases the chance that the response will be sufficiently detailed or that some of the information will be useful for coding. A combined question has proved to be far less effective;

(*d*)   The quality of occupational descriptions is generally poor when they are given by an informant other than the job holder. Every effort should be made to interview all employed respondents in person for the questions on occupation, but that is clearly difficult to achieve, in particular in a large-scale exercise like the census.

332.   Example E.5 is a hypothetical set of questions that may be a useful starting point for countries that need to develop and test questions on occupation for use in the national census.

Example E.5

Hypothetical questions, recommended for testing and use in countries

E.5.1   **(In the main job held last week) what was (your) work or occupation?**
- *Please give full job title and be specific, for example:*
  - Fruit picker
  - Legal secretary
  - Restaurant manager
  - Secondary schoolteacher
  - Cattle farmer
  - Registered nurse

Occupation:

_____

_____

_____

E.5.2   **What are your main tasks or duties in that job?**
- *Please give details. For example:*
  - Picking and carrying oranges and peaches
  - Preparing legal documents
  - Managing the operations of a restaurant
  - Teaching mathematics
  - Managing a cattle farm
  - Caring for the sick and administering medications

Main tasks or duties:

_____

_____

_____

5.   Occupation: issues of measurement

333.   Using pre-coded response alternatives with the occupational question(s) has the major advantage that it is by far the least costly solution, as one does not have to code the answers after interviewing. However, there are major disadvantages to pre-coding. Occupation is a complex topic and, even with a very well thought out set of response categories and with excellent training, interviewers or the respondents themselves may have great difficulty placing a job in the correct category. Both the validity (correct categorization) and the reliability (same categorization made by different interviewers of equivalent responses) of pre-coded occupational categories have been shown to be very poor. In addition, only a limited number of categories are possible with pre-coding and thus the details needed by many users of such statistics cannot be provided. Pre-coding is therefore not a reasonable alternative for this major variable in a census. The expense of a coding operation may therefore be justified in the light of the quality and relevance of the data that are generated.

334.   **Agricultural activities.** These activities are often relatively easy to describe, for example, with terms such as "vegetable farmer", "maize or rice farmer", "goat raising" or "cattle rancher". A brief description of the tasks performed can be given with two or three words as a response to the second question or second part of the question. A problem arises when the classification makes a distinction between those for whom the same activities are mainly for sale (referred to in ISCO-08 as "market-oriented skilled

agricultural workers") or mainly for home consumption (referred to in ISCO-08 as "subsistence farmers, fishers, hunters and gatherers"). Most classifications based on ISCO make that distinction, if subsistence farming is of importance in the country, which may require an additional question, since the difference cannot be expected to be reflected spontaneously in the reply given by the respondents. A possible supplementary question may then be, "Do you produce mainly for sale or is it mainly for your own or family use?". As a general rule such questions should actually be written into the questionnaire and not just given as notes to interviewers. In developing countries with a large number of farmers, it has generally been found that interviewers have been able to manage this addition to the occupation question with reasonable accuracy and that they usually have a good understanding of subsistence farming. More precise estimates or complex subdivisions such as "all cash crops/part cash/all subsistence" are best done by separate questions for the agricultural activities group or, preferably, in specialized surveys or agricultural censuses rather than in population censuses.

335.   **Other non-market production.** As noted in previous chapters there are many activities that are regarded as economic production even if they are not marketed or exchanged. Examples include fetching water or firewood; the processing of farm or natural produce (for example, making butter, brewing beer, extracting oil, making hats, mats or clothing, whether for sale or only for the household's own consumption); own house repair or construction; farm construction work such as digging irrigation ditches or building fences; and volunteer building of communal roads. When responding to questions such as "What kind of work does (…) do?" or "What are the main tasks and duties?", words reflecting those activities should be recorded as for any other job, with a job title and some description of the tasks performed. However, it will help overall coding if interviewers are told to record that such activities are mainly for own use, similar to the distinction between agricultural activities that are mainly for sale and those that are mainly for own use. If some of the activities are very common and there is no other way of identifying them separately, it may be necessary to create special occupational codes similar to that for subsistence farmer. It should be noted, however, that on farms in particular, those activities are relatively rare as a main activity separate from the general farm work. They are often carried out as secondary activities of farming families, although they can be a main activity in the agricultural off-season and in non-farm settings. In countries where they are likely to occur reasonably often, it is very important that they be allowed for, and interviewers should be given guidance about how to record such activities for the coding of occupation and industry.

336.   Interviewer training is the key to good occupational descriptions for subsequent coding. Interviewers should be briefed on the need for accurate and relevant information about the job. They should be thoroughly drilled on adequate and inadequate descriptions through a series of examples. The examples should be local and come from the queries and problem cases in recent censuses and surveys. Descriptions such as "manager—managing"; "public servant—office work"; "supervisor—supervising"; "clerk—clerical work"; "apprentice—helping"; "labourer—manual work"; and similar uninformative responses must be stressed as inadequate or they will occur frequently. Field practice should focus on the responses to the occupation questions. Supervisors should be instructed to send interviewers back to households when inadequate responses are given. Owing to time constraints in the census, however, it may not always be possible to do so. It is only with such a concentrated effort that the quality of occupational responses will improve and, with a similar effort on coding, the quality of the final statistics.

337.   Even if an occupation question has been asked several times before, it is worth testing it for possible improvements. Better wording, more examples and more ample

space should all be tested. As always, the test(s) should have clearly stated objectives and be quantitatively evaluated. Any questions with pre-coded answers must be thoroughly tested to minimize errors, to ascertain how interviewers are going to interpret the summary groups given, in realistic conditions, and to correct any problems.

C.   Place of work

338.   In the past, a question on the geographical location of the place of work has been included in censuses in some European and North American countries, but will be a new topic in most other countries if included in the 2010 round of censuses. Where that topic has been included, the actual address of the place of work has often been recorded when relevant and used to code the location to a fine geographical level. That has allowed detailed tabulations and mapping of place of residence by geographical location of place of work. Data on actual address of the place of work can also be useful for industry coding (see paras. 680-686) in countries where an establishment register has been developed that shows the industry code of each recorded establishment.

339.   In some countries, there may be concerns about the sensitivity of questions on the address of the place of work owing to fears that there may be follow-up to a respondent's employer. In many developing countries, it may not be possible to gather data on actual address of place of work because street addresses do not exist, and for proxy responses, the address may not be known. In those situations, it would be useful to consider collecting information on the village, suburb or similar low level of civil division.

340.   Additional questions have often been asked on the method of travel to work, and those and other variables have enabled the production of statistics on important travel-to-work patterns. That information has then been very valuable as a basis for transportation planning.

341.   While the information on place of work can be used to develop area profiles with respect to the employed labour force (as opposed to demographic profiles by place of residence), the primary objective is to link the place of work information to the place of residence. Therefore, the geographical location of the place of work should relate to the smallest civil division in which the economic activity is performed in order to establish commuter flows from the place of residence to the place of work.

342.   The type of place of work has also been included as a topic in many censuses and provides interesting information in its own right as well as providing supplementary information in the analysis of informal employment.

1.   Place of work: operational definition

343.   "Place of work" is the location in which a currently employed person performed his or her job.[4]

2.   Place of work: requisites for applying the definition

344.   As presented in *Principles and Recommendations, Revision 2*, information on both the geographical place of work (mentioned above) and the type of place of work is collected. For "type of place of work", three categories are sought:[5]

(*a*)   **Working at or from home.** This category includes persons whose work is directly at their home, for example, farmers who work and live on their farms; self-employed persons operating (work) shops or offices inside their own homes; persons working and living at work camps or military compounds; persons doing home-based work for a wage, such as sewing, tailoring or professional activities, for example,

[4] Or where a "usually employed" person currently or last performed the job.

[5] The categories and their definitions have not been formally adopted by the International Conference of Labour Statisticians.

accountancy or computer programming; persons producing non-agricultural items for own use at home; persons participating in informal sector activities, such as production for sale at home or sales at or adjacent to the home (small shops etc.); and those doing agricultural work on the farm near their home. Paid employees in agriculture living on the farm and domestic employees living with or near their employer would also be taken as working from home, even if they travel to another residence, for example, every weekend;

(*b*) **No fixed place of work.** This category is restricted to persons who travel from place to place for their work and do not report daily in person to a fixed address as a work base: it includes hawkers, travelling salesmen, long-distance truck drivers, seafarers, fishermen and some taxi drivers. It includes ambulant vendors, operators of street or market stalls that are removed at the end of the workday, construction workers working at different sites during the reference period, push-cart operators and so forth. Construction workers and the like based at their current work site for longer than a week should be classified to group (*c*);

(*c*) **With a fixed place of work outside the home.** All other employed persons should be included in this category, including all persons who move around in their job but have a fixed base location to which they report daily, such as bus and taxi drivers (with a base), train and airline staff, and operators of street and market stalls that are not removed at the end of each workday. This group may include individuals who travel to work, on a regular basis, across the border to a neighbouring country. Persons whose work involves changing sites, for example in construction, should give the location of their current worksite rather than the address of their employer's place of business, if appearance at the site will be required for at least one week.

345. It is for those classified under the last group that a supplementary question on the precise address has been added in many countries. Note that the above-mentioned categories provide only a geographical reference for the place of work. They do not make distinctions between the physical characteristics of the place of work, such as whether it is at a street corner, with a roof, in a shack or in a building, which may be relevant when determining whether the place of work is in the informal sector or not.[6]

346. There are no major conceptual problems with the preceding categories. It is, however, necessary to define how close to home "work at home" is, especially for informal sector activities. Those taking place on the same plot of land as the dwelling should definitely be taken as "at home". However, the definition is to some extent a function of the uses of the information and national circumstances; and in some cases "at home" has been defined to be at some distance and/or time from the home. It is preferable, however, to limit the definition to the structure and area around the home. The classification of those working on farm fields that are at a distance from the home requires careful consideration. Pre-testing should help to establish all such problem categories, and their treatment will have to be included in training and manuals for field staff.

3. Place of work: review of questions used in national censuses

347. For those countries that have included a question on place of work, the form of the questions suggests the following objectives:

(*a*) To determine if the activity is undertaken in the home, its vicinity or some other location, as in the international recommendations (example F.1);

(*b*) To obtain specific information about the actual location of the workplace (example F.2);

(*c*) A combination of both objectives (examples F.3 to F.6).

[6] For additional information on the topic, see ILO (2002).

348.   In question F.1 the emphasis seems to be on working from the home, either own home or in another's home. The second question relates to the administrative area in which the workplace is located. There is, however, no category for "no fixed location". The category "more than one parish" provides some indication of those who may work from no fixed location, but a specific category would be needed to obtain a more accurate count of those working from no fixed location (such as street vendors).

349.   Example F.2 presents a two-part question in which the first part seeks information about the place of work relative to the place of residence. The use of the term "sub-place" here is not specific and can be confusing, as it is not clear if it means the dwelling unit or compound, the locality or another geographical demarcation. However, some clarification of the term is given indirectly in the second part of the question, which is addressed to those who answer "no" to the first part of the question. The definition given for sub-place suggests that the intent of the question did not include determining the number of people who work at home or have no fixed work location. It should not be too difficult to extend the first part of question F.2.1 to cover those categories. The reason for limiting the information about the place of work in another country to the name of the country is unclear. In areas where cross-border commuting for work is common, it would be as useful to have the precise commuting patterns across the border as it is to have them for commuting within the country.

Example F.1

F.1.1	**Where is your/is …'s place of work located? (Read categories)**	
	○ In own home or yard	○ On a farm
	○ In another home or yard	○ Not stated
	○ Not in a private home	
F.1.2	**In which parish do you/does … work?**	
	○ Kingston	○ Hanover
	○ St. Andrew	○ Westmoreland
	○ St. Thomas	○ St. Elizabeth
	○ Portland	○ Manchester
	○ St. Mary	○ Clarendon
	○ St. Ann	○ St. Catherine
	○ Trelawny	○ More than one parish
	○ St. James	○ Not stated

Example F.2

F.2.1a & F.2.1b	**Place of work**
	If YES to X.X
	Does (the person) work in the same sub-place in which s/he usually lives?
	Y = Yes
	N = No
	Dot the appropriate box.
	If NO, where is this place of work?
	If NOT the same place, write PROVINCE *PR*, MAIN PLACE (city, town, tribal area, administrative area) and SUB-PLACE (suburb, ward, village, farm, informal settlement).
	If another country, write the name of the country in the boxes below.

350.   A format similar to the one being proposed was used in example F.3, where the first question relates to whether the individual worked at home or elsewhere, in which case the full address of the place of work was obtained as a second part to the question. The question used is more specific about the relative location and also clear about how live-in employees of the household are to be classified.

351.   Example F.4 has as the main question the "full address" of the place of work, with the two categories to be checked if the respondent worked at home or at no fixed location. The questions are simple enough to be used in almost all countries, but in the event that street addresses are not commonly used, the locality and related geographical information can be used, as in example F.2.

352.   Example F.5 is from a self-enumeration schedule. It is an example of the more detailed questions used in some developed countries. Note, however, that the three recommended categories (see para. 344) are there. In other countries the three basic categories could simply be asked without further elaboration, but in developing countries it will be essential to give examples or notes on the questionnaire for the common situations in that country. Presumably, "including farms" means residential farms. If broader coverage is intended, it would be preferable to include a separate category for "farms" since the "work at home" category is of increasing interest and value to analysts and labour legislators, especially in the context of home-based work and informal employment arrangements. Note that the question treats workplaces abroad differently from those within the country.

353.   In question F.6.2, special attention is given to "farms" as a place of work. The design of the question would imply that farms that are not the person's home should

Example F.3

F.3.1 Place of work or location of school	Outside home
• For a person who worked and also attended school during the week, give the place of work.	Home (including living-in employee) — Same *ku, shi, machi/cho, mura/son* — Other *ku, shi, machi/cho, mura/son*
• If the place is other than one's own home, write the address of the place of work or the location of the school (i.e., the name of the *to, do, fu* or *ken* (prefecture), and the *shi* (city), *ku* (ward), *machi/cho* (town), or *mura/son* (village)). Do not omit the name of the *ku*, if the address is in one of the 13 major cities.	(Write the address)
	Go to question X.X
The major cities are ...	*to, do, fu, ken*
	shi, gun
	ku, machi/cho, mura/son

Example F.4

F.4.1   What is (was) the full address at which you actually work(ed)?

1 ☐ Work mainly at or from home   2 ☐ No fixed place of work

Example F.5

F.5.1 **At what address did this person *usually* work?**	07 ○ Worked at home (including farms)
Example:	**Go to question** `Y.Y`
365 Laurier Ave. West	08 ○ Worked outside `the country`
Number ↑	**Go to question** `Y.Y`
Name ↑	09 ○ No fixed workplace address
Type ↑	**Go to question** `X.X`
Direction ↑	10 ○ Worked at the address specified below:
If direction (e.g., North, South, East or West) is a part of the street address, please include it.	**Specify complete address**
If street address is unknown, specify the building or nearest street intersection.	Street address (see example)
Note:	11 →
Some large cities are made up of smaller cities or towns called municipalities. Where applicable, identify the municipality rather than the larger city, for example:	City, town, village, township, municipality or Indian reserve
	12 →
	Province/territory
	Postal code
	13 → ☐☐☐ ☐☐☐

be classified in the first response category and that the third response category is only for farms that are the person's home. However, that distinction should be clearer. A possible alternative might have been: "home—not a farm,"; "home—located in farm"; "farm located outside the home"; "other fixed premises outside the home"; and "not a fixed place of work".

354.   More guidance is needed in both F.6.1 and F.6.2 on the treatment of taxi drivers, tradesmen and others who operate in a mobile way from a fixed business location. As indicated earlier, they should be classified as "with a fixed place of work".

Example F.6

F.6.1 **Branch of industry of place of work**	
Name and place of work	
_____	
_____	
District ☐☐ Commune/Municipality ☐☐	
Describe in detail the economic activity at your place of work	
_____	
_____ NACE ☐☐	
F.6.2 **Type of your place of work**	
Fixed premises outside home ☐ 1	
Home (not a farm) ☐ 2	
Farm ☐ 3	
Not a fixed place of work ☐ 4	

Note: NACE = Statistical Classification of Economic Activities in the European Community.

4.   Place of work: suggestions for enhancing the effectiveness of the questions

355.   The information in example F.7 may be presented on a prompt card or included in a questionnaire. For a landscape questionnaire layout, the notes or examples would have to be considerably reduced owing to the limited space, and it would be necessary to rely on the interviewer or supervisor manual and on training.

5.   Place of work: issues of measurement

356.   Place of work is a relatively simple topic, but as a new topic for most countries it will need thorough testing that focuses on problem groups (especially the group working from home) and how to assign them to the proposed categories. The response categories should be determined according to national needs.

Example F.7

Hypothetical question, recommended for testing and use in countries

F.7.1   In this job, was the person:

☐ Working at home

Include   • *Farmers and their employees living on that farm*

• *Domestic paid workers living an their employers' plot*

• *Home-based wage jobs*

• *Home production for own use or sale*

• *Fetching water and firewood*

• *Other informal activities based at or near the home*

☐ Working at or from an office, shop, factory or other fixed location not at home

Include   • *All activities with a fixed location or base, e.g. bus drivers, construction workers*

☐ Working (operating) without any fixed location

Include   • *Only activities not carried out at a fixed base*

• *Hawkers, travelling sales representatives or long-distance drivers, seafarers*

Chapter VII

Descriptive characteristics of the establishment

357. It is essential that once the main or reference job is identified, all characteristics should refer to that same main job throughout the questions that follow. The questionnaire design must be such that it will not confuse respondents or interviewers on that point. When any secondary activities are supposed to be recorded they must be clearly identified separately and placed so as not to confuse the flow of questions relating to the main job. The task is not usually difficult with current activity (see examples below) but can be quite difficult with usual activity.

A. Industry

358. As a census topic, industry is only slightly less common than occupation and has been included in several census rounds in many countries.

359. Industry is a key classifying item for most employer- and establishment-based employment and wage surveys. However, most of those surveys cover only larger, formal employers, and therefore a population census is usually the only possibility for obtaining comprehensive and detailed statistics of all employment in a country by industry. The census statistics on industry are thus a major basis for the economic statistics of most countries as well as for their labour market and social statistics. To facilitate comparability, it is essential that the industry classification used in the census be consistent with that used in other relevant surveys.[1] In previous censuses many countries have used the International Standard Industrial Classification of All Economic Activities, Revision 2, approved in 1968, and ISIC, Revision 3, approved in 1989. It is hoped that all countries will have changed to a classification compatible with ISIC, Revision 4, by the time of the 2010 round of censuses.

1. Industry: operational definition

360. Industry refers to the activity of the establishment in which an employed person worked during the time reference period established for data on economic characteristics (or last worked, if unemployed). It describes what the establishment does, not what the individual does when working for that establishment; for example, a person may work as a security guard in a department store, an accountant at a hotel or bus driver who drives passengers to the aircraft at an airport. For those who are self-employed without any employees on a continuous basis, the activities of the establishments that they represent will be those that they undertake themselves. For multi-establishment enterprises, the predominant activity of the establishment (the location where the person worked) should be recorded, rather than the predominant activity of the enterprise (legal entity).

2. Industry: requisites for applying the definition

361. There are generally two components to the desired industry question(s). The first part of the question should ask for the name of the place of work, with details of division

[1] The use of consistent classifications will not alone achieve comparability. Other factors that may affect comparability are outlined in paras. 30 and 110-111.

or branch for larger private enterprises and government institutions and establishments. That information may be recorded under the place of work variable (see chap. VI, paras. 338-342). Asking for the full name and address may also be useful, in particular when access to an establishment register is possible. The second part of the question must elicit a clear description of the main products or the main functions of the establishment.

3.   Industry: review of questions used in national censuses

362.   The questions used in censuses are grouped under the following three categories:

(*a*)   Questions with pre-coded categories (example G.1), as in the case of occupation;

(*b*)   One question that asks about the type of activity undertaken by the establishment (examples G.2 to G.4);

(*c*)   Two or more questions, a basic question on type of activities and one or more other questions requesting supplementary information on other characteristics, such as place of work or address of the workplace (examples G.5 and G.6).

363.   As with occupation, the use of pre-coded response categories for industry (see example G.1) is not desirable, as it is very difficult for field staff to allocate employed persons to the correct category. In addition, that strategy will limit the detail available for this variable. Although it is easier to draw up a reasonably sound pre-coded list for industry (based on the International Standard Industrial Classification of All Economic Activities) than for occupation (based on the International Standard Classification of Occupations), the list of categories will still be severely restricted and the resulting groupings are likely to be of very limited usefulness.

364.   A useful strategy employed in the example for Australia (see annex I, sect. D) is to combine the respondent's pre-coded response (question 42 in the Australian questionnaire) with more detailed information (in question 43) as inputs to the automated coding system,

Example G.1

G.1.1	If coded 1 (employed)	Type of main industry			
		Food crops	1	Plantation	2
		Fishery	3	Animal husbandry	4
		Other agriculture	5	Manufacturing	6
		Trade	7	Services	8
		Transportation	9	Others	0

Example G.2

Branche d'activité	Branch of activity
G.2.1 Marquer le genre d'activité de l'établissement où travaille le recensé ou le travail que fait le recensé indépendant.	**G.2.1 Mark the type of activity of the establishment where the respondent works or the kind of work if s/he works for her/himself.**
Cette colonne est "Sans objet" pour les personnes de moins de 10 ans.	This column is not applicable to persons younger than 10 years.
Pour ces personnes, inscrire 888 dans les grilles.	For these persons, write 888 in the cells.

together with details on name and address from earlier questions. The pre-coded response is important and useful supplementary data in the process (for more details on automated coding systems, see chaps. X and XI).

365.   In example G.2, the space allocated for recording activities in the column for the branch of activity is not adequate for providing any details. The question provides space for entering a three-digit code, which implies that the enumerator is to match the description given by the respondent with the codes for the activities, select the appropriate code and enter it on the questionnaire. That approach not only prolongs the interview but is also prone to errors, and there are no means of checking or correcting those errors after the interview is completed, since the actual response to the question would not have been recorded. The problems associated with interviewer coding, covered in part five, are also inherent in the approach illustrated in example G.2.

366.   Example G.3 presents one question about the type of activity (what the establishment is dedicated to or produces). Sufficient space is given to record the response.

367.   A stand-alone question requesting a description of the activity is unlikely to elicit the details needed to accurately classify industry to the third or fourth digit level codes presented in ISIC. However, it is recognized that census planners may be forced, owing to space limitations, to use only one question for industry, particularly with the landscape questionnaire layout. In such cases it will still be better to ask a combined question on the name of the location and the activity, rather than solely asking about the activity.

368.   Example G.4 is also a fairly comprehensive industry question, but it does not attempt a combination with place of work. It has very useful notes regarding the self-employed and paid domestic work. The manual and training for interviewers should provide guidance on how to record persons working from home and persons with no fixed place of work. For those cases, it would have been best to record their statements for the answers to G.4.1. The question includes the word "industry". Its use would be unacceptable for a single question, but in the example alternates are given in case the term was not understood, which is good practice. Question G.4.1 does refer to the person's employer or company, whereas it is the actual establishment or place of work that is recommended as the reference for industry. It is uncertain how large multi-establishment companies or similar government departments would be recorded.

Example G.3

En el trabajo principal, en el que trabaja más horas...

G.3.1   El lugar o establecimiento donde trabaja, ¿a qué actividad se dedica o qué servicio presta?

In the main job in which you worked the most hours...

G.3.1   What activity is the place or establishment where you work mainly engaged in or what services does it provide?

369.   The three-part question in example G.5 is captioned by the main question, "Where do you work?". The description of the activities the company of the employer engages in is presented with the work address and name of the company. The information requested should be easy enough for the respondent to provide, and there are no technical terms that would require explanations.

370.   The questions in the example G.6 cover both industry and place of work topics. They show how the problematic groups mentioned earlier—persons with "no fixed place of work" and persons "working from home"—can be identified, thus averting possible confusion when coding industry. The notes on the questionnaire are extensive and do take considerable space.

4.   Industry: suggestions for enhancing the effectiveness of the questions

371.   Industry questions have been asked in several censuses in many countries and thus there is usually some experience with this topic. However, many problems are still being identified with the resulting statistics. The following issues have been found to be of concern:

(*a*)   It is important to record as much information on industry as possible during the census so that accurate industry attribution can be derived. Lack of space on the questionnaire for recording the relevant information has been found to be the major factor resulting in inadequate industry descriptions. It is particularly the case if landscape format is used for the questionnaire;

(*b*)   The word "industry" should not normally be used in the question, as respondents' understanding of the word frequently does not correspond to what is intended. However, the word may be given as a heading for the question, as has been done in some cases, to guide interviewers. One of the objectives in the training of fieldworkers should be to eliminate any confusion that interviewers may have about the meaning of the term;

Example G.4

Company/business activity (P-19b)
If YES to P-18

G.4.1 What does the business do (main economic activity)?

Write the MAIN INDUSTRY, economic activity, product or service of (the person's) employer or company. For example, gold mining, road construction, supermarket, police service, healthcare, hairdressing, banking.

OR

Write the activity of the person if self-employed. For example, subsistence farming.

If doing PAID domestic work in a private household, write DOMESTIC SERVICE.

Use CAPITAL LETTERS only.

Example G.5

G.5.1.a geo-code	**G.5.1 Where do you work?**
	Work address: _____
	Name of company/organization: _____
	_____
G.5.1.b isic-code	Type of activity engaged in by company/employer: _____
	_____
	_____

Example G.6

G.6.1 What is the name of your employer?

Please do not use abbreviations.

Give the trading name if one is used.

Civil Servants—give names of departments.

Local government officers—give name of employing authority.

_____

_____

_____

G.6.2 What is the main thing which the organization you work for makes or does, at the place where you work?

Describe clearly what the employer (or yourself, if self-employed) makes or does.

Civil servants—write "government department".

Local government officers—give departments in which employed.

_____

_____

_____

G.6.3 What is the address of the place where you work?

If you are not reporting a fixed place, tick "no fixed place".

If you work mainly at home, tick "mainly at home".

If you are employed on a site for a long period, give the address of the site.

If you do not work regularly at one place but report daily to a depot or other fixed address, give that address.

☐ No fixed place

☐ Mainly at home

Please remember to give the postcode.

_____

_____

_____

(*c*)   The temptation to give too many examples should be avoided as it will clutter the questionnaire and may confuse rather than enlighten the interviewer. Examples should be selected on the basis of previous experience, noting in particular those whose identification has often caused problems for enumerators. Similarly, instructions to enumerators have to be restricted to the most critical ones only.

372.   It is useful to give enumerators good examples of goods or services produced and it may be necessary to give them effective training for such a complex question to succeed.

5.   Industry: issues of measurement

373.   Owing to its importance, the industry question should always be included in any pre-testing programme, even if the topic has been asked many times before. Most developing countries and many countries in transition are still searching for a means of improving the canvassing of this important topic. Aspects that need to be tested include the following:

Example G.7

Hypothetical questions, recommended for testing and use in countries

G.7.1 What is the name and address of your employer or business?

(*a*) Name _____

(*b*) Address _____

_____

G.7.2 What are the main goods or services produced by your employer or business?

_____

_____

- The most suitable order for the block of questions on industry;
- The means of obtaining good industry descriptions;
- The most effective examples and notes to be included in the questionnaire;
- The identification of establishments with different industry codes when they belong to the same large enterprise or government department.

374.   Good interviewer training is the key to the collection of accurate industry descriptions. Interviewers should be briefed on the need for accurate and relevant information about the activities of the place of work, using several words to give detailed descriptions. They should also be thoroughly drilled on the difference between adequate and inadequate descriptions through a series of examples. The examples should be local and come from the queries and problem cases in recent censuses and surveys. Unsuitable descriptions that occur often, such as "manufacturing", "public service", "teaching", "farming" or "selling things" must be stressed as inadequate. Field practice should also give sufficient emphasis to the descriptions given for the industry and occupation questions. Supervisors should be instructed to send interviewers back to households when inadequate descriptions are given. It is only with such a concerted effort that the quality of industry and occupation descriptions will be adequate and, with a similar effort on coding, that the high quality of the resulting statistics can be assured. Any lists of major enterprises (private and government), if used, should also be handed out when training interviewers and should be explained so that the interviewers are also fully aware of why it is necessary to record the location description for the different units.

375.   Special attention should be given to explaining difficult-to-classify categories such as production for own consumption, paid household staff and work at home or with no fixed location. These categories are explained below:

(*a*)   **Production for home consumption.** Production for home consumption can be significant in some countries. Subsistence agriculture is usually the most common relevant activity but other activities are also possible, particularly in the agricultural off-season. Those activities were covered in more detail (see paras. 334-335) in relation to the occupational classification of such activities. The same type of problem arises for industry. It is very important to have a strategy for identifying the activities in this category separately when they are significant. The International Standard Industrial Classification of All Economic Activities, Revision 3, did not have industry codes for subsistence agriculture nor did it have corresponding codes for manufacturing for home consumption or own-account construction. However, a strategy was developed, and the codes were presented in the *International Standard Industrial Classification of All Economic Activities (ISIC), Revision 3.1* (United Nations, 2002). As a result, that revision has a separate code (9600) for "undifferentiated goods-producing activities of private households for own use", and a similar code (9800) appears in ISIC, Revision 4;

(*b*)   **Paid staff for private households.** Workers in this group include mostly domestic servants, gardeners or watchmen working directly for private households. Their situation needs to be identified on the questionnaire. They can then be coded to a separate industry code, for example, 9700 for ISIC, Revision 4. Generally, either the first part of an industry question should serve to identify them as working from home, when they live on the premises, or interviewers should be instructed to record the private household as their place of work. That will differentiate them from persons working for security companies, for gardening and landscaping companies, or for domestic service supply companies, who should be classified to the industry of their employing company, rather than to 9700 under ISIC, Revision 4, even though they may only provide services to one household on behalf of the employing company. The procedure must be explained well to interviewers, as the group is usually of a significant number in developing countries;

(*c*) **Persons working at home or with no fixed location.** It is often confusing to ask for details about the place of work for persons working from home or with no fixed location.[2] If an address is recorded for such cases, it may only be that day's location for a hawker, or some description of the home location. There are two alternatives for these cases: (i) to continue to ask the question for all the employed but tell interviewers to write in "home" or "no fixed location", or (ii) to try to design the questionnaire so that the place of work is not required to be recorded.

B. Institutional sector

376. The institutional sector of employment is broken down into five main groups or sectors (see para. 377). The nature of and conceptual differences between the sectors makes it difficult for a population census to display the distinctions between them precisely. In the 2000 round of population censuses, although half the countries attempted to collect certain information on sector of employment, most of the attempts represented partial collection of the categories, notably government sector employment and private sector employment, which may appear within the status in employment response categories. Nevertheless, those replies have provided very valuable statistics. Furthermore, countries may derive the institutional sector from the name of the enterprise or establishment reported by the respondent, provided the name can be linked to a business frame that allows use of the institutional sector classification.

1. Institutional sector: operational definition

377. The term "institutional sector of employment" relates to the legal and social organization and institutional status of the establishment in which a job is located. Following the definitions provided in the System of National Accounts,[3] a distinction should be made between the following types of institutional sector categories, which are given in *Principles and Recommendations, Revision 2*. They are as follows:

   (*a*) The non-financial corporations sector;

   (*b*) The financial corporations sector;

   (*c*) The general government sector;

   (*d*) The non-profit institutions serving households sector;

   (*e*) The household sector.

378. Categories (*a*) and (*b*) may need to be further subdivided according to whether they are controlled by government or private entities[4] in order to achieve estimates of public sector employment that are separate from private sector employment, as required by many users.

2. Institutional sector: requisites for applying the definition

379. Owing to the complexity of the definitions and the difficulty in applying them in a large-scale exercise, the main relevant distinction to be made in a population census is to split paid employees into government-paid employees and other paid employees. That split will be the primary focus of the present section. The split is very important for labour policy analysts, as almost all employment estimates and policy plans attempt to provide separate statistics for the government and the other (mostly private) sectors. That is a result of the perceived difference in relevant policies for the two sectors.

380. The definition of government comprises:

---

2 It is necessary to take care that those who don't work in a fixed location but do report to an establishment somewhere are not treated as "no fixed location" (see also para. 344).

3 See Commission of the European Communities and others, *System of National Accounts, 2008* (United Nations publication, forthcoming), chap. IV; and *Principles and Recommendations for Population and Housing Censuses, Revision 2* (United Nations, 2008b), para. 2.335.

4 Some countries may use local definitions and practices to clarify whether an enterprise or institution is government- or private-controlled. Some countries record whether the unit is government- or private-controlled on their business register and are then able to use this information in census processing. Chapter 22 of *System of National Accounts, 2008*, paras. 22.13-22.25, provides guidance on this issue.

(*a*)   All branches of government: national, state, provincial or local;

(*b*)   Social security funds imposed and controlled by those units;

(*c*)   Non-profit institutions engaged in non-market production that are controlled and mostly financed by the units in the above two categories.

381.   The employees of all branches of government are generally not difficult to identify in a population census. Employees of social security funds can be identified, but the issue is whether extra steps need to be taken to ensure they are included in the government sector rather than with other financial or non-financial sectors, in which case a separate question, category or instruction will be required.

382.   Employees of non-profit institutions are also quite difficult to identify separately in a census. Non-profit institutions include schools, hospitals and trade associations, some of which are government-controlled but others which are partly or fully privately financed and controlled. When mixed types of non-profit institutions exist in a country, it is essential that clear guidance be given to field staff in regard to whether they are to be included or excluded from the government sector. The issue should be discussed with national accountants and labour policy analysts. Local terminology rather than national accounts terminology should be used in manuals or notes to interviewers (for example, "include all employees of government-aided schools, colleges and hospitals as 'government'").

383.   The situation for all semi-government or parastatal (partly government and partly private)[5] organizations needs to be made clear. In many developing countries there are many such organizations, though they are now being partly or fully privatized. For simplicity it would be best to exclude all such organizations from government and provide a separate category for them among the possible responses to the appropriate census question. A wide range of examples of such institutions should be included in the field manuals and covered during the training of field staff. The status and ownership of the post office, national railways, harbour boards, national airline, and the like, vary from country to country and over time in any country. It should be made very clear whether such organizations and institutions are to be classified in the government or in the semi-government category (as the case might be). Often interviewers and respondents think of some of them as government organizations even when they are partly or fully privatized.

384.   **Subdivisions of government**. Some countries will want to identify separately different categories of government, for example, national Government, state or provincial government and local government. Steps should be taken to ensure that the distinctions between the categories are clear and well understood by field staff and respondents. In some countries there are complex situations, for example, where the central Government pays and controls the teaching staff of schools but local governments pay and control other staff and the general management of the schools. Such situations always seem to lead to confusion even if explained well to field staff. If there are too many such complex situations, it is recommended that those subdivisions of government not be attempted, or that the procedures should be pre-tested.

385.   The industry codes for financial services will allow the further splitting of those employees in the financial corporations sector. Employment in non-profit institutions is very difficult to measure separately in a population census as the required information on the characteristics of the establishment (their non-profit status and who they service) generally cannot be expected to be known by many of the persons employed in them. Thus it is expected that paid employees of establishments in this sector will mostly be grouped together with those employed in private non-financial corporations, although some may also consider themselves to be government employees.

386.   **Subdivisions of other**. Some countries have subdivided the "other paid employees" into categories for missions, churches, non-governmental organizations or parastatals,

5 The term "parastatal" may also need a locally known definition, as in some countries the term means commercial enterprises in which the Government has a controlling interest or ownership.

when such organizations are significant employers in the country concerned. With any such set of subcategories it is recommended that the relevant questions and response alternatives be well tested to ensure they are easily and accurately understood. For example, the non-governmental organizations category has produced some surprising results when tested. In addition, some non-profit institutions are serving in the market sector (see United Nations, 2003).

3.   Institutional sector: review of questions used in national censuses

387.   Two types of questions have typically been asked:

(*a*)   Those that treat the institutional sector as a distinct topic investigated on its own (example H.1);

(*b*)   Those that combine the institutional sector with status in employment (examples H.2 to H.4).

388.   In example H.1, the question on the institutional sector (H.1.2) is very simple with two main categories, State/government and private, and is addressed to employees only. For many purposes this will suffice. However, as the above explanation indicates, there are subcategories under government that users might be interested in.

389.   The question in example H.2 combines status in employment and number of employees in the establishment with institutional sector. The combination unduly complicates the question and could increase the time that it would take to complete the interview. The quality of responses should also be of concern; employees do not always know how many workers there are in the establishment. It would simplify the task of enumerators and help to improve the accuracy of responses if separate questions were presented for the different characteristics, as in example H.1.

390.   In example H.3, the question provides more of the institutional sector categories than usual. It assumes that respondents (who may be reporting on behalf of other household members) know if their workplace is a "private, for-profit" company or a "private, not-for-profit" organization. It is strongly recommended that the understanding of any such categories be thoroughly tested before being introduced. "Government employees"

Example H.1

H.1.1 En el trabajo principal, ¿es...	H.1.2 ¿Trabaja en...	H.1.1 In the main job, are you ...	H.1.2 Do you work in ...
obrero o empleado?	un empleo estatal (nacional, provincial o municipal)?	Worker or employee?	State employment (national, provincial or municipal)?
	un empleo del sector privado?		Private sector employment?
	Ignorado		Don't know
patrón?		Employer?	
trabajador por cuenta propia?		Own-account worker?	
trabajador familiar?		Family worker?	

as a group are separately identified, but "self-employed" is not categorized by whether with or without employees. Instead the self-employed are categorized by whether their business is incorporated or not.

391.   The question in example H.4 is a reasonably standard "status in employment" question with "paid employees" split into two government categories and a private category for an institutional sector classification, in line with the earlier discussion.

4.   Institutional sector: suggestions for enhancing the effectiveness of the questions

392.   The institutional sector categories used in combination with status in employment seem to work well for a census. The major categories suggested for the topic can be identified by dividing the category "employees" in the status in employment

Example H.2

H.2.1	**Are/were you self-employed or working for someone else in your main job?**
1	Self-employed (no paid helper)
2	Self-employed (1-4 paid helpers)
3	Self-employed (5 or more paid helpers)
4	Employee (GOVT/GOVT CORP)
5	Employee (private 1-4 workers)
6	Employee (private 5 or more workers) **Skip to X.X**
7	Unpaid family worker
H.2.2	**Do you move all your goods daily; e.g., fruits, peanuts, newspapers, clothing, equipment?**
1	Yes (informal trader) 2 No

Example H.3

H.3.1	**Was ... *Read list. Mark (X) one box.***
	Employee of PRIVATE-FOR-PROFIT company or business or of an individual for wages, salary or commission
	Employee of a PRIVATE NOT-FOR-PROFIT, tax-exempt or charitable organization
	Local or territorial GOVERNMENT employee (territorial/commonwealth, etc.)
	Federal GOVERNMENT employee
	SELF-EMPLOYED in own NOT INCORPORATED business, professional practice or farm
	SELF-EMPLOYED in own INCORPORATED business, professional practice or farm
	Working WITHOUT PAY in family business or farm

Example H.4

H.4.1   Were you self-employed or working for someone else in your (main) job last week?

Self-employed	Worked for someone else
With paid help (employer)	As government employee
Without paid help	As employee of a foreign government
	As employee of private company/person
	As unpaid worker in family business/farm
	Not stated

classification into "government employees" and "other employees". All other status in employment categories fall into the "other" category in the broad institutional sector classification. When formulating the relevant questions, the categories of the two variables can be combined with little difficulty as shown in the set of examples H.2 to H.4 above. Although that approach should not generally cause problems, it should be tested. Example H.4 may provide a good starting point for testing.

393.   There is a tendency, when institutional sector is combined with other characteristics, such as status in employment, to use technical terms, as this helps to shorten the sentences or response categories. That temptation should be resisted and an effort should be made to use locally understood terminology to improve the quality and consistency of responses. Additionally, when two characteristics are combined in a question, some categories of each characteristic may be lost. The questions need to be carefully tested for their validity and for their potential uses in future analyses.

5.   Institutional sector: issues of measurement

394.   If the suggestion above to combine questions for institutional sector and status in employment is adopted, the main points noted for testing on institutional sector are the following:

(*a*)   Respondents' interpretation of government, particularly regarding "semi-government" or "parastatal" organizations. Identify necessary prompts and notes on the questionnaire, or instructions for manuals and training, to try to record the meaning of those terms as desired in that particular country;

(*b*)   Test any subcategories of "government" or "other" (particularly "profit/non-profit" as in example H.3) that are to be identified separately.

C.   Informal sector and informal employment

395.   At the Seventeenth International Conference of Labour Statisticians in 2003, a new concept of informal employment was added to the existing concept of employment in the informal sector.[6] Both concepts are described in section C. The difference between the two concepts may be summarized as follows:

(*a*)   The informal sector is enterprise-based, and therefore informal sector employment relates to all people working in enterprises of a certain size, household ownership and other characteristics;

(*b*)   Informal employment is job-based and relates to all those who have a job with certain characteristics, generally those who are workers with no written employment contracts, and those in employment that is not subject to labour legislation, pension or social security regulations, collective agreements and so on.

396.   In countries where a programme of labour force surveys exists, it is suggested that the survey be used, rather than a census, to collect information on the informal sector and on informal employment. In countries where a short form and a long form are used in the census, it would be more appropriate to include any questions on the informal sector and informal employment in the long form, and not in the short form.

397.   The size and structure of the informal sector is a topic of considerable interest and importance in many countries. The sector is considered to engage 50 per cent or more of the persons in non-agricultural employment, either as a main or as a secondary activity. A number of countries are obtaining reasonable national estimates for the size and structure of the informal sector, mostly using a household survey methodology. As such statistics are also in high demand for different regions and special population

6 See ILO, "Guidelines concerning a statistical definition of informal employment", Seventeenth International Conference of Labour Statisticians, 24 November–3 December 2003, available from http://www.ilo.org/ public/english/bureau/ stat/download/guidelines/ defempl.pdf; and Ralf Hussmanns, "Measuring the informal economy: from employment in the informal sector to informal employment", Working Paper No. 53 (Geneva, ILO Bureau of Statistics, 2004).

groups, data analysts are very likely to attempt to develop measurements of the informal sector from census results, even if those responsible for the census do not plan for such estimates to be prepared.

398.   The Fifteenth International Conference of Labour Statisticians adopted the first resolution concerning statistics of employment in the informal sector (ILO, 1993b), including a detailed review of definitions and data collection procedures. Precise identification of those working in the sector cannot be achieved with a census, in part as a result of the limitation on the number of questions that can be included in a census and as a result also of heavy reliance on third party reporting for census data collection. It should be possible, however, to obtain some selected indicators (the size of the sector and the distribution of activities and occupations within it) on at least some segments of the informal sector. That is important because of the significance of informal sector activities in many local communities.

399.   As summarized in *Principles and Recommendations, Revision 2*, the informal sector is part of the household sector,[7] one of the main categories of the institutional sector of employment variable. Informal sector units are neither incorporated businesses, which have a legal entity separate from their owners or shareholders, nor are they quasi-corporations, that is, production units, which, though neither incorporated nor registered, do keep a reasonably complete set of accounts so that the finances of the business are kept separate from those of the household(s) that owns or controls the business.

400.   The informal sector was a recommended topic for the 2000 round of population censuses, but now an additional topic, namely informal employment, may also be considered for inclusion in the census.

401.   As stated above, informal employment is a new concept in the international standards. As shown in figure V, the components of informal employment include the following:

(*a*)   Own-account workers and employers employed in their own informal sector enterprises;

(*b*)   Members of informal producers' cooperatives, for example, cooperatives not formally established as legal entities;

(*c*)   Own-account workers engaged in the production of goods exclusively for own final use by their household (for example, subsistence farming, do-it-yourself construction of own dwellings);

(*d*)   Contributing family workers working in formal or informal sector enterprises;

(*e*)   Employees holding informal jobs, whether employed by formal sector enterprises, informal sector enterprises or as paid domestic workers by households. Employees are considered to have informal jobs if their employment relationship is, in law or in practice, not subject to national labour legislation, income taxation, social protection or entitlement to certain employment benefits (advance notice of dismissal, severance pay, paid annual or sick leave and the like).

402.   Therefore, to measure informal employment in a population census it is necessary to collect information, not only on the status in employment (whether paid employee, employer, own-account worker without employees, contributing family member, and so on), but also on the conditions of employment for employees, including security of tenure, access to pension or social security and paid leave.

403.   Data on both informal sector employment and informal employment are needed since both types of employment exist and are of growing importance in the economy.

[7] See also Commission of the European Communities and others (1993), paras. 4.132-4.150 and 4.159-160, and chap. IV, annex.

Figure V
Conceptual framework: informal employment

Production units by type	Jobs by status in employment								
	Own-account workers		Employers		Contributing family workers	Employees		Members of producers' cooperatives	
	Informal	Formal	Informal	Formal	Informal	Informal	Formal	Informal	Formal
Formal sector enterprises					1	2			
Informal sector enterprises[a]	3		4		5	6	7	8	
Households[b]	9					10			

Source: Ralf Hussmanns, "Measuring the informal economy: from employment in the informal sector to informal employment", Working Paper No. 53 (Geneva, ILO Bureau of Statistics, 2004).

Note: Cells shaded in dark grey refer to jobs that by definition do not exist in the type of production unit in question. Cells shaded in light grey refer to formal jobs. Unshaded cells represent the various types of informal jobs:

Informal employment:	Cells 1 to 6 and 8 to 10
Employment in the informal sector:	Cells 3 to 8
Informal employment outside the informal sector:	Cells 1, 2, 9 and 10

a As defined by the Fifteenth International Conference of Labour Statisticians (excluding households employing paid domestic workers).

b Households producing goods exclusively for their own final use and households employing paid domestic workers.

The concept of informal employment is particularly important in relation to the quality of employment conditions, which is an aspect of decent work.

404.   There is some informal employment outside the informal sector, including employees holding informal jobs in formal sector enterprises, paid domestic workers employed by households, contributing family workers working in formal sector enterprises and own-account workers engaged in the production of goods exclusively for own final use by their household.

405.   For the 2010 round of censuses, any attempt at measuring informal employment and, to a lesser extent, informal sector employment may to a certain extent be experimental. It will therefore be useful to compare country experiences with those topics in order to improve measurement methodology over time.

1.   Informal sector: operational definition

406.   The population employed in the informal sector comprises all persons, irrespective of their status in employment, who during a given reference period were employed in at least one business unit or establishment for which business activities were carried out by households and for which no complete set of accounts was kept, that is, their business transactions cannot be totally and clearly separated from other household transactions (United Nations, 2008b, para. 2.339).

407.   The population that is informally employed comprises those with the status in employment indicated in paragraph 401 above. For those who are paid employees, it is necessary to further identify those with reasonably secure employment contracts (that is, of long duration or even permanent contracts), whether they have social security protection and whether they are entitled to certain employment benefits (advance notice of dismissal, severance pay, paid annual or sick leave, and so on).

2.   Informal sector: requisites for applying the definition

408.   The informal sector should not be confused with the concept of the "hidden" or "underground" economy or with illegal activities. In most countries, most if not all informal sector activities are legal and above ground, at least in the sense that the only laws and regulations broken will be those relating to operational licences or the payment of taxes and social security contributions.

409.   The resolution adopted by the Fifteenth International Conference of Labour Statisticians in 1993 recognized a subdivision of the informal sector, as follows:

(*a*)   Informal own-account enterprises, which can employ contributing family workers and the occasional paid employees but do not have paid employees on a continuous basis;

(*b*)   Enterprises of informal employers, who can employ paid employees on a continuous basis, but usually only on a small scale, for example, with less than five employees.

Thus, there are own-account workers, contributing family workers and paid employees in the informal sector. The problem is how to identify, separately, individuals working in such activities when measuring employment in the informal sector.

410.   For many people, informal sector activities are secondary in the sense that they may also have another job, and that job may often be considered the main one, even if it does not necessarily provide the highest income. Therefore, to obtain a more complete measure of the size of the sector, both the main and secondary jobs of the individual should be covered. Where possible, the population employed in the informal sector should be subclassified into two categories: (*a*) persons exclusively employed in the informal sector; and (*b*) persons employed both in and outside the informal sector.

3.   Informal sector: review of questions used in national censuses

411.   There is little experience with collecting data on the informal sector in censuses and no experience in collecting data on informal employment.[8] One may expect that few countries will wish to burden their census questionnaire with the additional questions on the characteristics of the job contract and conditions of work that are needed to measure informal jobs of employees. However, information from questions on some topics included in the census, such as status in employment and institutional sector, are directly relevant. For the measurement of informal employment, additional questions would be needed on the employment contract (whether it exists and if so, the duration) and on access to certain employee benefits. The following comments on possible questions deal primarily with the person's main job (economic activity). The possibility of a question on secondary activities and a question at the household level respectively, with the aim of obtaining a better measure of both total informal sector employment and total informal employment, is covered under the suggestions for enhancing the effectiveness of the questions.

412.   **Questions relating to the person's main employment.** The informal sector is a characteristic of the private enterprise in which the person works. It is important to design and locate any proposed question so that it relates to the correct sub-group of the total population. It would be most efficient for the informal sector questions to follow an institutional sector question and for only those working for private enterprises to be directed to it.

413.   In example I.1, combining a "yes" response with several characteristics of the job complicates the basic question used to determine economic activity status. In defining the categories, both industry and institutional sector variables are combined. Terms like "formal" and "informal", which are not easily or uniformly understood by fieldworkers (or even by statisticians in general), should not be used in a question without provid-

ing explanations. Moreover, the formulation "informal unregistered" may be incorrectly understood by some, who will suppose that there is an "informal registered" category somewhere. It is unlikely that the approach in example I.1 will yield reasonably accurate estimates of a sector for which measurement requires a more elaborate combination of institutional sector, industry and legal requirements.

414.   To identify informal employment, one must first determine the status in employment and then, for paid employees, ask about characteristics of the employment contract and employment conditions.

415.   The question in example I.2 combines a partial informal sector identification within the sector classification by subdividing the self-employed by whether they work in a "not incorporated" business or an "incorporated" one. Making that distinction would assist greatly in identifying the self-employed in the informal sector but would not cover all the informal sector establishments. For example, unincorporated businesses keeping complete sets of accounts will not be identified, so they will be excluded; nor will any paid employees working in the informal sector be identified from question I.2.1.

416.   Questions would need to be asked to determine whether an employment contract existed and, if so, its duration and whether the paid employee was covered by social security provisions, had paid annual leave, paid sick leave, and so forth.

417.   Some have asked for the number of employees or of paid employees working at a person's place of work. The results can be used for approximate informal sector identification.

418.   In the flow of questions in example I.3, all employed persons are asked to state the number of workers in the establishment where they work. Those working in the informal sector can then be approximated from a combination of responses to questions I.3.1, I.3.2

Example I.1

Any work in the seven days before 10 October

I.1.1   In the SEVEN DAYS before 10 October did (the person) do any work for PAY (in cash or in kind), PROFIT or FAMILY GAIN, for one hour or more?

1 = Yes: formal registered (non-farming)

2 = Yes: informal unregistered (non-farming)

3 = Yes: farming

4 = Yes: has work but was temporarily absent

5 = No: did not have work

If YES go to X.X

Example I.2

I.2.1   Was ... *Read list. Mark (X) one box.*

☐ Employee of PRIVATE-FOR-PROFIT company or business or of an individual for wages, salary or commissions

☐ Employee of a PRIVATE NOT-FOR-PROFIT, tax-exempt or charitable organization

☐ Local or territorial GOVERNMENT employee (territorial/commonwealth, etc.)

☐ Federal GOVERNMENT employee

☐ SELF-EMPLOYED in own NOT INCORPORATED business, professional practice or farm

☐ SELF-EMPLOYED in own INCORPORATED business, professional practice or farm

☐ Working WITHOUT PAY in family business or farm

Example I.3

I.3.1 En el trabajo principal, ¿es …
obrero o empleado?

patrón?

trabajador por cuenta propia?

trabajador familiar?

I.3.2 ¿Trabaja en …
un empleo estatal (nacional, provincial o municipal)?

un empleo del sector privado?

Ignorado

I.3.3 ¿Por ese trabajo le descuentan para la jubilación?
Sí   I.3.6

No

Ignorado

I.3.4 ¿En ese trabajo aporta por sí mismo para la jubilación?
Sí

No

Ignorado

I.3.5 ¿Recibe sueldo?
Sí   I.3.3

No

Ignorado

I.3.6 ¿Cuántas personas hay en total en el establecimiento o lugar de trabajo?
1 a 5

6 a 39

40 o más

Ignorado

I.3.1 In his/her main job, the person is …
Worker or employee

Owner

Own-account worker

Family worker

I.3.2 Does the person work for …
The government (national, provincial or municipal)

The private sector

Unknown

I.3.3 In that job, does the person's employer make deductions for retirement?
Yes   I.3.6

No

Unknown

I.3.4 In that job does the person make his/her own retirement contributions?
Yes

No

Unknown

I.3.5 Does the person receive a wage?
Yes   I.3.3

No

Unknown

I.3.6 How many people altogether are there in the establishment where the person works?
1 to 5

6 to 9

40 or more

Unknown

and I.3.6 (employers and employees of the private sector, own-account workers and family workers, in enterprises that have five or fewer workers). It is not possible with this set of questions to distinguish between regular employees and other types of employees or unpaid family workers, and the only other criterion that can be used to define the informal sector, besides the number of workers, is whether deductions were made for the pension.

4. Informal sector: suggestions for enhancing the effectiveness of the questions

419. Information about a number of basic characteristics of the establishment is required, depending on the definition adopted and the degree of precision aimed for. For example, in addition to being privately owned, the unit (*a*) must be unincorporated; (*b*) must not keep separate accounts from household expenditures; and (*c*) must be small in size, that is, with respect to employment.

420. Any information that has been collected from population censuses can at best give only partial coverage of employment in the informal sector. Some countries have tried to identify segments of the informal sector from characteristics other than those discussed above. One example is using the location of the place of work, specifically by identifying those persons who have no fixed work location or those who are working at home. Another approach is to ask a specific question to identify one segment of the informal sector.

421. Example I.4 provides a very restricted view of the informal sector, and that may not have been the intent of the question. Non-retail activities (including handicrafts, bicycle repairs from home, taxi operation, etc.) would be excluded by this question. The question may be useful for broader planning in some countries with a particular focus on a subsector of the informal sector.

422. Almost all informal sector activities can be carried out either as a main or as a secondary activity for the individuals involved. In some countries the number engaged in the informal sector as a secondary activity is almost as large as those engaged in it as a main activity. With the difficult economic circumstances in many countries, it is only by engaging in secondary activities that many who work as paid employees as the main activity can survive and support their families. In many cases one member of a family will engage in the informal sector activity as a main activity and other household members will assist as secondary activities. It should be noted, however, that very few countries have asked census questions about secondary economic activities. One country asked the complete block of economic questions again with regard to secondary activities, but for most countries there will probably be insufficient space for that approach to be used. If the complete set of questions was asked for any secondary activities, the same procedures presented above could be used for identifying persons engaged in the informal sector as a secondary activity.

423. Given the limited space and already curtailed number of questions, even for better tested questions on economic characteristics, the topic of the informal sector is best left for sample surveys designed specifically to collect this type of data. It is definitely not possible to ask a simple question such as "Do you work in the informal sector?" and expect to obtain reasonable results. Thus, a probable approach would be to (*a*) examine what can be obtained by using the other topics and variables and their classifications which

Example I.4

I.4.1 **Do you/does he/ she move all your/his/her goods every night: e.g., fruits, nuts, lottery tickets, clothing/shoes, etc.?**

1 ☐ Yes (informal trader) 2 ☐ No

are included in the census, and (*b*) look into possible additional questions and response categories that may be used to give more refined or precise identification of those working in the informal sector. It is strongly recommended that countries use this approach. Those steps should be taken whether or not it is decided to make a serious, official attempt at obtaining informal sector measures from the census, as users of the census results are very likely to ask for estimates of informal sector employment from the census, both nationwide and in local labour markets.

424.   If a more comprehensive measurement of employment in the informal sector is preferred (for example, in a census long form or for a country without a labour force survey), then example I.5, which is based on an actual labour force survey, might be considered. Similarly, example I.7 relates to the measurement of informal employment and is also drawn from a country's labour force survey.

425.   In example I.5, a series of filter questions gradually narrows the focus to the level of informal sector units. First, institutions and incorporated companies are identified separately from non-corporations (economic units that are owned by a sole person or conducted on a family basis). Units in the latter group may include quasi-corporations; the example has a question to identify how their accounts are kept[9] and hence whether they are informal or not. For informal sector units, there are questions that identify employment size, operational conditions and place of work. There are also questions on place of work for people in enterprises that are not in the informal sector. It is worth noting that the whole section applies to all people employed regardless of their status in employment. Finally, the example does not treat agricultural and farming activities as "informal sector", but some countries may elect to include those activities in accordance with the flexibility outlined in the international standards.

426.   For most small business operators in the informal sector, whether employers or own-account workers, the questions in example I.5 would not be difficult. For contributing family workers, interviewers would have to be trained to ensure that the answers will be the same as for the person(s), usually in the same household, who is/are the self-employed

9 In many countries, the number of quasi-corporations is low, and a considerable number of countries that measure employment in the informal sector in their surveys do not ask about account-keeping practices.

Example I.5

Hypothetical questions, recommended for testing and use in countries

I.5.1   **What is the name of the enterprise, business or institution you are working for?**
(Please write complete name of the enterprise, business or institution)

1. _____

2. The business does not have a trade name

3. It is a domestic unit (only for paid domestic workers) → **Go to the next section of the questionnaire**

9. Does not know

I.5.2   **What does the enterprise, business or institution do?**
(Please give details of type and materials of goods produced or services provided)

_____

I.5.3   **For interviewer only:**
(Please classify according to answers given in I.5.1 and I.5.2)

1. It is an agricultural or farming activity → **Go to the next section of the questionnaire**

2. It is an educational institution, or a hospital or a clinic → **Go to I.5.5**

3. It is a public institution or a non-profit organization → **Go to I.5.5**

4. It is a private activity or firm → **Continue**

5. Cannot be determined yet → **Continue**

I.5.4 **Is this business …**

(Please read the options and circle the one selected)

1. An independent, personal or family business → **Go to I.5.6**

2. A private incorporated company (trading company, corporation or corporations under other modality, banking company)? → **Go to I.5.10**

3. None of the above → **Continue**

I.5.5 **Do you work for …**

1. A government agency → **Go to I.5.10**

2. Non-government-managed institution → **Go to I.5.10**

9. Does not know → **Go to I.5.10**

I.5.6 **How many people work in this business or activity including you?**

(Please read the options until a response is given and circle the one selected)

1. Only one person

2. Two to five

3. Six to ten

4. Eleven to fifteen

5. More than fifteen

9. Does not know

I.5.7 **This business or economic activity …**

(Please read the options and circle the one selected)

1. Has a premise, site or office → **Go to I.5.9**

2. Does not have a site, nor office nor premise → **Continue**

9. Does not know → **Go to the next section of the questionnaire**

I.5.8 **So, where do you perform your work?**

(Please listen and circle the option mentioned)

01. Door to door

02. In the streets without a fixed place

03. In the street but always in the same place

04. Non-motorized vehicle (bicycle, tricycle, cart, boat, etc.)

05. Motor vehicle, automobile, motorcycle, light truck, etc.

06. At own address without any special equipment

07. At own address with special equipment

08. At the owner's address or the owner's property, or at place requested by client

09. Other site

99. Does not know

Go to the next section of the questionnaire

I.5.9 **In this business or activity …**

(Please read the options and circle the one selected)

1. There are book accounts kept or there are hired accountant services

2. Only a notebook or a booklet to write the business records is kept

3. No accounting records are kept

9. Does not know

I.5.10 **Where do you usually work?**

(Please read the options and circle the one selected)

1. At the facilities (or in the vehicles) of _____

(Please mention the firm's name given in question I.5.1)

2. At the facilities of another firm or enterprise, or at any institution where you are sent

3. Visiting at clients' location

4. At a construction site

9. Does not know

operator(s) of the same business. Paid employee respondents are much less likely to have the required knowledge about the characteristics of the business activity, such as whether the business is registered as a company or keeps a full set of accounts. It is preferable to measure employment size in terms of the number of regular paid employees, but employee respondents may not correctly differentiate between regular paid employees and unpaid family workers, although they may know approximately how many workers are engaged on a continuous basis. Since the reply to questions other than that concerning the number of employees will most often be "Don't know", it should be established through testing whether it is worthwhile to ask the complete set of questions of the employees. Problems are likely to be even more severe in third party reporting, that is, if another person answers questions on behalf of the actual respondent.

427.   **Identification of informal sector activities at the household level.** In many cases informal sector activity is a household activity with several members of the household engaged, some full-time and some part-time. Identification of the number of households with some engagement in informal sector activity, preferably by type of household and activity, is extremely valuable information in itself and is also useful for creating a sample frame for later informal sector surveys.

428.   Household information can be obtained during listing preparations for the census, provided the following problems, if they arise, can be resolved:

(*a*)   It is often difficult to match the individual household information from the listing with the replies obtained for the household during the census enumeration;

(*b*)   Informal sector activities are often intermittent. People stop and start easily, particularly with respect to the smaller informal sector activities and especially those that are seasonal. Listing information collected over a year or more prior to the census, as is often the case, may not relate to economic activities at the time of enumeration.

429.   Thus, although listing data can be useful both for the identification of the informal sector and for the provision of information on other enterprise characteristics (see *Principles and Recommendations, Revision 2*, paras. 1.51-1.52), it is far more useful to ask the question during the actual enumeration. The following question is hypothetical and untested, and is provided only for illustrative purposes.

430.   The question in example I.6 targets particular types of informal sector activities. This is an effective approach and it can be extended to other categories. It would have to be changed for each country to fit local examples and priorities. It could also be asked generally about small-scale business activities with a variety of examples.

Example I.6

Hypothetical questions, recommended for testing and use in countries
I.6.1   Did any member of the household engage in any of the following small-scale business activities in the last seven days? ☐   Making or repairing any items (e.g., clothing, pots, hats, furniture, metal products)? ☐   Preparing food or beer for sale? ☐   Selling food, clothing, etc. at the market, in a small shop, etc.? ☐   Transporting people or goods, using taxis, hand carts or the like? *Interviewer*: *For each of the above record one of the following codes* Yes ... 1 No ... 2 *Note*:   *Must be own or family business with less than five employees on a regular basis.* *Not working as a paid employee.*

431.   Example I.7 is a set of questions that were designed to measure informal employment in a country's labour force survey. Such a detailed set of questions will generally not be appropriate for inclusion in a population census, but they may be useful if a programme of labour force surveys does not exist.

5.   Informal sector: issues of measurement

432.   As this is a new topic in most countries and, in particular, in a census, any questions related to informal employment or the informal sector should be subjected to thorough testing. It is strongly suggested that a test of any such questions be followed by complete re-interviews using a detailed household survey questionnaire to check whether the identified informal sector activities really are informal sector activities, whether any relevant informal sector activities were missed and whether the employment was really informal in nature. Tests would need to cover a variety of rural and urban areas and various types of informal sector activities. They should be combined with test derivations of informal categories during processing and analysis, as explained below.

Example I.7

Hypothetical questions, recommended for testing and use in countries

I.7.1	**Is [Name] employed on the basis of a contract or agreement?**
	Yes 1 → **Continue**
	No 2 → **Go to I.7.4**
I.7.2	**Is the contract or agreement of a limited duration?**
	Yes 1
	No 2 → **Go to next section**
I.7.3	**What is the duration of the contract or agreement?**
	Daily 1
	Less than a month 2
	1 to 2 months 3
	3 to 6 months 4
	7 to 12 months 5
	Over 12 months 6
I.7.4	**Does [Name]'s employer pay social security contributions for [Name] (pension fund, unemployment fund)?**
	Yes, certainly 1
	Yes, possibly 2
	No 3
	Don't know 4
I.7.5	**Does [Name] benefit from paid annual leave or get compensation for unused leave?**
	Yes 1
	No 2
	Don't know 3
I.7.6	**Does [Name] benefit from paid sick leave in case of illness?**
	Yes 1
	No 2
	Don't know 3
I.7.7	(*For women only*) **Does [Name] benefit from maternity leave if she wishes to have a baby?**
	Yes 1
	No 2
	Don't know 3

433.   At present the only reasonable option in many censuses is to derive estimates of employment in the informal sector from answers to other questions, as no specific question designed to identify the informal sector is known to be have been asked. The responses to the status in employment question, for example, provide the most useful information for such derivations. The self-employed with no employees (own-account workers) and contributing family worker categories are the most useful ones for identifying possible informal sector workers. If those engaged in agricultural work can be excluded by using the appropriate industry codes or other means, then most of the remaining persons in employment in the two categories will be working in the informal sector in many developing countries. Careful consideration of other industry and occupation categories, combined with local knowledge, may lead to the exclusion of other jobs and a more refined measure. However, similar possibilities are difficult to find for the self-employed with employees and for paid employees. Thus, this approach is useful for obtaining rough measures for only two components of the informal sector workforce. However, for some countries it may be the only approach they can afford, and it may be better than making no efforts at all. The estimates can be further refined if the number of employees (less than or more than the size limit defined for the informal sector) is included in the status in employment categories.

434.   Similarly, estimates of some of the components of informal employment, as shown in figure V, can be made from the answers to the question on status in employment. However, the identification of employees with informal jobs will not be possible without more detailed questions, and it will still be necessary to identify those employed in informal sector enterprises.

435.   In most of the countries with a large informal sector, analysts (official or otherwise) have frequently tried to obtain statistics on the informal sector from previous censuses by making use of statistics classified according to sector, occupation, industry and status in employment, if available, and on that basis derive the groups that they think will approximate the informal sector or some of its activities. For example, it would be possible to take the self-employed who are coded under the industry code "street sellers and hawkers" to approximate that segment of the informal sector.

436.   It is easier initially to identify what is not in the informal sector, and to make sure that the persons engaged in those activities—formal sector and subsistence activities—can be identified separately and thus excluded from consideration either for additional questions and/or during the tabulation stage. The following two groups can clearly be excluded from the informal sector or can be separately identified, using information obtained with the questions on status in employment, sector of employment, industry and occupation:

(*a*)   **Employees of recognizable formal institutions.** All paid employees working for the government, or semi-government or parastatal enterprises (or NGOs or international organizations where significant). These employees can be identified principally from the appropriate sector and industry categories;

(*b*)   **Persons engaged in agricultural or non-market production activities** as own-account **workers** or paid employees, or as private household employees;

(i)   **Agricultural activities.** All agricultural activities, whether undertaken for home consumption or for the market, including those carried out in unincorporated enterprises, which are part of the household sector, are usually excluded from the informal sector. Population censuses can and usually do, however, provide valuable measures for the identification of persons engaged in such activities. Industry classifications based on ISIC, Revision 4, may be constructed so as to allow separate

identification of subsistence or traditional agriculture and commercial agriculture. Occupational classifications based on ISCO-08 may also allow for a distinction between market-oriented agricultural workers and subsistence agricultural workers, where that distinction is important. It is recommended that, if possible, persons engaged in agricultural activities in the household sector for home consumption should be identified separately from those involved in production for the market, and excluded from statistics and discussion on the informal sector;

(ii) **Non-market production activities.** These activities include the processing of primary products for home consumption, such as making butter from milk; making oil from oil seeds; making mats from straw or reeds; weaving cloth from fibres; and building or carrying out major repairs on the family farm or house, for example, making a fence or digging an irrigation ditch; and fetching water and collecting firewood. Persons engaged in the preceding activities are to be counted as employed if such activities are important in the national economy and should be included in the household sector. Since these non-market activities are also closely tied to agricultural activities, it is recommended that they not be included in the informal sector. Such processing and building activities, however, are often done partly for sale (to be considered for inclusion in the informal sector) and partly for home consumption (not informal sector). They may be the only or main economic activity in farming communities during certain periods, such as in the agricultural off season. Such activities cannot usually be separately identified by the use of standard industry and occupation codes. Jobs in establishments whose only or main activities, during the reference period for data on economic characteristics, consist of processing or building for own use should be coded to the appropriate industry category (for example, manufacturing, construction). Jobs that only involve processing or construction work should similarly be coded to the appropriate occupation group (for example, handicraft workers, building construction labourers). Workers in these jobs will be difficult to differentiate from those who perform similar work for a market and are thus to be considered for inclusion in the informal sector. The issues outlined above should be carefully considered in the design of census questionnaires and procedures, particularly in countries where such non-market activities are significant. Interviewers will need guidance on how to record specific situations, to ensure that special codes are used to determine whether jobs should be classified in the informal sector. Note that questions on status in employment may be designed so as to identify persons working on their own or family farm separately from those working in small-scale non-agricultural businesses that may be in the informal sector;

(iii) **Employees of private households.** Paid employees of private households, such as maids, cleaners, guards and so forth, are included in the household sector and are in some countries considered to be part of the informal sector. Consequently, the resolution (see ILO, 1993b) recommends that they should then be shown separately. A special industry code is usually used for such employees, for example code 9500 in ISIC, Revision 3 (code 9700 in ISIC, Revision 4). The occupational classification should ideally have separate codes for each of the occupations,

so that this component of household sector employment can easily be identified through cross-classification with the industry variable. Then the users of the statistics have the option of keeping them within or outside the informal sector according to descriptive or analytical needs.

437.   **Employees of private households.** Paid employees of private households, such as maids, cleaners, guards and so forth, are included in the household sector and are in some countries considered to be part of the informal sector. Consequently, the resolution (see ILO, 1993b) recommends that they should then be shown separately. A special industry code is usually used for such employees, for example code 9500 in ISIC, Revision 3 (code 9700 in ISIC, Revision 4). The occupational classification should ideally have separate codes for each of the occupations, so that this component of household sector employment can easily be identified through cross-classification with the industry variable. Then the users of the statistics have the option of keeping them within or outside the informal sector according to descriptive or analytical needs. Informal sector employment is derived as follows:

$$\text{Total workers} - RIC_p - OBCA_p - COM_p{\geq}x = \text{Workers in private businesses}$$

where RIC_p = Persons working in registered incorporated companies;

$OBCA_p$ = Persons working in other enterprises (businesses) with complete accounts;

and $COM_p{\geq}x$ = Persons working in companies with x or more paid employees on a continuous basis.

438.   The specified exclusions cannot usually be determined from the answers to questions on other topics included in the census. Relevant questions would thus have to be asked, and the types of questions possible and whether respondents can be expected to have that knowledge were considered previously in the present chapter.

439.   The upper limit on the number of employees for an own-account enterprise to be regarded as an enterprise of the informal sector is usually decided in each country according to the local circumstances. If there is a business register that supports employer- and/or establishment-based statistical surveys, it often includes only establishments above a certain size, often expressed in terms of the number of paid employees. If that is the case then the informal sector should be complementary to the scope of that register. In countries where there is a value added tax (VAT) and a lower limit of total sales for VAT registration, then such registration may be a possible criterion for defining the scope of the informal sector, at least if no other reasonable formalization criteria are available (note that with VAT non-registration as the criterion there will be informal sector units with total sales larger than the threshold for VAT registration, either because of deliberate tax evasion or because the registration requirement has not yet been complied with). Surveys have shown that the number of informal sector enterprises declines rapidly as the number of paid employees on a continuous basis increases, since most enterprises with high levels of employment are usually not informal for one of the reasons given above.

Chapter VIII

Working time and income

440.   In general, the measurement of working time and of income from employment should relate to the total hours or income over all jobs. However, often those characteristics are measured only for the main or "reference" job. Once the main job is identified, it is essential that all characteristics refer to the same main job. Questionnaire design must be such as not to confuse respondents or interviewers on that point. When any secondary activities are identified they must be clearly separated and placed so as not to confuse the flow of questions relating to the main job. Doing so is not usually difficult with current activity (see examples) but can be difficult with usual activity.

A.   Working time

441.   "Working time" is a very valuable topic for many countries, both developed and developing. Working time covers both working hours and working time arrangements. A growing proportion of employees and self-employed persons have non-standard working hours or arrangements. In those situations, observing only the number of persons employed will provide a rough estimate of the total amount of work done. However, it is necessary to collect information about the hours actually worked in all jobs by those employed to obtain estimates of labour input with any reasonable degree of precision and to provide comparability over time and between groups. Information about working time is also essential for studies of its relationship with income, underemployment and productivity; for some national accounts measures; and for assessment of new working time arrangements and the extent to which excessive hours are worked. A question about working time was asked in about half of countries in the 2000 round of censuses. It is recommended that working time for all economic activities, main as well as secondary, be recorded for all employed persons and not just for paid employees.[1]

442.   Statistics on normal hours of work are useful to users who are interested in the prevailing working time situation, especially in comparison to situations when respondents were absent from work for all or part of the reference period (for example, due to illness or vacation) or worked longer hours in the reference week. As with the measurement of hours actually worked (see below), the collection of data on normal hours of work should preferably cover all jobs (if more than one) and should be collected for all employed persons and not just for employees.

443.   Many countries use household surveys to obtain national estimates of working time. Conceptually better measures and a larger variety of measures can be obtained by the more thorough yet more time-consuming methodology that such surveys can use, in particular to measure the key concept of hours actually worked. However, it is primarily through a census that measures for small areas and for small groups of the employed population may be obtained. Measures of working time, such as hours paid for, contractual hours, overtime, absence and the like, from surveys of establishments or enterprises usually cover only large, formal employers and cannot easily provide breakdowns by age, sex or occupation. They may also measure working time only in respect of jobs and not persons.

[1] In 2008 at the Eighteenth International Conference of Labour Statisticians, revised international standards concerning the measurement of working time will be discussed. The proposals will include the additional measurement of time on activities outside the SNA operational production boundary, such as cleaning, cooking, repairs, transport and care of dependent persons, which is a significant proportion of often unmeasured productive activity.

1.   Working time: operational definition

444.   Working time is the total time actually spent producing goods and services or performing activities related to or enhancing production, including time spent on short pauses that may be in the nature of work-related breaks or rest periods during the reference period adopted for economic activity in the census. It is recommended that if the reference period is short, for example the week preceding the census, working time should be measured in hours, as hours actually worked. If the reference period is long, for example the 12 months preceding the census, working time should be measured in units of weeks, or of days where feasible. It is, however, not recommended that any attempt be made to establish the actual hours worked for such a long period, as recall and calculation errors are likely to be high. The preferred measure in that instance should be hours usually worked.

2.   Working time: requisites for applying the definition

445.   Working time should include the total time spent producing goods or services during the reference period, including any overtime, time spent in preparation for work (which includes training to facilitate production for the economic unit) or in maintenance or repair of equipment, as well as short breaks. For example, teachers should include their preparation and similar work-related time in addition to the time actually spent in front of a class, and so should persons on call in essential services (such as health and security) whose movements and/or behaviour are restricted even if not at work. Hawkers and other sales workers should include time waiting or looking for customers. However, long meal breaks should be excluded as should time off for holidays, vacation, sickness or industrial disputes. Time spent commuting (to the workplace and back) should be excluded unless work is carried out, such as during a train ride, and travelling is part of the job, for example, for taxi drivers, and workers travelling to or between variable work sites, such as field projects, fishing areas or meetings. The total working time should include all activities defined as economic activities as detailed earlier, and if practical, time spent on secondary activities should be recorded separately from the time spent on the main activity. Separate recording will allow for more accurate tracking of time and will also provide a sound basis for analyses comparing wages and salaries for different categories of workers.

446.   If the block of economic questions asks about current activity, then the question(s) about working time should also relate to the current (short) reference period. The objective thus is to measure hours actually worked in the current reference period. Such measures will include some zero hours for persons temporarily absent from work (on holidays, sick leave, and so forth) throughout the reference period. Reduced hours as compared to "usual" hours will also be recorded if any part of the reference period is taken, for example, for holidays or sick leave. All overtime, paid or unpaid, should be included. Thus the concept of working time as it relates to the current (short) reference period is a measure of hours actually worked rather than hours paid for or normal hours established by laws or regulations.

447.   Although *Principles and Recommendations, Revision 2* (see paras. 2.323-2.325) recommends measuring the hours actually worked during the short period just before the census reference date, some data users may prefer hours usually worked in a corresponding short reference period. That would be a measure of the hours worked during a normal or typical week, including overtime hours regularly worked, whether paid or unpaid. Days and hours not usually worked should be excluded as should unusual periods of overtime. In some censuses, questions have been asked about both actual and usual hours worked. Users of the statistics should be consulted regarding the most useful measure(s). Reference should also be made to the ILO international standards for measurement of working time (expected to be updated in December 2008 as a resolution of the International Conference of Labour Statisticians).[2]

2 The most up-to-date information on international standards for labour statistics can be found on the ILO website, http://www.ilo.org/global/lang--en/index.htm.

3. Working time: review of questions used in national censuses

448.   One or more questions on working time have been asked in censuses, as follows:

    (*a*)   A single question on actual or usual working time during the reference period, on the main activity or all economic activities combined (examples J.1 to J.3);

    (*b*)   Separate questions on working time concepts for the main activity and for other (secondary) activities (example J.4).

449.   When the questionnaire uses only a single question, the accuracy of the answers for working time will usually be poor. Asking and recording the hours actually worked each day in the reference period and in each job can improve the accuracy. However, that procedure takes more time and space on the questionnaire than is often possible in a census operation.

450.   The question in example J.1 combines a basic activity question with the collection of data on working hours. Although J.1.1 prompts very well on some of the problem categories of economic activity and the subsequent questions (not shown) identify temporary absence from work and probe for unemployment, it does not allow for recalling other activities in addition to the main economic activity. In this setup, the emphasis is more on determining who did "any work" than at measuring working time.

451.   Question J.2.1 deals with hours actually worked. It is not clear, however, whether all work done or only the main job is to be included. The responses to this question will most likely be mixed, with some individuals focusing on one job and others including all jobs. Training of interviewers will need to aim at standardizing the way respondents are guided to provide the intended information.

452.   Question J.3 makes it clear that all employment should be included. It deals with hours usually worked rather than hours actually worked. Depending on how com-

Example J.1

J.1.1	**Last week, how many hours did this person spend working for pay or in self-employment?**	◯ None ▶ **Continue with the next question**
	Include:	
	• *Working directly towards the operation of a family farm or business without formal pay arrangements (e.g., assisting in seeding, doing accounts);*	**OR**
	• *Working in his/her own business, farm or professional practice, alone or in partnership*	◀ Number of hours (to the nearest hour)
	• *Working for wages, salary, tips or commissions.*	**Go to question X.X**

Example J.2

J.2.1	**How many hours did you/ .... work during that week?**
	⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
	⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
	⑨ ⑨ Don't know/Not stated → **Skip to X.X**

Example J.3

J.3.1	**How many hours per week did you generally work, in all places of employment?**
	Hours a week ☐☐ I worked

mon multiple job holding is, it may be more effective to separate the main job from other jobs, to help respondents' recall.

453.   The questionnaire from which example J.4 was taken was designed for data processing using an optical mark reader (OMR). It has a relatively detailed question on working hours and asks for main and secondary activities separately. It asks for the hours usually worked using the term "normal" hours, rather than the hours actually worked in the last week. A complete series of questions would be to ask for the actual hours first for both main and secondary activities separately and then the hours normally worked. The sequence of actual followed by usual hours has been shown to improve accuracy in labour force surveys. Although the questions in example J.4 prompt for the inclusion of overtime, the example might also have benefited from the inclusion of prompts on types of economic activities, as in example J.1 above.

4.   Working time: suggestions for enhancing the effectiveness of the questions

454.   Example J.5 asks respondents to recall the days on which they worked, whether there was any time off and whether overtime was worked. Those questions are likely to produce a more accurate recall of the actual hours worked in addition to providing information about overtime.

455.   It is recommended to ask and record an estimate of the number of hours actually worked in all jobs. The answer recorded should normally be rounded to the nearest hour as working time expressed as decimals or fractions of an hour may be inaccurate. The asking and recording of individual hours also allows a review of the detailed distribution of working hours. That information will be particularly useful if some users of employment statistics wish to set a limit higher than the recommended one-hour criterion[3] for the minimum hours worked for those persons who are to be considered as employed in certain tabulations, to analyse full- and part-time work or excessive working hours, or to measure working time in relation to activities outside the SNA operational production boundary, including "unpaid non-market services".

456.   However, in some countries pre-coded ranges have been used for the answers to the question on working time to help the respondents. If that is done, then it is desirable to provide categories (based on national experience, such as from the labour force surveys,

3 The available evidence, usually from labour force surveys, of the distribution of workers according to hours actually worked has generally shown very low numbers of persons working very few hours per week, for example, below 10 hours.

Example J.4

J.4.1   How many paid jobs did you work at during the week of 13th-19th May?

☐ Jobs   ⓪ Not stated

① ② ③ ④ ⑤+

J.4.2   How many hours do you normally work in your (main) job in a typical week, including overtime whether you are paid for it or not?

☐☐ Hours   ⓪ Not stated

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

For persons with two or more jobs

J.4.3   Excluding your main job, how many paid hours do you normally work in your other job(s) in a typical week?

☐☐ Paid hours, other job(s)   ⓪ Not stated

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

if any) for replies that cluster on particular values, such as 35, 40 or 45 hours per week. That procedure may, however, provide another source of error as some interviewers with inadequate training may record the wrong category. There will also be a rounding error, as the midpoint of each range has to be taken as the basis for the calculation of averages and other descriptive statistics.

457.   Example J.6 provides a model that will allow for the collection of data on the amount of paid and unpaid overtime usually worked in addition to usually worked hours.

458.   It is not recommended that a question using the categories "part-time" or "full-time" be included in the questionnaire. Even with notes to interviewers, respondents will often apply their own inconsistent interpretations of these terms. Output categories for tabulations can of course be called full-time and part-time after reviewing the overall distribution of the working hours concept(s) measured in the census.

Example J.5

Hypothetical questions on hours actually worked, recommended for testing and use in countries

Hours actually worked

J.5.1   On which days did [YOU] work last week (in all jobs)?

	Yes No		Yes No		Yes No		Yes No		Yes No		Yes No		Yes No
Mon	☐ ☐	Tue	☐ ☐	Wed	☐ ☐	Thurs	☐ ☐	Fri	☐ ☐	Sat	☐ ☐	Sun	☐ ☐

J.5.2   Did [YOU] have any time off from [YOUR] job(s) on those days?

Yes ☐   No ☐

J.5.3   Did [YOU] work any paid or unpaid overtime on any day last week?

Yes ☐   No ☐   Other ☐

J.5.4   How many hours did [YOU] actually work last week (not counting time off but counting overtime)?

☐ ☐

Example J.6

Hypothetical questions on hours usually worked, recommended for testing and use in countries

Hours usually worked

J.6.1   How many hours a week do [YOU] usually work in your (main) job (business), that is, excluding meal breaks and overtime, whether paid or unpaid?   ☐ **[A]**

Interviewer: If work pattern is not based on a week, make notes and give an average over the last four weeks.

J.6.2   In addition to this, do you work paid overtime?

Regularly   ☐
Occasionally   ☐
Never   ☐

J.6.3   On average, how many hours of paid overtime do you usually work in a week?   ☐ **[B]**

J.6.4   Do you work unpaid overtime?

Regularly   ☐
Occasionally   ☐
Never   ☐

J.6.5   On average, how many hours of unpaid overtime do you usually work in a week?   ☐ **[C]**

Interviewer: Total [A] + [B] + [C]   →   ☐

5.   Working time: issues of measurement

459.   For countries where working time is a new census topic, it is strongly recommended that it be included in the pre-testing programme. Testing should focus on evaluating the accuracy of the answers by conducting a second interview using more detailed household survey questions and the wording, layout, prompts and examples most suitable for improving accuracy.

460.   It is, however, worthwhile to train interviewers to ask about each day and then add up the hours for all the days in the week before recording the answer. For rural farmers in many countries, interviewers will have to impute the time from other information, as such workers often do not have clocks and watches and can only answer in terms of sunrise, sunset, half-days and so forth. The time spent on non-market activities, such as production for home consumption and fetching firewood and water, will be very difficult to assess accurately. In their census, some countries collect working time only for paid employees and rely on other sources, such as labour force or time-use surveys, for measuring the hours worked of persons with less standardized or supervised activities.

B.   Income

461.   Income is not a common topic in population censuses. It was asked in only about a quarter of the countries in the 2000 round of censuses. The topic is recommended in *Principles and Recommendations, Revision 2*, but that source states that the "collection of reliable data on income, especially income from self-employment and property income, is extremely difficult in general field enquiries, particularly in population censuses. The inclusion of non-cash income further compounds the difficulties. Collection of income data in a population census, even when confined to cash income, present special problems in terms of burden of work, response errors, and so forth. Therefore this topic is generally considered more suitable in a sample survey of households. Depending on the national requirements, countries may nonetheless wish to obtain limited information on cash income" (see para. 2.332).

462.   The value of home production and the value of income paid in kind (such as food, clothing and accommodation), are extremely important components of the total income for individuals and households in most developing countries, but measurement is extremely difficult and complex even in household surveys. Asking questions on those categories is not the most efficient use of resources in respect of interviewing time and is therefore not recommended for a population census.

463.   The measurement of cash income itself raises many problems, and the major ones are covered below. For economically active persons it is recommended that only earnings over all jobs be collected for the individual. However, it is also recommended that, if possible, an item of total household income be collected, including any cash income from interest, dividends, rent, social security benefits, pension and life insurance annuity benefits.

464.   It should be noted that censuses in some developed countries have a long history of collecting detailed cash income information at the individual level. When detailed data are collected by source of the income, the results will allow the identification of economically depressed areas in addition to providing statistics on the source of the incomes in these areas. That information, used together with other characteristics from the census, will help in the development of social and economic policies to assist such areas and/or to monitor the effects of existing programmes to help them.

465.   Income is often thought to be a sensitive topic, but the degree of sensitivity seems to vary significantly between countries as well as between groups within the same

country. Some groups, such as owners of relatively large businesses, sometimes refuse to answer questions on the topic in any format, but paid employees and informal sector operators have been generally cooperative in most countries, within the limitations of their capacity to provide accurate estimates of their income, since they may not have written records. In at least one developed country, arrangements are made with the approval of respondents to access income tax records directly from the tax office in order to minimize respondent reporting burden and to improve data quality.

1.   Income: operational definition

466.   Income can be defined as follows:

(*a*)   For workers in paid employment, gross monthly earnings in cash and/or in kind from the work performed by each employed person. The cash[4] income figure should also always be the gross amount (that is, before deducting any contributions to pay income tax, social security, pension, or union dues and similar types of deductions). For employees in the formal sector, their gross cash earnings may be available from their pay slips. The concept of employment-related income of workers in paid employment is a broader concept than earnings and includes all benefits received from employment-related social security (net of contributions made) as well as benefits received as a result of previous employment;

[4] The term "cash" includes cheques, bank deposits and other forms of monetary income.

(*b*)   For the self-employed, income from employment should be the gross output or takings, minus any operating expenses (for example, for staff wages, materials and supplies, electricity, gas, water). In theory, depreciation of capital assets and production subsidies should also be deducted, but it will be possible to obtain that information only from the self-employed who operate formal businesses with formal accounts. As with earnings, the income from the employment of the self-employed should be recorded gross, that is, before any deductions are made for contribution to income taxes, social security, pension or any other similar types of contributions. Employment-related income for the self-employed includes all employment-related social security benefits (net of contributions) as well as benefits received as a result of previous employment;

(*c*)   The total annual income in cash and/or in kind of households regardless of source. In addition to the employment-related income of employed members, the total income of the household should include income that does not result from work, whether it is received by persons who work or not. The total would therefore include, for example, interest, dividends, rent, social security benefits that are not covered in (*a*) or (*b*) above, pensions, transfers (such as remittances and income from begging) and life insurance annuity benefits of all members of the household.

2.   Income: requisites for applying the definition

467.   What income to measure is a major conceptual and practical issue. It is useful to distinguish between the following:

(*a*)   Income from paid employment;

(*b*)   Income from self-employment;

(*c*)   Other income (investment income, including rental income; income from social security and insurance schemes; pensions; and so forth).

468.   The concepts involved in determining income are not simple to grasp, and respondents may be unable or unwilling to provide exact information. Even if respondents are willing to provide the information, it may rarely be possible for employees to report gross figures as they are most commonly aware of only the net amounts received,

or for the self-employed to report their mixed income for the relevant reference period. If the aim is to measure solely income from current or usual employment, then a measure of items (*a*) and (*b*) is sufficient and the questions need be targeted only at the employed population (current or usual). If welfare is more the concern, then the other incomes must be included and the question addressed to all individuals. However, that will make the questionnaire more complex. Note that rental income is not uncommon in many low-income areas of developing countries where a significant number of the aged or widows often earn a living by the renting of rooms and the like. Obtaining estimates of the rental income is, however, often very difficult.

469.   One approach is to ask individuals for income from paid employment and self-employment and to obtain a broad measure of total cash income at the household level.

470.   Ideally, current income would be measured with current employment and usual income with usual employment; that is to say, the income measure should correspond to the employment measure. Except for some regularly paid employees, very short-term income measures (for example, for the last week or the last month) can often be unrepresentative. Rarely is the same reference period for income received suitable for all the employed. For owners of large businesses, a year is often the most convenient as they can obtain that information from their accounts or tax records. A year is also the best period for income from activities with significant seasonal variations in earnings, such as agriculture. However, for small businesses with limited records, particularly in the informal sector, such a yearly measure is almost impossible to report. For employees an annual figure is also difficult to report unless they have a copy of the tax return easily available. One possibility is to vary the period according to the type of employment with either a current or usual measure. However, that approach would be more complex and take more space than is available on most census questionnaires.

471.   An approach that can be adopted is to take a current, longer period, usually a month, and ask for details about incomes from all types of employment with notes to interviewers to multiply daily earnings and divide annual earnings, in order to obtain a monthly average. That operation does of course add to the complexity and inaccuracy of the resulting estimates. Another approach is to allow flexible reporting periods and later standardize the income periods to monthly figures.

472.   It is further recommended that the income from all work activities be collected, rather than for the main activity only. In censuses that include questions about secondary activity, it is very desirable to collect data on income from both the main and the secondary activity separately, and to include both income in kind and income in cash. That is particularly so in countries where the informal sector is important and often provides a secondary activity. International experience suggests that questions on secondary activities are not well reported in household-based data collections.

3.   Income: review of questions used in national censuses

473.   Pre-coding of income ranges is often used, particularly in some developed countries. It is easier to record the answers when pre-coded ranges are given for the weekly, monthly and annual figures. It is also thought that respondents are more willing to give a range for income than they are to state the precise value. However, there is a loss of accuracy in obtaining ranges only, as the calculation of average incomes and the like would normally have to be based on the midpoint of the range. If ranges are to be used, then good base data on the distribution must be available from pre-tests, household surveys or other sources to ensure that the ranges given are suitable for the relevant income distribution. In some developing countries, ranges have been found to confuse interviewers and thus be a source of error. Good pre-testing is essential before their introduction.

474.   The sequence in example K.1 is interesting in that it allows a variety of periods in the first question and then asks about the income for that period in the second question. The income information is collected in ranges and a flash (or prompt) card is used to assist in determining the range from the respondent. The use of prompt cards is a common survey technique also being used in some censuses. The questionnaire was designed for processing using an optical mark reader. It is also interesting that a "Don't know" category was allowed on the questionnaire; this is not always a good practice, as interviewers may overuse it, or respondents when uncertain may prefer to use that option instead of selecting the category most likely to be correct. Question K.1.1 is definitely worded for paid employees, and interviewers would have to rephrase it for the self-employed. There is also no allowance for secondary employment, and it is uncertain whether total income from all jobs was to be recorded or only income from the main job. This should be explained in the interviewers' manual and in training. Even if secondary activities are rare, some guidance is necessary for when they occur.

475.   The question in example K.2 differentiates between the main job on which other economic characteristics have been collected and any others that the individual might have had in a specified month. The question also requests gross income, which not all respondents might know, even if they were reporting on their own incomes. In general, unless it is common practice for employees to receive a printed account of their salaries and corresponding deductions, it should not be assumed that individuals will know their gross income.

476.   In example K.3 (which is intended for self-enumeration), the questions cover both sources of income and total gross income by range. It asks the respondents to think about all the possible sources of income, but does not seek to collect data about each. In some developed countries, it has been common to ask for the detailed income from each source. Questions with pre-coded ranges take space for the ranges and, as seen under K.3.3, explanatory notes can at times be extensive. The exclusion of visitors in K.3.1 is interesting. This could also be done during processing provided that visitors were identified elsewhere on the questionnaire. The aim presumably was to provide income data only for usual residents and to sum to household totals without visitors. That approach may not be appropriate for censuses in all countries.

477.   No mention is made of income in kind, such as food and clothing, in example K.4. Income in kind may have been excluded and was presumably not stated since most respondents do not think of it as income. However, the fact that this type of income

Example K.1

K.1.1	**What was …'s last pay/income period?**
	1. Weekly
	2. Fortnightly
	3. Quarterly
	4. Annually
	5. Other (Please specify) _____
	6. None
	7. Not stated
K.1.2	**What was …'s gross pay/income during the last pay period, that is, before income tax or other deductions?** *(Present flash card)*
	Interview: For self-employed persons obtain 'net income', i.e., receipts less business expenses.

Income group
                0 10 20 30 40 50 60 70 80 90
                0 1 2 3 4 5 6 7 8 9

☐ Don't know

Example K.2

Qual foi o seu rendimento bruto do mês de julho de 2000?			
K.2.1 No trabalho principal?			
	0	Não tem	
	1	Somente em benefícios	
	R$		,00
K.2.2 Nos demais trabalhos?			
(Quando só tiver um único trabalho,	0	Não tem	
registrar "Não tem" nos demais trabalhos)	1	Somente em benefícios	
	R$		,00

What was your gross income for the month of July 2000?			
K.2.1 From your main job?			
	0	None	
	1	Only benefits (non-monetary)	
	R$		.00
K.2.2 From your other jobs?			
	0	None	
(If you have only one job, mark the box "None")	1	Only benefits (non-monetary)	
	R$		.00

is to be excluded should still be noted in the manuals and during training for the rare cases when interviewers are asked about it. The prompts to be used on the questionnaire and the exact questionnaire wording need to be established through pre-tests. As with other questions, every effort should be made to avoid technical jargon and concepts in the actual questions.

478.   To include an income question when the questionnaire has a landscape format can be difficult owing to the space needed for the essential notes to interviewers. However, the recording space for answers is usually adequate. No country example of an income question using the landscape format was located.

4.   Income: suggestions for enhancing the effectiveness of the questions

479.   *Principles and Recommendations, Revision 2*, recommends household income measures as well as individual measures. Household cash income is a major general welfare measure and, in countries without separate data collection programmes on income distribution, data on it are needed for studies on poverty as well as on other welfare-related issues. Depending on the question asked at the individual level, the household estimates may be derived by summing the individual measures or, as suggested earlier, a separate, broader question may be asked at the household level, if there is an opportunity to add in other income that is not included in the individual measures.

480.   Example K.5 is based on a question drawn from a developed country's questionnaire. It relates to the measurement of total person income and not just income from employment. The answer is recorded in a grid that shows ranges based on known income distributions in the country and indicates the period over which the income was received (weekly, fortnightly, monthly, yearly).

481.   The decision to include an income question should not be made lightly. Pre-tests on the topic should be extensive, particularly when the topic is asked for the first time in

Example K.3

K.3.1 Are you a visitor to the household?

☐ Yes *(do not complete the remaining questions)*

☐ No *(go to next question)*

K.3.2 Which kinds of income do you receive?

(Tick all the boxes that apply)

☐ No source of income

☐ Salary/wages

☐ Income or profit from self-employment

☐ Pension from a former employer or spouse's former employer

☐ Child benefit

☐ Disability benefits such as:

  ◆ disability living allowance

  ◆ disability working allowance

  ◆ attendance allowance

  ◆ severe disablement allowance

  ◆ invalid care allowance

☐ Other state benefits such as:

  ◆ retirement pension

  ◆ widow's pension

  ◆ family credit

  ◆ unemployment benefit

  ◆ income support

  ◆ sickness/invalidity benefit

  ◆ maternity allowance

☐ Interest or annuity from savings or investments

☐ Other sources of regular income such as:

  ◆ student grant or scholarship

  ◆ rent from property

  ◆ maintenance

☐ Other sources not already mentioned

K.3.3 What is the total gross income (before any deductions for Income Tax and National Insurance contributions) that you usually receive from all sources?

■ *Count all income including:*

  ◆ *salary or wages*

  ◆ *income or profit from self-employment*

  ◆ *occupational pensions*

  ◆ *child benefit*

  ◆ *disability benefit*

  ◆ *state retirement pension*

  ◆ *family credit*

  ◆ *unemployment benefit*

  ◆ *income support*

  ◆ *sickness/invalidity benefit*

  ◆ *maternity allowances*

  ◆ *interest or annuity from savings or investments*

  ◆ *student grant/loan*

  ◆ *rent from property*

■ *Do not deduct:*

  ◆ *taxes*

  ◆ *national insurance contributions*

  ◆ *superannuation payments*

  ◆ *health insurance payments*

Per week	**or**	**Per year**
☐ Nil		☐ Nil
☐ Less than 40		☐ Less than 2,000
☐ 40-79		☐ 2,000-3,999
☐ 80-119		☐ 4,000-5,999
☐ 120-159		☐ 6,000-7,999
☐ 160-199		☐ 8,000-9,999

Example K.4

K.4.1 Ingresos por trabajo

¿Cuánto gana o recibe (Nombre) por su(s) trabajo(s) o actividad(es)?

Anote en pesos la cantidad recibida y el período

Si no recibe ingresos anote "0" en Período

Período: A la semana 1
A la quincena 2
Al mes 3
Al año 4

☐☐☐☐☐☐   ☐☐

Anote con número   Período

K.4.1 Income from work

How much did (Name) earn or receive for his/her work (activity/activities)?

Record in pesos the amount received and the period

If received no income record '0' for the period

Period: A week 1
A fortnight 2
A month 3
A year 4

☐☐☐☐☐☐   ☐☐

Record the amount   Period

Example K.5

Hypothetical question, recommended for testing and use in countries

K.5.1 **What is the total of all wages/salaries, government benefits, pensions, allowances and other income the person usually received?** *MARK ONE BOX ONLY* *Do not deduct: tax, superannuation contributions, health insurance, amounts salary sacrificed, or any other automatic deductions* *Exclude income in kind, such as free or subsidized housing, clothing, transport* *Include the following:* **Pensions/allowances** ▪ Family tax benefits ▪ Parenting payment ▪ Unemployment benefits ▪ Newstart allowance ▪ Rent assistance ▪ Pensions ▪ Student allowances ▪ Maintenance (child support) ▪ Workers' compensation ▪ Any other pensions/ allowances **Other income** ▪ Interest ▪ Dividends ▪ Rents (exclude expenses of operation) ▪ Business/farm income (exclude expenses of operation) ▪ Income from superannuation ▪ Remittances ▪ Any other income **Wages/salaries** ▪ Regular overtime ▪ Commissions and bonuses **Information from this question provides an indication of living standards in different areas.**	☐ $2,000 or more per week ($104,000 or more per year) ☐ $1,600-$1,999 per week ($83,200-$103,999 per year) ☐ $1,300-$1,599 per week ($67,600-$83,199 per year) ☐ $1,000-$1,299 per week ($52,000-$67,599 per year) ☐ $800-$999 per week ($41,600-$51,999 per year) ☐ $600-$799 per week ($31,200-$41,599 per year) ☐ $400-$599 per week ($20,800-$31,199 per year) ☐ $250-$399 per week ($13,000-$20,799 per year) ☐ $150-$249 per week ($7,800-$12,999 per year) ☐ $1-$149 per week ($1-$7,799 per year) ☐ Nil income ☐ Negative income

a census. If possible, pre-tests should be followed by extensive re-interviews using household survey methodology to establish the level of accuracy of the census measures and thus to advise users. Early pre-tests should establish the prompts necessary for interviewers and identify problems with agricultural and informal sector income in particular. As interviewers are required in some cases to make calculations, a check on their accuracy is also essential. Users of the resulting statistics should be consulted during testing to ensure the question(s) are eliciting the type of information that they require and with a degree of accuracy that they regard as acceptable.

482. An example of a much more detailed question to collect individual incomes by source is question 52 of the Canada 2006 questionnaire (see annex I, sect D). That example collects exact values, but some countries may prefer to collect the data in income ranges (see also paragraph 473).

5. Income: issues of measurement

483.   Even when income is restricted to cash only, the income information in a census is usually very approximate. Regular wage earners may know their regularly received pay, but the self-employed and particularly the small informal sector operators are very unlikely to be able to give good estimates of their income. The same also applies to agricultural income, which is often seasonal. Measurement of income received for casual, temporary and intermittent work will also very often be approximate. A major problem that is often encountered is that one member of the household gives the census replies for all household members. That person does not often know other household members' income and only very rough estimates, if any, may be obtainable. Owing to problems of this nature with accuracy, some countries have restricted their income measures to paid employees only, even in labour force surveys. Others have excluded agricultural income and measure that component only in specialized agricultural surveys.

Part Four

Data processing, tabulation, analysis and dissemination

Chapter IX

Data processing, tabulation, analysis and dissemination

A. Introduction

484. Data processing relates to the activities involved in converting census responses into a computer-readable data file of records that have data that are as free of errors as possible and that are ready for tabulation and further analysis.

485. Unless advanced technologies are used for data collection (for example online Internet completion and return of electronic questionnaires, automated telephone interviewing, hand-held devices, and so forth), most countries will have paper questionnaires that must be converted by key entry and/or optical recognition into computer data files for analysis and tabulation. The data entry process also applies to other census topics and is covered in *Principles and Recommendations for Population and Housing Censuses, Revision 2*. Consequently, those topics are not mentioned further in the present chapter.

486. When pre-coded responses are not given for a particular question, clerical, computer-assisted or automated coding may also be required. That issue relates especially to occupation and industry and is covered in chapters X and XI.

487. The classifications used in each census should, to the extent possible, be comparable with previous censuses in order to facilitate continuity and comparability over time. Similar considerations should apply for comparison with other sources of data.

488. Owing to the size and complexity of census operations, it is likely that errors of one kind or another may arise at any stage of the census. The present chapter outlines specific points that will help to identify and correct as many of those errors as possible (data editing[1] and amendment), the derivation of selected data from census responses (for example, whether a person is employed in the informal sector), and the output phase (tabulation and dissemination of census results).

489. Every national census organization should establish a system of quality assurance and improvement as an integral part of its census programme. The data processing to detect and correct errors is only part of that programme. There is no single standard quality control and improvement system that can be applied to all censuses or even to all steps within a census. Census designers and administrators must keep in mind that no matter how much effort is expended, complete coverage and accuracy in the census data are unattainable goals. However, efforts to first detect and then to control errors should be at a level that is sufficient to produce data of adequate quality for the main uses, while offering the opportunity to significantly reduce costs and at the same time improving the quality of the census within the constraints of the budget and time allotted.

490. Automated data capture, repair and coding systems (see sect. G below) increase data quality greatly and also introduce a different set of risks compared with traditional census processing approaches. If not properly monitored and managed, data quality problems can remain undetected until late in the process when cost and timing constraints limit the options for any corrective activity.

[1] *Principles and Recommendations, Revision 2*, para. 1.311, refers to critical edits that have the potential to block further processing as well as to non-critical edits that do not have that potential. Some responses are impossible (outside valid ranges or totally inconsistent with other answers, such as data recorded on economic activity when the person is below the minimum age). Some statisticians also refer to fatal edits (impossible responses) as opposed to query edits (answers that need to be checked or confirmed, but which may be possible but unlikely).

491.   Some methods of measuring data quality from data capture processes, such as substitution rates or measures of key entry errors, are inadequate, as those forms of monitoring simply measure the overall incidence of errors but not the significance of the errors. Indeed that approach could lead to considerable extra expenditure for the correction of trivial errors that lead to no appreciable gain in quality. For that reason, data quality should be measured at the topic response level rather than at the individual character or numeral level. It should be done in two ways: independently processing a sample of records using manual processes and comparing the results for each of the records with those obtained through the automated systems; and in aggregate by comparing the overall data for an area with the expected results based on other information for that area (for example, from the previous census or other data sources). Where manual or computer-assisted processes are used, it is useful to establish a quality control system in which a sample of records is coded independently by a second operator. Any discrepancies in the results can then be resolved by an expert or experienced operator. That will allow calculation of estimates of error rates as well as the identification of problem areas and cases where re-training may be necessary.

492.   The above-mentioned process should be undertaken continuously during processing with a focus on early detection of quality problems and an understanding of any systems or processes that have contributed to them. The amount of error that is acceptable and the degree of intervention and systems or process change undertaken will depend on the assessment by the census agency of the overall fitness of purpose of the output and the overall cost and timeliness impacts. This will vary from topic to topic.

493.   In chapter II (concerning planning and design for a population census), the need to involve data processing staff at an early stage of census planning was already mentioned, as well as the importance of appropriate form design (to facilitate data entry and with sufficient space for codes to be entered). Chapter II also mentioned the possibility of processing a sample to produce advance results. The present chapter covers more detailed aspects of data processing for particular topics and the tabulation of census results.

B.   Data processing aspects for particular topics

1.   Economic activity status

1.1.   Current and usual economic activity status

494.   The usual activity reference period and the current activity reference period do not necessarily overlap, for example, if the period for usual activity covers to the end of the previous month and the period for current activity is defined as the previous seven days from the enumeration day. In the gap between the two reference periods, the person may have changed activity to become, for example, employed, unemployed or retired. More generally, a person's economic activity status during the longer reference period for usual status may be different from that person's current status over the shorter period used for determining their current status. Thus, for those countries collecting both current and usual economic activity status, it is not recommended to force any consistency between those data items during data processing.

495.   In addition, most of the not economically active categories can change status from the usual measure to the current measure and consistency in that instance should not be forced either. For example, a person could be a student during all months of the usual measure but be engaged solely in housework or have started work in the seven days of the current measure.

496.   If retired or aged is given as a category for the not economically active, it is strongly recommended that a lower age limit be put on the category (such as 45 years of age) for

the editing procedures. The limit should not be set too high as some people are retiring or being retrenched from wage employment earlier in recent years. It would be particularly wise to remove persons in the younger age groups from the category, as their inclusion may lead to incongruous results, such as having persons "aged or retired" that are 20 years old.

497. Cross-tabulations of usual by current activity are a useful method to check for large numbers of unusual cases before publication. It may be necessary to go to small area data to find explanations or possible errors of interpretation. For example, a large new factory or a natural disaster may explain odd-looking differences between usual and current activities in particular areas. Such cases are worth explaining in the presentation of the results.

1.2. Employment

498. The determination of whether a person is currently employed may depend on the answer to more than one question (especially if there is a separate question to identify those who are temporarily absent).

499. The determination of whether a person should be classified as employed or not is almost always through a derivation process. All possible combinations of answers to the questions on employment should be allowed for. The derivation should be automatic (by computer program) and will also check that the proper sequence of responses has been followed. Inconsistencies between responses to different questions (including questions being answered when they should be skipped over) must be resolved.

500. It is useful to do an age distribution of the categories of the employed. Extreme occurrences, such as employed persons over 80 years of age, and any unusual values at unusual ages in general should be investigated.

501. Tabulations of the characteristics (age, occupation and so forth) of the temporarily absent should be done as part of evaluation of this category.

1.3. Unemployment

502. Given the wide variations in the reference periods used, especially when two criteria are used in the questions to derive the unemployed, and in the way either one or two of the criteria may be applied, information about the definition and the question(s) used needs to be provided as part of the results. In addition, when questions on both "seeking" and "available for work" are used it would be useful and would facilitate comparisons across countries if some of the tables showed the two components, as well as the resulting estimates when they are combined to identify the "unemployed".

2. Status in employment

503. It may be necessary to derive the standard status in employment categories from some of the responses given. This is relatively simple with modern tabulation software, and most statisticians and data processing staff are now accustomed to much more complex derivations from survey data.

504. There are no viable consistency checks for the International Classification of Status in Employment categories using age, sex or education level since the categories can be found among employed persons of all ages, educational levels and sex. Some checks (for example, employers aged under 18 years) may demand investigation but may still be correct.

3. Occupation

505. As stated in chapter XI, it is strongly recommended that countries code their occupational data in a way that makes it possible to produce statistics in accordance with the International Standard Classification of Occupations. ILO has provided consider-

able assistance in the use of occupational classifications, with the major documentation being contained in chapters X and XI below; see also Hoffmann and others (1995) and Hussmanns, Mehran and Verma (1990).

506.   Data consistency edits can be made between occupation groups, industry groups and age and education levels: for example, most persons classified to a "professionals" category will have educational attainment which is higher than the "no education" or "primary school" level.[2] Care should be taken in the use of such edits, however, so as to avoid the risk of biasing the relationship between such variables as the level of educational attainment and occupation (for example, mismatches between occupations and education levels to study inadequate employment situations). Analysts may be particularly interested in identifying the number of persons who do not have the level of education normally expected for particular occupational groups. Checks are also possible between occupation and sector of employment: for example, "government tax and excise officials" should be in the government sector. Note, however, that in the government sector there will be a large number of jobs given occupation codes other than that for "government administrative worker". Occupational information should be coded at the highest level of detail made possible by the recorded responses (see chap. X for further details).

507.   As previously indicated, any unusual or implausible combinations should be investigated.

4.   Place of work

508.   Checks of the consistency between the type of place of work and other selected variables (such as occupation and institutional sector) may identify possible errors to be investigated (for example, doctors or public sector employees operating from a market stall).

509.   Similarly, the geographical location of place of work would normally be within reasonable commuting distance of the place of residence, and large discrepancies would need to be explained when disseminating results. An exception concerns commuting workers who work far from home and come home only for the weekend. In that case, the use of the "usual" concept to define place of residence may not be appropriate for some analyses.

510.   In respect of tabulations, the main categories of type of place of work are useful information in themselves for labour market analysts. The cross-classification of those location categories by status in employment and industry and occupation will give very valuable information on the structure of employment. That will be the case particularly for the informal sector, and also for the identification of other categories that often are of special interest, such as "paid employee home-workers". Those target groups should be clearly identified at an early stage of census preparations, and the processing steps necessary to extract information about them should be tested to ensure that the target groups are being correctly identified.

5.   Industry

511.   The coding of industry in population censuses and surveys is covered in some detail in chapters X and XI. There are useful suggestions also in Hussmanns, Mehran and Verma (1990) as well as in other ILO papers and documents. Although industry coding is an established topic, there may be problems with industry data collected in censuses, and the points presented below are particularly noteworthy.

512.   The industry description and code should refer to the establishment where the person actually works and not to the legal unit to which the establishment may belong. For example, a major company in a country, "XYZ Ltd", may be well known but may engage in a variety of activities at different establishments around the country. Generally,

2 Designers should be aware that there is a tension between occupation and educational attainment, and that checks should not be overdone.

industry should be coded for each separate establishment (usually at separate locations), not only for the main activity of XYZ Ltd. When there is access to a reasonably up-to-date register of establishments, and where industry is coded for each unit, the industry question should aim to collect the detailed name and address of the person's place of work (see para. 338). Then coding can be done by reference to the register and will be compatible with employer surveys based on the same register. The issue is explored further in chapter XI (see paras. 680-681).

513.   Unfortunately, few countries have a comprehensive updated register of employers with separate identification of all establishment locations. Thus, in most countries the detailed name and address of the establishment can be used as a source for the correct industry code for only some of the responses, and a description of the activity at the establishment has to be used as the basis for assigning codes for the rest.

514.   However, even if a register of establishments (or of employers) is very poor, it is often useful to have a limited list of major companies and enterprises with their industry codes when creating the coding indices described in chapter X. For example, XYZ Ltd may be dominantly an iron and steel mill but may also have XYZ Coal Mines and XYZ Iron Ore Mines as separate establishments. Entries in the coding index for industry reflecting that structure will help to ensure that respondents working at the separate establishments are given their relevant industry codes and not simply coded to the code given to XYZ Ltd. A similar listing of major government departments and organizations is also very useful in coding respondents working for the government to their correct industry codes: for example, to code those working with the country's health administration separately from those working in government hospitals.

515.   There are a limited number of consistency checks of industry codes with the results from other questions. They include checks that jobs in establishments coded to industry classes related to public administration are in the government sector (but note that many government sector workers will work in other industry groups), and that only certain occupations are associated with the industry category for "activities of households as employers of domestic personnel". Other edits are generally not definite and depend on the situation in a given country. Particular attention should be paid to any codes identifying separately "production for home consumption", including subsistence agriculture.

516.   Informal sector derivations should be carefully checked to ensure that only relevant industries are shown. Some limited checks are possible of industry against status in employment: for example, self-employment cannot occur for "activities of households as employers of domestic personnel", or for "industry classes related to public administration". A full range of edits should be prepared before data processing commences, but cross-tabulations with other topics should be done early and progressively to locate improbable occurrences that will need explanation or possibly further edits. It is recommended that this not be left to the final tabulation phase since the resolution of such queries can delay the results considerably.

6.   Institutional sector

517.   If the question on institutional sector is combined with the status in employment question(s) as presented in some of the examples shown previously, it will not be necessary to do a consistency check between status in employment and institutional sector (for example, government sector workers being classified as self-employed or unpaid helpers). However, in cases where both status in employment and institutional sector data are collected separately, it is important that they be cross-checked. There are no possible checks against sex, age or education level. Checks against the industry and occupation classifications are also possible as mentioned above.

7.   Informal sector

518.   The design of a procedure to measure employment in the informal sector should be clearly thought out and well tested. Some computer derivation may be necessary and will need very careful consideration and prior testing.

519.   Careful checks and cross-tabulations of status in employment categories by industry and occupation groups will be essential to edit any informal sector data, and re-coding is likely to be necessary to remove absurdities.

520.   If the derivation procedure is properly designed, then consistency checks with the institutional sector will not be needed; otherwise consistency checks will be required to ensure that there is no public sector or company employment in the informal sector.

521.   Such preparations will prove to have been well worth the extra effort when the demands for small area and regional estimates of informal sector employment are received.

8.   Informal employment

522.   The procedure to identify those in informal employment should also be carefully designed. There are no viable consistency checks since those in informal employment may be in any occupation, industry or institutional sector. If information is collected to derive whether a person is in the informal sector, then it would be useful to check that all who are engaged in the informal sector are also classified as informally employed.

9.   Working time

523.   If a person is recorded as temporarily absent from his or her main job then the actual hours recorded for that main job should be zero. If a person has more than one job reported then working time should be calculated for all jobs for both hours actually worked and usual hours, if the data was so collected. If hours actually worked are recorded separately for all jobs then it would be reasonable to expect that usual hours would also be recorded separately for all jobs.

524.   Consideration should be given to compiling statistics of average hours actually worked as well as distributions of the employed in five-hour groups of hours actually worked. In presenting the distributions of employment by working time, care should be taken to ensure that the distribution shows both low hours and excessive hours.

10.   Income

525.   When income is collected for the same reference period as the measurement of economic activity status, then only persons classified as employed should have income from employment. Unemployed and inactive persons may receive other types of income, but not income from employment. Persons with secondary paid or self-employment jobs should have income from the second activity as well.

526.   In cases where the reference period for the measurement of income differs from the reference period for the measurement of economic activity status, then no viable consistency checks are suggested.

527.   Total household income should always be equal to or more than the sum of the individual job-related incomes in the household. Codes for "not stated" must be allowed for this question in particular, with a clearly defined strategy on how to deal with such cases in the processing of the question, for example, by imputation or by exclusion from averages.

528.   Extremely low and high incomes should be checked since they can distort averages. Those checks can be done by status in employment and by occupation.

C.   Data quality and verification

529.   Before results are disseminated, it is important to ensure that economic activity estimates are checked and confirmed. That requires an evaluation of the data quality and comparisons with other data sources, where available. As previously stated, such evaluation and verification will be facilitated if concepts, definitions, reference periods, scope (including age cut-off) and the like are comparable between sources. Post-enumeration surveys are also an important part of the evaluation (see also *Principles and Recommendations, Revision 2*, paras. 1.274-1.277 and 1.379-1.400.) Quality assessment techniques or strategies include checking coefficients of variation, the involvement of subject-matter experts in the evaluation of data and the production of test tables at various geographical levels to identify where additional checks may be required. Such preliminary tables should be prepared and evaluated early in the output-generation phase. If significant concerns about data quality are identified, it may be necessary to consider reprocessing to detect and amend unforeseen errors and rerunning the early tables.

530.   The disseminated results should include a summary of the evaluation, an assessment of the data quality and an explanation for users on the differences between estimates from different sources.

D.   Dissemination

531.   A census is not complete until the information collected is made available to users in a form, and to a timetable, suited to their needs. Thus, in disseminating the results of the census much emphasis should be put on responsiveness to users and on high standards of quality in the production of statistics. A major strength of the census is that it supplies statistics for small areas and small population groups, and the tabulation and dissemination systems should make full use of that strength by generating detailed tables of results.

532.   Annex II of the *Handbook* lists the tables relating to economic characteristics that are recommended in annex II of *Principles and Recommendations for Population and Housing Censuses, Revision 2*. That annex also includes other recommended or additional tables of interest, for example on the economic activity of persons with disabilities and on the economic activity of the foreign-born population.

533.   The tables in annex II of the *Handbook* have generally been recommended for production for national, regional and provincial levels with an urban/rural dissection, and even for production at the locality level. Some have been prepared in considerable detail (for example, those cross-classified by occupation and industry), and therefore it would be more appropriate to release the more detailed tables electronically (see below) or in publications and abstracts that relate to only those small geographical areas.

534.   The large size of census files can affect the timeliness of the results. It may not be necessary to process and tabulate all census results before dissemination. Rather, the data can be released in stages, so that some census results are released before others. If that is done, it is suggested that the results on economic characteristics be released at a time close to the time when other related topics, such as education, are released.

535.   As indicated in *Principles and Recommendations, Revision 2* (paras. 1.208 to 1.209), there are several ways of making the results of a census available to the user:

(*a*)   As printed reports containing standard and pre-agreed tabulations, usually at the national, regional or local district area level, that may be obtained from government agencies or directly from booksellers;

(*b*)   As unpublished reports (often referred to as abstracts) comprising standard tables but produced for either smaller geographical areas or population sub-groups not otherwise included in the published reports; they may often be requested by users, who may have to contribute towards a proportion of the marginal costs of their production;

(*c*)   As commissioned output produced from a database, comprising customized cross-tabulations of variables not otherwise available from standard reports or abstracts;

(*d*)   As microdata, usually available in restricted format only and supplied under strictly controlled conditions.

536.   A range of products and output media should be available to meet the changing requirements of users. Major users should be consulted in advance to determine their needs. Users will also need supplementary metadata covering definitions, classifications, and coverage and quality assessments.

537.   Each of the different forms of output is presented below.

1.   Printed publications

538.   Owing to production costs, printed publications may eventually play a secondary role in the dissemination of the main census results, though paper still provides a medium that does not readily deteriorate and does not require the user to have any necessary hardware, software or technical skills.

539.   It is suggested that only the basic and recommended set of tables appear in the printed publications that are prepared for broad user and community use, and that even then those tables do not need to include statistics for small geographical areas. National, regional and provincial statistics would generally suffice.

540.   Those printed publications (and output via the Internet, CD-ROM and other media) should include graphs, maps and textual analysis as appropriate.

2.   Electronic dissemination

541.   The release of some outputs (especially of detailed cross-classified results and results for small population groups, and some of the optimum set of tables) may be possible only by distribution through the use of high-capacity electronic media. However, when data are provided in electronic form, special attention should be given to providing users with easy means of data retrieval.

542.   The outputs and relevant metadata should be accessible in standard formats as well as in common database and spreadsheet format for easy retrieval and manipulation.

3.   Microdata

543.   National statistical agencies should also consider the possibility of releasing data files of unidentifiable unit records so that advanced users can undertake more detailed analysis. Such releases should be subject to careful controls on data security and confidentiality. For more information, see "Principles and guidelines for managing statistical confidentiality and microdata access" (United Nations, 2007).

4.   Confidentiality and data security

544.   The computer systems handling census data should have strict safeguards to prevent unauthorized access to the information (see also United Nations, 2007, as mentioned above). Care should be exercised to avoid the inadvertent disclosure of information (for example occupations or income) in respect of identifiable individuals through

the statistical results of the census. Special precautions may apply particularly to statistical output for small areas.

545.   Measures to ensure disclosure control may include some, or all, of the following procedures:

(*a*)   Swapping of some records between households of similar demographic characteristics located in different small geographical areas within the same higher-level geographical area;

(*b*)   Restricting the number of output categories into which a variable may be classified, such as aggregated age groups rather than single years of age;

(*c*)   Where the number of people or households in an area falls below a minimum threshold, suppressing statistical output—except, perhaps, for basic head counts—or amalgamating with that for a sufficiently large neighbouring area;

(*d*)   Randomly modifying or rounding data before the statistics are released;

(*e*)   In the case of microdata or public use samples, removing all information from databases relating to name, address and any unique characteristics that might permit the identification of individual respondents.

5.   Graphical and related output

546.   Products should be developed that allow statistical and geographical information to be delivered together with geographical information systems to meet as widespread an interest and with as much flexibility as possible, commensurate with assurances on confidentiality. When databases are provided with associated graphing and mapping capabilities, their usefulness will be greatly increased.

547.   Ideally users should, themselves, be able to generate graphs and/or maps easily, and then to print or plot them or make the images available for other users. Several census organizations have produced such a product, sometimes in cooperation with a commercial company.

548.   Printed and electronic output should include summary graphs and maps.

E.   Metadata

549.   The provision of adequate metadata is essential to allow users to understand the meaning of statistical output. A metadata system contains, in particular, definitions of terms and variables, classification schemes and an assessment of the quality of the estimates. For variables for which international standard classifications have been used, those classifications may be included in the metadata system. For variables that are not classified by international standards, the definitions, concepts and classification schemes actually used in the census should appear in the metadata system.

550.   A metadata system may provide supplementary information on characteristics of surveyed and published data. Such information is essential for users to understand the strengths and limitations of the statistics and to use the data properly in more detailed analysis.

551.   The population census for the 2010 round of censuses has to ensure comparability with the data from the previous population censuses, to the extent possible. Nevertheless, it should include new elements relevant to development that has taken place during the time since the previous census. Thus, the metadata for the population census around 2010 should follow the metadata system from the previous population census with an update in line with the needs resulting from development since the previous population

census. With the rapid changes in information technology that are taking place, the design of the newer metadata system may also need to change.

552.   When disseminating the results of the census, the associated metadata should also be disseminated. This disseminated metadata should include an evaluation of the quality of the census results, including at topic level, and an evaluation of its comparability with other sources of statistics on economic activity.

F.   Emerging technologies for data processing and output

553.   Technology has been used to assist in all phases of population censuses for many years, and particularly in data processing, analysis and dissemination. Well-established approaches and technologies might well be the most viable option for many countries, and adoption of a new technology or approach should be considered only where there is a sound understanding of the new approaches and technologies and the developments can be managed. There should also be a clear understanding of both the risks and the benefits.

554.   The following sections cover only some of the current technologies. For more information, see *Principles and Recommendations for Population and Housing Censuses, Revision 2* (United Nations, 2008) and *Handbook on Census Management for Population and Housing Censuses* (United Nations, 2001). Further technological development is certain and will affect future census taking.

G.   Data capture and coding methods

555.   Several data capture technologies have traditionally been used, such as key entry and optical mark recognition (OMR). The last decade has seen significant improvements in intelligent character recognition (ICR), data repair, imaging and automated coding technologies that have reduced the cost of census processing and the time it takes to do it. They have also improved data quality. Those trends are likely to continue.

556.   For most countries, the most cost-effective option is likely to be a combination of digital imaging, intelligent character recognition, repair and automated coding whereby the census forms are first processed through scanners to produce an image. Recognition software is used to identify tick box responses and to translate handwritten responses into textual values. Confidence levels are set to determine which responses require further repair or validation. Automated repair typically involves the use of dictionary look-up tables and contextual editing. Operator repair can be undertaken on images not recognized. This is cost-effective only for those questions where there is a high probability that the repaired data can then be automatically coded.

557.   Automatic coding uses computerized algorithms to match captured responses against indices. The responses that cannot be matched are then passed to a computer-assisted coding process. That methodology eliminates inconsistencies caused by human error inherent in manual and computer-assisted coding. The quality of the automated capture and coding needs to be carefully monitored during processing to ensure that the system is functioning as specified. However, the process does offer the opportunity to improve the quality of the data and has the potential to reduce processing costs significantly.

558.   In computer-assisted coding, a coding clerk typically keys a truncation of selected words from responses written on the form or captured in some other way. Matching index entries are displayed on the computer screen and, when the index entry most closely matching the response is selected, the appropriate classification code may be written automatically to a data file or database.

559. Census agencies also need to consider how the data are going to be held through the processing stream. Traditionally, census processing has been conducted using a flat file that gets progressively updated, with the earlier version of the file retained for backup and recovery processes. Typically that procedure has been allied with batch processing where a discrete group of forms (typically for an enumeration area) are processed together. Thus, data from the forms will be entered, edited and coded as a group, allowing a high degree of workload control. Databases allow information to be held and processed at the individual field level. That provides a greater degree of flexibility: once census data are electronically captured, it is possible to maximize both processing efficiency and quality, as similar responses can be readily grouped and coded together. However, holding the census data in a database requires more complex systems to manage and deliver work. Consideration also needs to be given to backup and recovery mechanisms.

560. Systems such as those described above typically require far more extensive development and testing than a traditional census processing system. A number of factors need to be considered throughout the design process and integrated into the development of the system, such as the work organization of the remaining clerical processes.

H. Output

561. Traditionally, census output has been conceived in terms of generating tabulations, usually for sophisticated clients well acquainted with the census data, its structure and content, and other metadata. Less sophisticated users have traditionally relied on static products, such as publications, that generally contain a limited range of outputs.

562. Internet dissemination allows for much improved service to users by permitting the design of products to meet the needs of different types of census data users, from the novice to the sophisticated; the cost-effective dissemination of a much wider range of census data; and improved usability of the census data.

563. The main strength of a census in a developed statistical system is to complement the information provided by other data collection methods such as surveys with a focus on small domain statistics, that is, information for small geographical areas and for small population groups (both social and economic). Internet dissemination can support both uses of the data. For small geographical areas, Geographic Information System (GIS) technology can be used as a means for defining areas of interest in searching for data as well as for mapping of the outputs of the search. A range of packages can be used to focus in on populations of interest from large predefined matrix tables.

564. The Internet dissemination system should provide freedom for clients to specify the form of the output, whether as hard copy or as a data file that can be exported into a range of commonly available statistical analysis, tabulation or mapping packages.

565. Some countries may wish to consider providing access to clients to submit tabulations directly online to be run against the census unit record file. Again, protecting the confidentiality of the census data is a prime consideration in such systems. In addition to implementing confidentiality procedures (such as random rounding), there may be a requirement to limit the size of tabulations that can be submitted through this method.

Part Five

Coding of occupation and industry

Chapter X

Preparations for coding occupation and industry

566. Chapter X explores the various strategies and preparations that are necessary to ensure effective and reliable coding of industry and occupation responses, and also considers the costs and advantages of the different alternatives. It outlines the objectives and main strategic choices to be made and presents the main organizational factors determining the effectiveness and success of the coding and processing tasks. The chapter describes the development of the major tools, the coding indices and how to use them effectively.

567. The basic tools include such international classifications as the International Standard Classification of Occupations (ISCO)[1] and the International Standard Industrial Classification (ISIC),[2] which are available on the Internet. Since requirements will vary from country to country, the chapter covers common issues that must be addressed in any census operation when questions related to occupation and industry are included in the census questionnaire.

568. At the time of release of the present manual, the most recent versions of the classifications were ISCO-08 (forthcoming) and ISIC, Revision 4 (United Nations, 2008a), endorsed by the United Nations Statistical Commission at its thirty-seventh session in 2006. In general, it is preferable to adapt the international classifications for national or regional use rather than to use them directly in the collection of national statistics. In many countries such adaptations may already exist (or are under development) for use in economic and labour statistics. Wherever possible such adaptations should also be used for the census. In their economic and labour statistics, some countries may still be using national classifications based on earlier versions of the international standards (for example, ISCO-88 or ISIC, Revision 3.1) and will need to make a decision whether to update their national classifications in time for use in the 2010 round of censuses.

A. Objectives

569. The main aim of the coding and processing of the information given by the respondents about their place and type of work (about the industry and occupation in which they work) is to determine and record correctly to which of the categories in the respective classifications the jobs belong, at the most detailed level of the classification possible on the basis of the information provided in the response. That task has to be completed within an overall processing plan for the census, to a pre-specified timetable and either within pre-specified cost limits or in a fashion that will minimize cost, given the specified data requirements.

570. The development of a census-processing strategy with the preceding aims needs to consider many aspects and requirements of the processing task, including the following:

　　(*a*)　The existing data processing capacity and infrastructure;

　　(*b*)　The type and format of the information to be processed;

(*c*)   The volume of data to be processed and the throughput rates required;

(*d*)   How the processing of the information about industry and occupation is embedded within the total data processing task for the census;

(*e*)   The level of detail required to satisfy important user needs in the national context, as well as for international reporting.

571.   The precise impacts of those aspects will depend on the choices made with respect to some strategic aspects of the coding and processing of census forms, as explained in the following section.

B.   Strategic coding and processing options

1.   Process all cases or a sample only

572.   Coding of industry and occupation is typically one of the most expensive and time-consuming operations in the processing of a census. To reduce costs, to make the management and quality control of coding easier and to enable earlier production of results, consideration should be given to obtaining and/or coding the information for a sub-sample only.

573.   A policy to code a sub-sample only may be implemented at the data collection stage by fielding longer and shorter versions of the census form, so that only a sample of the population answers the questions designed to capture the information needed to code industry and occupation. To avoid error-prone field sampling procedures and administration of different forms or schedules, one may alternatively collect the information from everyone but process it for a sample only. Either form of sampling will significantly, almost proportionately, reduce the operational costs of coding but not the costs of preparing for the coding operation. However, sampling means the introduction of sampling-related errors (imprecision in the resulting estimates), and that will be an important concern when estimates are to be produced for small population groups or small geographical areas. As the provision of small area and small group data on a nationally consistent basis is seen as one of the major functions of the census in most countries, many users may see the processing of industry and occupational information for a sample as contrary to the objective of the census, unless the sample size is sufficient to meet those requirements. That perception may be especially prevalent if there is a regularly conducted, annual or quarterly labour force survey that already provides sample-based national statistics for employment and unemployment, classified by occupation and industry.

574.   Other considerations relevant to the coding of a sample or sub-sample are the following:

(*a*)   The use of a sample for the coding of industry and occupation will mean that sampling imprecision will make comparisons between categories and over time more difficult;

(*b*)   Sampling may involve a serious loss of descriptive power, particularly for small areas, small industries and small occupational groups;

(*c*)   It may hinder the production of summary socio-economic indices for small areas or groups, as they are based on the occupation variable.

575.   If the decision is made to collect industry and occupation information for only a sample of the population, then attention must be given to the design of the sample, balancing the requirements of statistical precision against those of operational simplicity and robustness. The sampling fraction should in principle be determined by balancing the precision of estimates required for the smallest aggregates of the population for which

separate figures are to be produced against the saving of cost and time resulting from reducing the processing load. For operational convenience, sampling has often been done by selecting all households in a whole enumeration area to be included in the sample, thus sending whole bundles of questionnaires into one processing stream or the other. Such an approach may, of course, result in very high clustering of particular occupations and industries, and will give valid statistics only at high geographical levels, thus contributing to further loss of precision.

576.   If the sampling within enumeration areas is done manually it is also important that the sampling procedure be simple (for example, by selecting every nth case, with $n = 5$ for a 20 per cent sample, 10 for a 10 per cent sample and 20 for a 5 per cent sample, as appropriate). The sample selection should be done on a probability basis. In practice, it can be conveniently approximated by taking every nth census household in each data collection area (enumeration district) and including all members of those households in the processing sample. In that case, owing to variable household size and to the clustering of individuals within households, there will be a departure from a true simple random sample of individuals (the units of analysis for household-based occupational statistics). An effort should be made to estimate the sample design effect, but guidance on how it should be done is beyond the scope of the present publication.[3]

2.   Field or office coding?

577.   With respect to census coding, the following choices are available:

(*a*)   The respondent codes himself/herself to a predefined category;

(*b*)   The enumerator codes in the field, either during the interview or before the questionnaire is forwarded for further processing;

(*c*)   Specially trained coders code in connection with consistency checks of the questionnaire and data entry.

578.   The choice is a strategic one with respect to the balance between costs, quality of coding and resulting statistics, and control of the coding process.

2.1.   Coding by the respondent

579.   In practice, coding by the respondent means that respondents are requested to place their job in one of a set of predefined categories presented to them in written form on a questionnaire, or on a card read or handed over by the enumerator.

580.   The main advantage of this approach is that it is the least expensive of the possible coding procedures. It can easily be adapted for use with optical mark recognition and intelligent character recognition technologies resulting in even greater cost savings. The main disadvantage is that it results in lower quality data, in terms of reliability, validity and detail.

581.   The lower level of reliability results from the difficulty of assuring consistency in the ways the respondents relate the predefined categories to the job about which they are asked to give information. The content of each predefined category has to be described with a limited number of words, normally in the form of a category title, and it may be very difficult to convey to respondents the intended understanding of the ways in which their jobs relate to the different categories. The fact that the number of possible categories to choose from has to be severely limited will also limit the number of possible mistakes to be made. This is of little comfort, however, when many users of industry and occupation statistics need to make much more detailed distinctions and to work with much more homogeneous categories than those obtainable when using this approach.

[3] For details on design effects, see for example Leslie Kish (1995).

582.   Nevertheless, the cost savings are such that some national statistical agencies have been asking respondents to code in their census, at the expense of one of the main advantages and purposes of the census: the possibility of providing statistics for (relatively) small groups in a consistent manner for the whole country. The quality of self-coded occupation and industry data is so restrictive that it would probably be better to augment the sample size of existing sample surveys as an alternative method for obtaining more detailed occupational data.

2.2.   Coding by the enumerators

583.   Coding by enumerators can take two forms:

(*a*)   The enumerator assigns the response to a pre-coded alternative during the interview, on the basis of the respondent's answer(s) to the standard questions; or

(*b*)   The enumerator writes down keywords of the respondent's answer(s) and then codes the response after the interview, but before the questionnaire is forwarded to the processing centre.

584.   The cost advantages of the first possibility are almost the same as for respondent coding. The main difference is that the interview is likely to take a little longer because of the need for the enumerator to understand and "translate" the information received to the appropriate category. The procedure may be used if the respondents themselves are not expected to read the questions and write the answers. The effect of the procedure on the validity and usefulness of the resulting statistics is similar to that described in section B.2.1, "Coding by the respondent". On the one hand, the reliability of the coding may improve relative to coding done by the respondent if the enumerator has been well trained and has received more detailed instructions on what types of jobs the different pre-coded categories are supposed to cover and where there are ambiguities that will require probing. On the other hand, the enumerator may misunderstand what the respondent reports and therefore not select the correct category.

585.   The advantage of the second approach, in which the enumerator writes down the response that is to be coded later, over the use of categories re-coded by respondents or enumerators, is the possibility that it offers for much more detailed coding and thus for much more valid and useful results. An enumerator coding outside the immediate interview situation may be given complete coding indices to assist with the coding process, as well as other coding aids—including the possibility of conferring with supervisors. Another advantage is that, as the enumerators gain experience with coding, they will become more aware of the type of information that is required to code correctly and improve the recording of those characteristics with interviews. The main disadvantage is that, since they are more numerous and geographically scattered, the enumerators cannot be given the same amount of training, supervision and support as that given to specialized coders, to assure coding reliability. Narrowcasting of satellite transmissions has been successfully used to provide interactive training of census field staff in remote locations and should be considered as a solution to the training of geographically scattered enumerators.

586.   A statistical consideration with the second approach is that all coders have irreducible idiosyncrasies that produce biases in the distribution of codes, which they allocate, relative to the mean of all coders. It can be shown that, if biases of similar magnitude are imparted by all coders, whether in the field or in the office, the effect on the overall variance of the results will be less if each of a large number of fieldworkers codes a relatively small number of cases than if each of a relatively small number of office coders code a large number of cases. Depending on the relative sizes of the work quotas of field and office workers, the total variance of estimates with field coding may still be lower even if fieldworkers code in a more variable (less consistent) way than office coders. Note,

however, that the above reasoning is valid mainly for national aggregates. For smaller areas the biases of individual field coders may not cancel each other out, and it may be possible to distribute the workload among central coders to prevent clustering of their biases. The best solution is, however, to provide the coders with training and tools, which can minimize systematic errors and biases.

587.   It may be possible to save time and reduce costs and operational complications by using field coding in circumstances where industry and occupation are the only items requiring office examination and coding. In that case, the field-coded question- naires could be passed straight on to computer data entry. However, fieldworkers may be thought unsuitable in terms of background and training to act as coders. In addition, field coding sacrifices to a large extent the important advantages of a controlled and supervised coding environment that can provide direct feedback on coding quality and an incentive for correct coding. Such direct feedback can be provided to office coders in a census through quality control procedures but will be much more difficult to provide to enumerators, as field operations are likely to have been finished by the time serious problems are discovered, and field supervisors will normally not be equipped to give high priority to quality control of coding results.

588.   The arguments in favour of using field coding, as compared to specialized coders detached from the data collection process, are closely linked to having a permanent field staff that can be trained cost effectively in coding and can accumulate experience. Field coding may therefore be a realistic option for continuous labour force and similar surveys, as well as for local administrative offices, but the ad hoc and much larger-scale operations like population censuses should normally include the coding of industry and occupation in the central processing operation.

2.3.   Office coding (including computer-assisted coding)

589.   Specialized coders, located in a few processing centres, may specialize entirely in the coding of one variable or in coding in general, and/or they may carry out the coding as one part of an integrated data entry, coding and data control operation for the census. The exact context will depend on the whole organization of the processing operation, but the concerns outlined below should be given serious consideration.

590.   With a large coding operation, such as for the census, the coding staff will become thoroughly accustomed to the practical routines of the task, the rates of work required and so on. Specialized coding units may become strongly production-minded and iso- lated from the rest of the organization, with little incentive to assess their product in terms of external validation criteria. It is therefore necessary to establish rigorous measures of quality control in coding operations. In addition, individual coders may find short-cut methods that reduce the laboriousness of their task but may also incorporate errors or unjustified assumptions, if not actual violations of the coding instructions.

591.   Although it may also occur when coding is done by enumerators in the field, such unofficial departures from, or additions to, the specified coding procedures can be a particular problem in a specialized office coding environment. They may become institutionalized to the point where no distinction is perceived between them and the coding rules derived logically from and designed to support the system of classification in use. Supervision and routine systematic audits of the coders must be employed to ensure compliance with the established coding rules and procedures. More generally, issues of external validity and internal consistency of coding may fall into abeyance unless the routine procedures of the unit include specific and properly designed checks on the levels of coding validity and reliability achieved.

592.   Coding may be one element in a larger processing task for each operator. This may make the coders' tasks more interesting and can be a useful method of using employee resources efficiently. An additional argument for this solution is that fewer persons will handle the census forms, and that simplifies the control task. (Note, however, that to simplify the discussion the larger processing context is mostly ignored in the rest of chapter X.)

593.   The large-scale coding exercises mounted for national censuses of population generally rely on special recruitment and training of inexperienced coding staff. The performance of staff that are inexperienced may limit or delay the acquisition of the adequate coding habits described. Nevertheless, the coding results are likely to be poor unless recruitment, training and coding procedures are well planned, well executed and supervised. A particular problem is that, owing to resource limitations or practical difficulties at a time of heavy stress during the preparations for census processing, there may be insufficient time to establish rules and routines and generally test the coding procedures before production coding has to start. If this happens the organization may quickly become overwhelmed by the sheer volume of documents and data to be handled, with a consequent loss of control and a severe drop in standards, which may prove difficult to recapture and reverse later on.

594.   In general, existing coding procedures used in regularly conducted labour force surveys should be adapted for use in the census. This avoids the need to "reinvent the wheel" and helps to ensure that coding is consistent among collections.

3.   Level of coding

595.   The purpose of the coding process is to determine and record, from the responses obtained from the respondent, the code for the category in the classification to which the job of the respondent belongs. The coding can be seen as a translation process, whereby the coder translates the industry and occupation responses into the correct categories; only in a minority of cases will the words of the responses be the same as those used to designate categories in the classifications. The raw materials for the process are the responses, and the tools the coders can use are the coding index, the coding instructions and the persons responsible for answering queries.

596.   The coding process should therefore be designed to find and record the most detailed codes supported by the responses. The more information retained after the coding, the more valid and therefore valuable the resulting statistics can be for the users. A compromise option, often overlooked, is coding at different levels within different parts of the structure. However, traditionally, the most common procedure has been to decide that coding should be done at a particular level of the classification structure, for example the three-digit level, no matter what information has been provided in the response. The arguments for this have commonly been as follows:

    (*a*)   That it would be too costly to code to a larger number of categories, both in terms of coding errors and in terms of staff hours required;

    (*b*)   That the responses would not support coding to more detailed categories;

    (*c*)   When coding only a sample, it would not be possible to publish results for the more detailed categories owing to a lack of observations.

597.   However, closer examination of those arguments in the light of the experience gained by statistical agencies has revealed the following:

    (*a*)   The marginal costs of coding to a larger number of categories in the classification, in other words, to a lower level of aggregation, are rather small in terms of increased error rate as well as in terms of work hours needed for coding. The error rate for aggregate

categories does not seem to increase. The most significant cost associated with coding to a lower level in the aggregation structure is the associated increase in the required sample size;

(*b*)   Experience clearly shows that the industry and occupation responses are very uneven in the level of detail they will provide. Many responses will support detailed coding, especially if the questions are formulated along the lines outlined in part three. At the same time, a significant number of responses will not support the level conventionally chosen. By insisting on a predefined level, the coding process may therefore both lead to an unnecessary loss of information for a large part of the returns and to a misrepresentation of the data quality for other parts;

(*c*)   The similarity criteria used to define the categories in an industry or occupation classification are mostly defined with reference to the nature of the production process and the type of work performed, without much regard for the number of employed persons in the resulting categories. Consequently, the number of jobs that can be found in categories defined at a particular level in the classification may differ greatly. Even in the few cases where statistical balance has been adopted as an important consideration when constructing the classification structure, the number of jobs in a category such as "customer service clerks" defined at an aggregate level may, for example, be smaller than that of a category, such as "shop sales assistants", defined at a lower level in the structure of the classification, but within another high-level category;

(*d*)   In addition, the tabulation of industry and occupation statistics may involve both the merging of categories and cross-classification with other variables such as age, sex and region. Even in cases where the data cannot be published, the microdata can be extremely useful for internal analysis. Consequently, one should not restrict tabulation possibilities during the coding process.

4.   Coding of vague and difficult responses

598.   Most industry and occupation classifications specify residual categories of "type x industry/occupations not elsewhere classified (n.e.c.)". "Not elsewhere classified" categories are designed to take care of activities and jobs that belong to the more aggregate category but that are not similar enough to the specified subcategories within it to belong in any of them, and which in themselves include too few cases to warrant separately specified categories. Responses should be coded to an n.e.c. category only when they match index items of that category. The not elsewhere classified categories should not be used to code responses that the coders cannot assign to any of the specified categories. Such responses can exhibit the following characteristics:

(*a*)   Be too vague and imprecise to allow the coder to determine to which category the job belongs;

(*b*)   Indicate that the establishment (or job) in question produces a combination of goods or services (or the job involves tasks and duties) that cut across the distinctions made in the industry (occupation) classification;

(*c*)   Represent a type of production or work not covered by the classification.

599.   The proper way of handling such responses will depend on the type of case, as follows:

(*a*)   Vague and imprecise responses should be coded to the level in the aggregation structure supported by the information contained in them; they should not be forced into any particular detailed category where it is likely that only a small proportion of the jobs would fall if one had an adequate response. For example, in one census, 15 per cent of the jobs coded to the major group "clerks" could not be coded to any of the more detailed categories within that major group. It would have represented a significant distortion of the

results if they had all been placed in one particular more detailed category together with the jobs that properly belonged to that category. A common method of dealing with this type of response is to provide entries in the coding index for commonly occurring vague responses. Those entries are assigned the code for the relevant higher category, followed by trailing zeros. For example, in Canada the census was coded to the level supported by the responses and then the responses were allocated proportionally to the more detailed categories in a transparent manner;

(*b*) The classification of establishments with an uncommon mix of activities or jobs with an uncommon mixture of tasks or duties should, as far as possible, be made on the basis of the general priority rules of the classification. Such responses should preferably be left to expert coders or raised as queries for the classification experts. Disruptions of the coding process can be minimized if the responses are given a special code and the questionnaires put aside for later examination by the experts. The same treatment should be given to responses that seem to represent establishments with activities or functions not covered by the industry classification and jobs with tasks or duties not covered by the occupation classification. The reporting of these types of difficult cases is an important input to the process of updating, maintaining and possibly expanding and revising the relevant classification and coding index.

600. The information to be used as a basis for the coding of industry or occupation is first and foremost that contained in the responses to the respective industry and occupation questions. However, in a significant number of cases where that information is not sufficient for the coder to choose between possible alternatives, ancillary information provided by the responses given to other questions may provide the basis for making the choice. (For occupation the most important type of such information is often that pertaining to the industry of the workplace.) However, it must be emphasized that the use of such ancillary information should be specific and strictly controlled, to avoid an undermining of the descriptive and analytical use of the variables, in particular when they are being used together or for cross-classification. This means that rules for proper use of such ancillary information—when and how to use it—must be incorporated into the coding indices and the coding instructions.

C. Planning and organizing coding operations

601. The remarks that follow are based on the assumption that coding is to be carried out by specialist coding staff in the context of the final processing of the census, since that is the solution most commonly chosen for census operations by national statistical agencies. In most cases it is easy to deduce, from the information that follows, the guidelines that should apply to field staff coding or to other decentralized arrangements.

602. Assembling the right resources to process industry and occupation information in the right places at the right time and managing those resources efficiently are fairly complicated tasks. They require anticipation and cooperation between different parts of the census organization and must be coordinated or integrated with the other processing tasks. Large volumes of documents and data have to be handled and, because of the interdependence of different stages within the overall processing plan, the penalties for failures of operational or quality controls, in terms of delays and cost increases, can be heavy. The main outline of the processing plan for a population census, particularly as it affects requirements for finance, staff, equipment and premises, may need to be worked out long before the start of the actual processing. Managerial staff who will be involved in the planning and supervision of the coding operation and professional staff responsible for designing the classification and coding procedures, training coders, updating classifications and coding indices, and interpreting results need to collaborate closely at the planning stage.

1. Finance and resources

603.   A substantial amount of money is required to support the processing of a census, and it will have to be estimated and provided for under appropriate budgets. Estimates for each part of the processing task often need to be made several years in advance and integrated into the financial planning and procurement procedures of the responsible agency to ensure adequate provision. Such budget considerations demand early decisions about resource requirements, which may in turn precipitate strategic processing decisions that have resource implications (for example, staff numbers and pay rates, number of processing offices and use of computer-assisted techniques). It is important to ensure that financial, resource and operational planning are coordinated, so that the technical assessment of requirements determines bids for resources, rather than vice versa.

2. Expertise, experience and rehearsal

604.   The coding of industry and occupation demands special expertise in those who plan, manage and supervise the operation. The processing of each census relies heavily on the technical information, expertise and experience gathered at the previous similar exercise. Information and experience from that exercise may be documented in detail, but practical expertise is likely to reside very largely in the heads of a small number of experienced staff. Staffing continuity in key positions is therefore extremely desirable. However, circumstances change, and it is not possible to rely entirely on documented or undocumented experience from the last census. Staff working with the labour force survey or other related surveys where information about industry and occupation is collected on a regular basis will often have very relevant experience and tools, and they should be consulted. If and when outside advisers are used it is also extremely important to verify the validity of their experience for the local context and to have the opportunity to modify their estimates of costs and time requirements, based on concrete experience. A processing rehearsal is very important as an aid to planning and estimation, and the results of such a rehearsal need to be available early and at a stage when it is still possible to adjust plans for the main operation in the light of them.

3. Estimating coding rates

605.   Some key planning parameters, for example, coder work rates and effective coding throughput rates, can be estimated reliably only from well-documented previous experience or well-planned tests. Certain problems arising from the scale of the full operation, such as its effect on problems of recruiting, maintaining and controlling staff, may be hard to test in advance. Experience also shows that performance rates vary significantly over time during the processing period. Coding rates are much lower and query rates much higher early in the process than later. There is also the danger that relaxing controls can cause a drop in standards towards the end of the process, beyond what is warranted by the improvements in the coding operation that result from the early improvements in the coding tools used and from on-the-job learning by the coders, supervisors and classification experts.

4. Coding staff

606.   It is important to make good estimates both of the number of coders required and of the numbers of first-line supervisors needed to control the coding process, as well as of the number of specially trained staff needed to resolve queries.

607.   In a census operation, the large volume of work to be processed within a limited period will require special staff recruitment and training, both for initial recruitment and

for anticipated staff turnover in the course of the task. Thorough enquiries and consultations should be undertaken about likely sources of suitable staff recruits, since financial constraints are likely to prevent actual recruitment until the last moment. There may be external pressures to employ particular groups of persons, even when their suitability cannot be guaranteed. Common criteria should be defined and applied in the selection of all staff. The terms on which staff are to be employed, including the minimum acceptable level of education, pay rates, grading, disciplinary guidelines and rules for hiring and firing, need to be carefully defined. It is important to provide adequate time and resources for staff training at both the coder and supervisor levels, and to recognize that the specialists, who are to advise the supervisors and resolve the more difficult queries, normally cannot, and should not, be recruited and trained for a particular census operation but be part of the permanent competence of the organization.

608.   Persons with the following characteristics are best able to perform the tasks of the coder:

(*a*)   Literate, with an aptitude for the job and a suitable temperament;

(*b*)   Willing and able to speak up when a problem is detected;

(*c*)   Clerically accurate and careful;

(*d*)   Willing and able to follow detailed instructions conscientiously, with an understanding that he or she should not alter or improve upon such instructions prior to consultation with and clearance from supervisors; prepared, however, to raise queries in cases of genuine doubt;

(*e*)   Honest and trustworthy;

(*f*)   Persistent and willing to work steadily for long periods;

(*g*)   Able to work reasonably rapidly and to maintain a steady level of productivity.

609.   Persons responsible for recruiting and selecting coding staff should have the above-mentioned characteristics in mind. Several of them, such as (*a*), (*c*) and (*d*), are best assessed through an objective screening test (which may also be applicable to other types of routine clerical tasks). Success at the initial recruiting stage will help to ensure retention of qualified staff and to limit the need for recruitment and training of replacements while production coding is in progress.

610.   Proper training of coders is very important. Coding of industry and occupation is best learned through practical instruction in specific procedures, such as document handling routines, use of the coding indices and so on, interspersed with supervised practice on appropriate, specially designed exercises. It is to be expected that trainees will learn at different rates; in this case, retention of what is learned and the ability to stay on task are key attributes. The training period can also be used to identify persons who are the most qualified and able to follow the coding instructions precisely and those who may be better suited to other types of census processing work. It is important to remember that staff recruited to replace coding staff who leave before the end of the coding operation will also need to be trained.

5.   Coding teams and supervisors

611.   Production coding on large jobs is best organized by allocating coders to teams, each under a first-line supervisor. The supervisor's role and work tasks need to be carefully specified and are likely to include the following: controlling workflows; monitoring and maintaining work rates; enforcing work discipline; motivating staff; resolving and recording coding queries; applying quality control procedures; and so forth. For first-line supervisors, the need to have previous experience in coding operations depends on their role in query

resolution. In principle, the coding operation may be organized in a way that gives the operational supervisors a very limited role in query resolution. Then, it is not essential for them to have had prior experience of industry and occupation coding. However, in most cases it will be preferable to give supervisors the responsibility for first-line query resolution owing to their close contact with the coders and the shorter response time and higher capacity relative to classification experts. Supervisors with responsibility for query resolution should be given a good understanding of and training in the classifications and coding systems.

612. The number of coders allocated to each supervisor is important. Typical ratios are between 6 to 1 and 12 to 1, but the appropriate ratio needs to be assessed in each case, taking account of the flow of work that supervisors will be required to manage. Maintaining a reasonable workload among supervisors will help to avoid bottlenecks, improve the reliability of coding and ensure that staff are able to report problems and queries in a timely fashion. It is also important to maintain staff morale and discipline and to sustain productivity rates. Particular problems may arise where coders expect that it will be difficult to find a new job after the end of the coding operation; in such circumstances, special bonuses may be effective in maintaining productivity rates. It may also be possible to retain to the end of the process only those staff with a potential for more permanent and long-term employment with the statistical service. Their experience will be valuable both for other surveys and for the documentation and explanation of census procedures for the benefit of the users, including those who will be preparing the next census.

6. Coding tools

613. It will be necessary to provide appropriate documentation, procedures and training materials not only to coders, to guide the coding process itself, but also to supervisors. The basic tools required by coders include the following:

(*a*) **Coding instructions.** They should cover all operations that the coder is required to carry out. The procedures and instructions for handling all relevant items and operations should be integrated. The instructions relating to the coding of industry and occupation will need to be particularly clear and specific with respect to (i) the order in which checking, coding and editing tasks are to be carried out; (ii) the procedure for analysing verbatim material for significant terms; (iii) the use of the coding index or other coding documents; (iv) the circumstances and procedure for using ancillary data; and (v) the circumstances in which to refer a "difficult to code" response to a supervisor for query resolution;

(*b*) **Coding index.** This is the key coding document through which verbatim terms incorporated in job titles, descriptions of tasks and the like, are translated into codes. Coders should not be encouraged to interpret verbatim responses in terms of their own conception of the purpose or criteria of classification, but rather to follow in a conscientious way the instructions laid down for consulting the index. For those reasons it is essential that the index be clearly set out, explicit and easy for coders to use. The use of the coding index, instructions and procedures should allow for updating in the light of decisions made in resolving queries and problems that arise and are dealt with in the course of coding;

(*c*) **Query resolution procedures.** A query occurs when coding clerks cannot assign a code using the specified procedures and indices. There should be clear instructions on when and how the coders should raise queries, how to record and report them, and their resolution. Queries are the most useful inputs to both immediate and future work to update the coding index and the classification itself. Early in the census processing, it may prove necessary to carry out such updating frequently if the coding indices, or even the classifications, prove to be out of date or inadequate for other reasons. The updating of coding indices should be done by, or in close collaboration with, those responsible

for the standard classifications so that consistency between collections can be maintained and any problems with the classification itself can be identified;

(*d*)   **Legal and administrative forms.** Coders should sign a legally binding undertaking to maintain the confidentiality of census data. Other documents used by both coders and supervisors are likely to include forms for recording queries and their resolution; for controlling the flow of work and reporting progress; for quality monitoring; and so on. To ensure that target throughput rates are achieved, the productivity of coders and coding teams needs to be monitored and progress charts maintained. Special measures for motivating coders should be used, for example, the posting of productivity and error rates for coding teams or for the best individual coders.

7.   Coding problems and queries

614.   A coding query occurs when a coder is unable to assign a classification code to a census or survey response using the standard coding procedures and tools. No matter how carefully coding instructions and the coding indices have been prepared, it can be guaranteed that large numbers of detailed queries will be raised in the course of a major coding operation. This happens if the indices prove to be out of date or incomplete in some respects. Another reason for queries is that actual responses will be more varied than anticipated by the index constructors, even with the most carefully analysed pre-census tests and well-designed questions and instructions to enumerators. Any revision of the structure of the classifications since the last census or survey may also lead to a new crop of problems in the treatment of vague and inadequate responses at the borderlines between categories.

615.   Evidence of shortcomings in the documentation that comes to light in the course of production coding will need to be rapidly and consistently addressed and fed back to the coders and their supervisors in the form of amendments to their tools. Appropriate procedures for reporting and recording queries need to be established in advance. In addition, decisions need to be made to resolve them and to incorporate any consequent amendments in the coding documentation and procedures. The roles of supervisors in processing queries and amendments need to be defined, for example, how and when they should communicate with the coding experts, and how new versions of the tools should be distributed to the coders. Particular care is needed in the coordination of query reporting and document amendment where coding is being carried out in several different locations, for example, if different provincial or local offices have responsibility for processing the census returns. All coding sites and teams need to be informed and given updated tools as quickly as possible.

8.   Quality assessment and quality control

616.   To provide adequate information on quality of output, explicit allowances for the resource and time costs of formal quality control, rather than casual observation by supervisors and ad hoc checking of coded output, need to be built into the processing plan. Those costs will cover the establishment and staffing of a quality control unit responsible for acceptance testing of coding during the start-up of the coding operation and for assessment of the reliability and consistency of the operation as a whole. A procedure for sampling and sub-sampling the work of the coders for quality control purposes needs to be defined, and the quality control unit should be staffed at a level that will enable it to keep pace with the main coding operation. Coding schedules must make allowance for corrective action (for example, 100 per cent checking) in the case of batches that fail a quality control test.

617.   Built-in quality control procedures will also be required. It is necessary to design separate procedures to handle (*a*) online acceptance testing of coders' work, and (*b*) overall monitoring and assessment of performance.

618.   Coders' performance must meet criterion levels of accuracy in following coding instructions. The aim of acceptance testing is to identify deficiencies in performance rapidly so that corrective measures can be taken. The aim of overall monitoring is to estimate average levels of coding accuracy and inter-coder consistency for the entire coding exercise. Estimates of coding reliability need to be supplemented by estimates of the validity of coding if a balanced overall assessment of the quality of the statistical output is to be made. Estimates of validity may be obtained from a post-enumeration study in which the whole data collection, coding and editing process is repeated for a sample of census cases. On the basis of such quality assessments it may be possible to separate the contributions made to total variance by source of variance, that is to say, due either to data collection or to coding.

9.   Premises, infrastructure and equipment

619.   The large volume of questionnaires and of work entailed in census processing requires suitable office space and the entire infrastructure necessary for properly supervised clerical operations, as well as for easy movement, storage and retrieval of forms. A special requirement is for security of documents bearing personal details. Proper attention must also be paid to the functionality and capacity of the desks, chairs, shelves, filing cabinets, lighting, heating and ventilation and to the adequacy of paper, pencils and other stationery. Attention to such factors can boost the morale of the staff, prevent high staff turnover and encourage attention to work quality and speed. Suitable premises need to be specified, identified, costed, approved and booked well in advance. If coding staff are to operate with special equipment, such as computer terminals or optical readers, special arrangements may be needed to estimate the requirement, identify suitable and reliable equipment and carry out tests, estimate and provide for capital expenditure and depreciation, provide for replacement in case of breakdowns, go through procurement procedures and so forth.

10.   Processing in one location or in several

620.   Census processing creates a substantial but temporary demand for suitable staff and premises. Recruitment of extra staff and other managerial tasks may be more complicated if all processing is carried out in a single location. There may also be other cost and logistical arguments for carrying out processing at one or several locations other than, or in addition to, the central census processing office. In any case, it needs to be borne in mind that for such relatively complex tasks as census coding and processing, maintaining consistency in coding between coders and coding teams, while difficult, is important. One reason for this is that census coding inevitably generates large numbers of queries. For example, in the 1981 census of the United Kingdom of Great Britain and Northern Ireland, more than 30,000 coding queries were processed, and it is thought that many more were dealt with informally. A large number of queries may lead in some cases to amplifications or changes to coding indices and instructions, which then need to be applied in a consistent fashion (the task is much easier if all documents are in electronic format). The maintenance of consistent production and quality control standards is generally more difficult to achieve when coding is carried out in several locations than if it is centralized in one place. It will be important to establish possibilities for communication by telephone, fax and/or electronic mail, especially if the operations are not centralized.

11.   Handling of documents

621.   The coding of occupation and industry will normally be an integrated part of the total processing of the information on the census questionnaires. In that case the main concerns will be (*a*) how to receive the forms; (*b*) how to store them; and (*c*) how to allocate them to staff so that one can control and ascertain that forms have been processed. It should be easy to find individual forms that for some reason need to be rechecked. If each form has to be handled by more than one person, for example, because different persons carry out the coding of different variables or because data entry is done by special operators, then the flow of documents must be planned to avoid bottlenecks and loss of forms. All movements of questionnaires from one location, such as a workstation or office, to another and to storage should be carefully recorded.

622.   Capturing the responses by intelligent character recognition to create electronic records is an option that has the potential to significantly reduce processing costs and problems caused by the flow of forms from one process to another. A less expensive and technically simpler option is to scan the questionnaires and to work with image files of the documents. When considering the cost of either intelligent character recognition or scanning, it should be noted that an archive is created as part of the process.

12.   Use of automatic or computer-assisted coding

623.   Until very recently, most census operations reserved computer usage for "downstream" applications. That is, after data had been collected, manually transposed to data collection sheets and keypunched onto cards or paper tape or entered on magnetic tapes or disks, the machine-readable data would be fed into a computer (usually a large mainframe computer) and verified through a series of computer edit checks. That process would involve passing the data through a computer program to ensure that any out-of-range codes or non-allowable combinations of codes were detected and reported for further investigation. Those types of controls are still valid and important, but computer edit checking cannot detect invalid coding within the range of valid codes if errors in data collection and coding do not generate non-allowable combinations of data. Apart from edit checking, data processing by computers was usually reserved for the final stage of the whole operation, the production of statistical tables.

624.   More recently, data processing has benefited from the development of database management systems as well as a wide range of applications for the vastly enhanced capacity of personal computers, standing alone or linked through internal networks. For large-scale data processing applications there are both mainframe and personal computer versions of the systems, which allow the non-specialist to design screens and forms and to develop a system for the processing of data at all stages, including data entry, edit checking, management of the processing facilities and the presentation of statistics for publication. Previously, the manipulation of census information or large-scale social survey information via such systems was prohibitively expensive in terms of programming and the hardware requirement for online data storage. Technical advances have now made large-scale database management a feasible proposition for such applications.

625.   It is also only relatively recently that real progress has been made in bringing computers to bear on the tasks carried out, in an inherently slow and laborious way, by census coders. However, the situation is changing rapidly, and systems for automatic coding (AC) or computer-assisted coding (CAC) of industry and occupation are now used in a number of countries. Automatic coding and computer-assisted coding improve the consistency of coding and reduce its time and cost. Both types of system

read captured responses. The responses can be captured by intelligent character recognition or input manually. Manual input requires operators with appropriate keyboard skills that may not be available in some countries. The success of intelligent character recognition varies by language, alphabet or characters used and requires responses to be written clearly and inside the designated boxes. Successful intelligent character recognition operations may require the use of special quality paper and ink, as well as handling procedures for the questionnaires before, during and after the data collection that will protect them from humidity, sun and other spoiling influences. That may prove difficult to ensure in many countries.

626.   So far, no workable system has been developed that can fully automate the human decision-making task in coding occupation and industry. Some automatic coding systems allocate codes automatically to more than 70 per cent of the cases, but the development costs have often been high and there have been problems in making the systems sufficiently "intelligent" to simulate reliably the performance of trained human coders. The reported error rates for responses coded by many of the systems are of the same order of magnitude as those of human coders. Moreover, the residual need for human intervention in the coding process in a substantial proportion of cases reduces the benefits that such automation could have on simplifying, speeding up and reducing the cost of data processing. Both manual and automated systems will have difficulty with vague and contradictory responses. Those limitations do not, however, eliminate the gains that can be achieved using an automatic coding system, in particular if it is integrated into a data entry and processing system that starts with optical reading of the questionnaires.

627.   Computer-assisted coding uses the computer to assist the coder to find and assign the correct code. Typically the coder inputs a keyword or the first few characters of the keyword. The index items appropriate to that keyword appear on the screen with the appropriate coding instructions. The process uses few keystrokes, as typically the coder can scroll through options presented on the screen. Computer-assisted systems maintain the advantages associated with manual coding but at the same time can enhance the quality, consistency and speed of the coding.

628.   Countries that have adopted automatic coding for such classifications as occupation and industry typically adopt a three-stage coding process. The first stage involves attempting to code all responses using the automatic coding system. In the second stage, responses that have not been assigned a code automatically are coded using computer-assisted coding. The cases that cannot be coded using standard coding procedures, usually a small number, are then dealt with by specially trained staff using query resolution procedures.

629.   The process described above presents a number of opportunities to introduce quality control measures. First of all, it is good practice to pass some or all of the captured response data through the automatic coding system before finally assigning codes to the data file. Coded data from the test runs can be sorted in code order together with the captured response, and the index entry matched, so that expert staff may identify any systematic errors or new response types. Adjustments can then be made to coding indices and, if necessary, to tolerances in intelligent character recognition systems. The procedure can be repeated several times before making a final pass of the data through the automatic coder and can lead to improvements in both accuracy and the number of responses coded automatically.

630.   Additionally, a sample of responses assigned a code by the automatic coding system can be selected for coding by computer-assisted coding, and any discrepancies analysed and corrected. Similarly, a sample of responses for computer-assisted coding

can be coded by two coders and any discrepancies referred to expert coders. This will make it possible to monitor errors. Coding staff who make a large number of errors can undergo further training or be assigned to more suitable duties. The procedures also allow the identification of problem response types and the establishment of error rates and inter-coder consistency rates.

631.   Although computer systems have the potential to yield dramatic improvements in the quality of data, to control the environment within which data are collected and to reduce the amount of time that elapses between data collection and data dissemination, the true costs of implementing the system must be explained. The costs must:

    (*a*)   Include a realistic estimate of the rate of depreciation of hardware, although it may frequently be put to good use for statistical surveys after the completion of the census processing;

    (*b*)   Take account of the dependence on specialist programming and systems analysis skills for the implementation of the required software.

632.   If it is intended to introduce an automatic coding or computer-assisted coding system, trials of the hardware and software and of the machine/operator interface need to be conducted well in advance. If there are other surveys (perhaps a labour force survey) that collect occupation and industry variables in a similar manner it would be advantageous to develop and test an automatic coding or computer-assisted coding system with those surveys. Until the feasibility and operational robustness of the machine-based system have been established, it is prudent to make parallel plans for reversion to a manual and/or clerical system as a fallback position.

633.   To minimize the risks of development, and probably the cost, one should seek to acquire the rights to operate a system that has already been tried and tested. In selecting a system, ease of operation and adaptation to national circumstances (for example, the language used by the operators, the classification and the coding index) should be given priority. The operational advantage of an automatic coding or computer-assisted coding system is likely to depend more on the type and cost of data registration and the suitability of the coding index than on any particular feature of the search and decision-making algorithms of the system. However, fast response times and an easy-to-understand interface between the computer and the operator should also be important considerations when selecting a computer-assisted coding system.

13.   The problem of different languages

634.   In the preceding commentary, little reference has been made to the problems encountered in countries where the population uses more than one language in daily life, although the issue of multiple languages in questionnaires was addressed in chapter II (see para. 76). On the assumption that enumerators will know the language of the respondent and therefore can write down correctly in that language the answers to the industry and occupation questions, the best solution will be to make sure that the coding indices can reflect those answers, as follows:

    (*a*)   Separate coding indices for each major language will provide the best solution where it is operationally feasible;

    (*b*)   Otherwise it may be possible to create multilingual coding indices to allow the coder, whether the enumerator or a specialized coder, to find an index entry that corresponds to what was written down as the response of the respondent. A multilingual coding index will be larger than a single language one, but not necessarily dramatically larger; in many countries the terminology reflecting modern sector activities and jobs will be common to many languages, and it will mainly be the terminology

reflecting traditional activities and jobs that will differ. Such activities and jobs are normally less varied than those in the modern sectors, and there may therefore be fewer terms in each language describing them.

635.   If the coding index can be constructed in only a single language, then someone, usually the enumerator, will be required to translate from the response to the language of the index. The problem with this is that the correct translation of occupational terms will require not only a good general knowledge of the two languages in question, but also knowledge of the particular area of work in order to understand precisely how particular terms for activities, products, services and jobs are used in the local context. Very few persons will normally be able to satisfy that requirement over the whole range of work situations covered by a population census or labour force survey.

Chapter XI

The development and use of coding indices

636.   The present chapter is based mainly on experience from English-speaking industrialized countries, as the limited documentation readily available on the development and use of coding indices and coding procedures has originated mainly in such countries. It is difficult to assess the extent to which the documented experience is transferable to other languages and cultures. While that caveat should be kept in mind when reading the following text, the experience from those English-speaking countries may provide a good starting point for work and experiments in other languages.

A.   What is a coding index?

637.   The process of coding industry and occupation information involves the task of matching responses to questions in a census or survey with index entries to find the appropriate classification codes. The coding index is the key instrument for the matching process. The index can take the physical form of a durable printed publication, a loose-leaf binder, a computer printout or a machine-readable file within a computer system. The matching can be carried out by a person (the coder), by a computer or by means of interaction between the coder and a computer.

638.   Detailed industry and occupation coding has traditionally been carried out using clerical (manual) procedures. Relevant information is recorded verbatim in the field by the respondent or by enumerators and the resulting raw data are brought to one or more central offices. There, clerical staff scrutinize each case, decide to which industry or occupation category to allocate it and record an appropriate code on a document, or directly onto a computer-readable medium, for further processing. It should be noted, however, that automatic coding and computer-assisted coding systems are emerging rapidly. Users of manual coding should consider the eventual adoption of such systems in the long term and begin preparing for them far in advance of their implementation.

639.   The coding index is the principal instrument for linking the words used in the various parts of the response to the numerical code that represents the allocation of that response to the corresponding category of the classification. The coding index guides the coder by listing information, for example, keywords that can be found in the responses, and indicates how different responses are allocated to the detailed or more aggregate categories of the classification, depending on the nature of the information in the response and on the instructions for the coding process. Thus, the census coding index must be a reflection of the types of responses that the respondent writes on the census form or that the enumerator writes on the basis of information received from the respondent. It should contain the words and expressions that the respondents will use when asked to give the information about their place of work and their job in the census enumeration.

640.   It is important to recognize that a coding index is different from the index of categories specified in the classification, which is usually just a list of the titles given to those

categories that are separately defined in the classification. The titles chosen for the categories in the classification are designed to be as descriptive as possible for the category content, given that only a few words may be used. Only a few of the titles will correspond to the terms used by individuals when asked about the activities of their place of work or about the tasks and duties of their job. The coding index is also different from a list of titles or terms that may have been chosen to be descriptive of the content of the categories. That type of list may also contain entries that may never be used as a title by any person describing their job or place of work. When such a list exists, it may, however, serve as a useful starting point for the construction of a coding index.

641.   Since it has to be in place before census coding operations start, the coding index has to be constructed in anticipation of what the census responses will look like. The basis therefore will have to be actual responses to similar questions in the last census, in household surveys carried out after that time and in census pre-tests. Terms and expressions concerning the types of economic activities in which people participate and the jobs that they have can also be found in advertisements for products and services (for the industry index) and job vacancies (for the occupation index), and in registrations of vacancies and job seekers in employment offices.

642.   The collection and coding of the elements to be included in the coding indices should be carried out by experts on the respective classifications to ensure that they are done correctly. The work will be painstaking and time-consuming, but the investment involved in the collection and coding of up to five, ten, twenty or even thirty thousand index entries ahead of the census will prove well worth the effort in terms of the speed and reliability with which hundreds of thousands or millions of census forms, depending on the size of the workforce, the complexity of the economy and the coding strategy chosen, can be coded during census operations.[1]

643.   At the start of the census operations the coding index must be considered as incomplete. While provisions must be made to update it during the whole period of the census coding operation, it will probably be necessary to do so much more frequently and with more new items early in the process than later. The updating process should be an extension of the query resolution process in the sense that the nature and outcome of resolved queries should be made available to all coders as soon as possible, in case they encounter the same situation. The best approach is to issue a new version of the coding index. It is better to issue a completely new version of the coding index than to issue additions to an existing index, since the new entries will belong in different places in that index. Furthermore, if the coders have to transfer the information from a note on additions into the main coding index, there is a danger that they will make mistakes. New, complete versions of the coding indices issued frequently in the first weeks of the coding operation will also reduce the danger that individual coders will assign codes on the basis of their own notes on the coding of particular responses. Such notes can easily be the source of systematic differences between coders in the coding of responses not reflected in the initial version of the coding index.

644.   During census operations, the physical form of the coding indices must reflect the temporary nature of the version currently in use at any point in time. Each issue should be precisely dated and, when choosing the form of the paper versions, the main consideration should be the speed of reproduction. Both paper and electronic versions should be issued with orders to destroy earlier versions.[2] If using paper, a binder is very adaptable if space is initially left on each page for several additional entries. It is then possible to reprint and replace one page without requiring changes to other pages. If using an electronic index it is best to keep the index on a server. Changes to an index on a server are immediately reflected at all workstations. If, for reasons of efficiency, it is necessary for copies of indices to be held on workstations, arrangements need to be in place to ensure that updates to all workstations are made automatically for the server.

[1] It is in many ways correct to regard the coding index as the ultimate manifestation or embodiment of the classification. However, it should be kept as a working tool and not be given the status of being an official part of the classification, for the following two reasons: (*a*) it will be necessary to update the coding index during the census operations and when it is used later in other coding operations. That flexibility may be difficult to achieve if it is a formal part of the classification; (*b*) in order to reflect actual responses the index may need to include terms, e.g., brand names, which are commonly used to describe places of work or types of jobs, but which may be trademarks protected by copyright or may be difficult to include for other reasons in an official publication. Such problems will not usually arise when the coding index is used only as an internal working document and is made available to other organizations only for use as such.

[2] Copies of each version of the index need to be retained by management and by the team responsible for the classification, however, to assist in analysis of data quality and to allow for recovery in the event that an error introduced into the index results in systematic errors in the coded data.

645.   Only when it is clear that further significant additions to the coding index will not be forthcoming during the remainder of the census coding operation should it be issued in a form more suited for dissemination to other users, such as government departments, survey agencies and academic users. Then a well-bound quality publication may be considered, but a format more compatible with its status as a working tool and with regular, although less frequent, updating may be preferred: for example, a ring binder or computer printout may be the appropriate format.

646.   Up to this point, the presentation of the coding operation and coding tools has been relevant to both industry and occupation coding, as the issues are the same for the two variables, and when both are present in the census those issues and the preparations should be addressed jointly. However, as the information sought for the two variables is different, the responses obtained will also be different, and therefore the content and construction of the two coding indices will be different. They will therefore be presented separately in the remainder of chapter XI. Sections B and C concern the occupation coding index and its use, while sections D and E concern the industry coding index.

B.   Developing and updating the occupation coding index

1.   Sources of information for construction of the index

647.   The process of updating a coding index for occupations should be one part of the general process required to maintain the national standard classification of occupations. As new ways of organizing work between or within enterprises or new technologies are introduced, new jobs will appear with new combinations of tasks or new types of tasks associated with them. The new jobs may be given new job titles by their incumbents or their employers, or may be referred to under existing job titles. At the same time, existing jobs may be given a new title without their tasks and duties being changed in any significant manner, as a result, for example, of reorganization of the enterprise or of a change in their placement in a wage hierarchy. Thus, there is a need to keep track of job titles and the associated job descriptions, monitoring the relationship between that information, index entries and the associated occupational codes. If the preceding tasks have not been undertaken by the custodian of the national standard classification of occupations or by any other party or entity, then it will be necessary either to update the occupation coding index or to create it from scratch for the census.

648.   A full-scale job content monitoring exercise cannot be used for census preparations as that would be too time-consuming and costly. The most realistic alternatives would be to carry out (*a*) post-coding reviews of recent survey operations; (*b*) job vacancy reviews; (*c*) consultations with job placement services; and (*d*) a job monitoring exercise for the industries known to be either restructuring or undergoing technological change.

649.   The first procedure makes use of the tools and members of the coding team(s) used for recent survey(s). Coders are a good source of information on the adequacy of the index that they have used. Their suggestions for improving the index and for additional or revised entries must be recorded and investigated. Ideally, preparations for the use of post-coding review procedures for the development of index material should have been made an integral part of the design of the data processing operation for the respective surveys, as it is essential that information that may be useful in updating the classification and index be carefully collated and retained. Such information typically consists of records of problems encountered, queries raised and decisions and amendments to the working instructions adopted in the course of coding. A sample of the actual coded responses captured in machine-readable format can be very useful in classification development, index

development, development of automatic coding or computer-assisted coding systems and coder training. Such procedures should be made part of the normal routine for any continuous or regular surveys, such as labour force surveys, as well as for the registration of occupations, which takes place in the local offices of the employment services.

650.   Job vacancy descriptions available from job advertisements in newspapers and journals, on bulletin boards and the Internet, or from registrations at local offices of the employment services may provide a useful source for constructing or updating information on job titles and detailed job descriptions, and therefore also a coding index. The descriptions may be especially helpful where they have been coded to the occupational classification as part of the job vacancy recording process in an employment service. From such announcements of job vacancies, it should be possible to find a reflection of the impact of technical and organizational change on the allocation of tasks to jobs, and to develop proposals for new entries for the index (and the classification). The advantage of that approach is that it does not require an expensive initial search for contacts, as follow-up enquiries to a vacancy of interest can use the name, address and contact person of the employer found in the vacancy notice. The main disadvantage is that job vacancies that have been advertised in newspapers or on bulletin boards, or submitted to and recorded at an employment agency, normally cover job vacancies for only a limited range of occupations and industries.

651.   In cases where no national index of occupational titles exists, the index developed for the International Standard Classification of Occupations (ISCO) may be a useful starting point for the development of a national index. It should be stressed, however, that even for English-speaking countries using a national classification based on ISCO, the ISCO English index would need to be modified for use in national circumstances, as the words in the index need to reflect national usage and there is great variation between countries. Direct translation of the English, French or Spanish ISCO indices into other languages will not yield good results. Those indices are useful sources of information concerning the content of categories in ISCO and on the correct placement of some internationally used occupation titles.

652.   In some countries employment agencies have established standard procedures for the collection of job vacancy materials. For example, when employers contact an employment agency, the agency may create a computerized record of information containing the job title and a brief description of the main duties or tasks associated with that job title. Those records may be used within a word-processing system or may form part of a database of vacancy information. In some cases they may have been linked, or coded, by the national employment service to the national standard classification of occupations.

2.   Organization and structure of the index

653.   When organizing the material for the coding index of occupations, the first issue concerns the structure of the index itself. Basically the choice is between two different approaches, which can be described as "all-inclusive" and "structured".

654.   When the all-inclusive approach is adopted, every distinct type of response found in the process of coding should, in theory, have an entry in the index, although allowance may be made for misspellings and inversions of words that are without consequence for the meaning of the response. An advantage of this approach is that it may be possible for coders, when faced with an obscure job title and/or task descriptions, to find that title and/or those tasks listed in the index. A significant disadvantage is that an all-inclusive index may become very large and its sheer size may slow down the process of searching for the appropriate entry in the index, and thereby slow down the coding, whether the coding is done manually or is computer-assisted.

655. Furthermore, large verbatim indices may create the impression that coding is a simple task, involving a straightforward matching between a response and an index entry. However, no matter how large the index (and some occupation indices have over 30,000 entries) it will always be the case that a significant proportion of responses fail to match the index entries exactly. In those cases the coder has to be given rules and/or use judgement to make the best match. Consequently, in a number of statistical agencies the approach has been to develop a structured index.

656. A structured index does not try to reflect every possible response because it is accompanied with instructions to the coder on how to break down the available response into keywords and qualifying nouns or adjectives. The primary entries in the index are the keywords. If a keyword in itself is not sufficient to uniquely identify the category, an appropriate qualifying word (or phrase) must be added to distinguish between the possible alternatives having the same keyword. If that is not sufficient to resolve all ambiguities, a second or higher-order qualifying word should be used. The following examples may illustrate the system for transforming an occupational response into an entry in a structured coding index according to the following format:

Response: → keyword/first qualifying word/second qualifying word

Cost accountant: → accountant/cost

Drilling machine operator: → operator/machine/drilling

Aircraft instrument maker: → maker/instrument/aircraft

Room maid: → maid/room

Marine biologist: → biologist/marine

Capstan lathe setter-operator: → setter-operator/capstan lathe

657. The following examples, based on the coding index used for coding occupation in the 1986 population census in Australia, will illustrate the use of the qualifying words as well as the way instructions about the use of the index can be incorporated with the index entries. The codes given are from *Australian Standard Classification of Occupations—ASCO Coding System: Unit Group Level* (Australian Bureau of Statistics, 1992):

5999	Researcher/market/interviewing
2909	Researcher/market/statistician
2907	Researcher/market (*except above*)
2701	Researcher/accountancy
2107	Researcher/agricultural
2907	Researcher/anthropology
2999	Researcher/assistant to parliamentarian
2107	Researcher/biological sciences (*except medical*)
2101	Researcher/chemistry (*except medical*)
2109	Researcher/medical
3103	Researcher/toxicology
2000	Researcher (*no additional information about type of research*)
8919	Restaurateur/assisting in kitchen
4705	Restaurateur/cooking
1503	Restaurateur/supervising staff and administration
6505	Restaurateur/waiting on tables
1503	Restaurateur (*no additional information about specific tasks*)

3999	Retoucher/photographic
4503	Retoucher/printing
1311	Secretary/assistant/senior government officer/computing division
1307	Secretary/assistant/senior government officer/distribution division
1313	Secretary/assistant/senior government officer (*except above*)
6503	Secretary/club/tending bar
1599	Secretary/club (*except above*)
1201	Secretary/trade union
5601	Secretary/receptionist
5101	Secretary (*no additional information about specific tasks*)
5101	Secretary (*except above*)
4405	Signwriter[3]
4921	Silverer/glass
4923	Silversmith
2815	Singer

3 Note that although the term "signwriter" may not appear in dictionaries as a single word, it needs to be included in the coding index if it is given as a single word in census responses. An entry "Writer, sign" would also need to be included in the index.

658.   The much smaller number of index entries in a structured coding index than in an all-inclusive index is a result of the restriction of the index to keywords and essential qualifying words when possible and of the use of the instructions "except above". Consequently, it is possible to exclude from the index words used in responses that are immaterial to the selection of the correct category.

659.   In a structured index for coding occupation, the keyword is the single word in the relevant response that alone can serve as an occupational title, however imprecise. The qualifying words will usually indicate some form of specialization of tasks. Sometimes the keyword may be precise and in itself suffice as an index entry, such as, "silversmith" in the examples above. However, the keyword may also be very ambiguous, as in the "researcher" and "secretary" examples above. Note that the qualifying words in some of the "secretary" examples do not serve to distinguish between specializations but between occupations that are very different in nature.

660.   The construction of a structured coding index must reflect and support the coding rules to be used for assigning occupation codes on the basis of the responses to the relevant questions, using, when necessary, permissible ancillary information given in other responses and indicated in the index as qualifying words. This means that one should organize the index alphabetically first in terms of keywords, then in terms of the first qualifying word. The entries that also have a second qualifying word should be listed before those that do not, and the "except above" instruction should follow the entries with qualifying words. The keywords listed in the index must reflect those which can be selected from the permissible parts of the responses, and the qualifying words must reflect those that can permissibly be selected, in the order in which they should be selected.

661.   In English, it will be normal to use as keywords those that can be found in either (as first priority) the title component of the occupational response or (as second priority) the task component. First priority for qualifying words will be those normally found in the title or task components of the response. Second priority, and based on the rules for using industry information in the coding of occupation, should be given to words found in the response for industry or name/type of employer.

662.   The advantages of having a structured coding index are threefold. First, it causes the coder to search for index entries in a way that is consistent with the coding rules.

Second, it speeds up the task of coding by restricting the coder's search through the index owing to the smaller number of entries. Third, when the index is searched by either computer or human, it reduces the risk of finding matches with words that are irrelevant for the coding decision and therefore reduces error.

663.   Some words may be found to be in common usage in job titles, but can be ignored for the purpose of creating a coding index. For example, such words as "senior", "junior" and "executive", when given as qualifying words, may not carry information about the tasks that constitute a particular job. It is possible to exclude such words from the index and to instruct coders to ignore them.

C.   Using the occupation coding index

1.   Using the occupational response

664.   Coding can be seen as a process in which the task of the coder is to translate the information provided by the recorded responses to the appropriate code in the occupational classification structure. The main tools for the translation are the coding index and the coding instructions—including instructions on when a response should be treated as a query to be resolved by supervisors or expert staff. The instructions should specify the following:

   (*a*)   How the translation process should be carried out;

   (*b*)   What items to look for in the occupational response and in what order;

   (*c*)   What type of ancillary information to use from other responses;

   (*d*)   When such ancillary information can permissibly be used and how to use it.

Ideally the coding index should have been constructed to reflect and support the use of these instructions.

665.   The starting point should always be the response given to the occupational question(s), specifically, question(s) on the type of work in the person's own job and on the usual or main tasks and duties carried out in the job. The question should normally elicit a job title and a few words on main tasks. When using an all-inclusive coding index the coder should start by marking the words that are relevant for the search in the index. When using a structured index the coder should identify the keyword, normally part of the response pertaining to the job title, and look for it in the index. The part of the response relating to tasks should then be used where necessary to supplement or modify the information provided by the job title. Alternatively, those tasks may need to be transformed to a title, for example, "baking bread" to "baker, bread"; "cleaned school" to "cleaner, school". Transformation of a task response to a title should be performed when there is no proper title response or when there is no index entry corresponding to the title given, as may be the case for "civil servant", "helper" and other non-informational titles. If the title is inconsistent with the tasks or if the occupational responses are not sufficient to determine a detailed occupational category, then the coder should choose one of the following three alternatives:

   (*a*)   Look at the form for recorded ancillary information of a specified type for further clarification;

   (*b*)   Use an appropriate code for inadequately specified responses;

   (*c*)   Refer the case to supervisors as a query.

The coders should be given clear instructions and training on the circumstances that would allow any of the alternatives to be chosen. Coders should not be expected to make judgements that require an understanding of the classification structure.

2.   Using ancillary information on industry or name and type of employer

666.   Most modern occupational classifications, including ISCO-88 and ISCO-08, are designed according to the principle that "occupation", meaning a particular pattern of work tasks and skills that constitute an individual's job, should be kept conceptually separate from "industry", meaning the sector of the economy in which the job is performed. Thus, an "electrical maintenance fitter" may work in any of a range of different industries, and that person's occupation cannot be validly deduced from a knowledge of the industrial category of the employing organization. Without violating the above-mentioned principle, it must nevertheless be recognized that certain occupations or occupational titles are to be found solely or predominantly in particular industrial sectors. In such cases, knowledge of the industry may clarify an occupational title or description that is inadequate for coding purposes. For example, a "face worker" working in the coal mining industry may be deduced to be a miner engaged in coal cutting.

667.   In other cases the descriptions of work activities used to identify an occupation are best formulated in terms of, for example, the nature of the material worked with (such as wood, rubber or leather). That information may be deducible from knowledge of the industrial sector in which the job is located and may again help to clarify a vague occupational description. For example, the occupational term "coil winder", used on its own, is ambiguous since coding depends on whether what is wound is some form of metal wire, some form of textile product or some other material. Knowledge that the job is located in a textile-manufacturing establishment may be sufficient to resolve the ambiguity with a reasonable degree of certainty.

668.   Thus, since some interrelationships between occupation and industry are inherent in the industrial structure of the economy, they can be made use of to improve the accuracy of occupational coding. However, there are costs as well as benefits in that practice. In the first place, there is always some danger that inferences from industry to occupation will be based on incorrect or out-of-date assumptions about the distribution of occupations across industries. In the second place, when coding is being done on a large scale, coding work rates and inter-coder consistency are important considerations. Coding rates are likely to be slowed if the coder routinely considers extra sources of information to arrive at a code, particularly if the extra information is itself hard to interpret. In such circumstances, there is also a danger of increasing inter-coder variability, since different coders will tend to interpret the information in different ways. The latter two problems are minimized if:

(*a*)   Industry is coded in advance of, or at the same time as, occupation so that no further interpretation of verbatim responses is required for the occupation code;

(*b*)   Coders are allowed to use data on industry only where the responses to the specific questions on job title and activities are inadequate to determine an occupational code;

(*c*)   The choices that an occupational coder can make on the basis of the industry code are exhaustively specified through index-referenced instructions.

669.   A simplified example may clarify the above. A coder encountering the job title "coil winder" would be instructed in the coding index under "winder, coil" to look first at the description of job activities for information on the type of material wound. In cases where no indication of that was found, the coder would look next at the pre-allocated industry code. If it was code "x", standing for "textile manufacturing", the occupational category would be determined as "textile yarn winder"; if the industry code was "y", standing for "electrical machinery manufacturing", the occupational category would be determined as "wire winder". If the industry code were other than

x or y, the occupation would be placed in an appropriate "inadequately described" category by the coding index.

3. Use of other ancillary information for coding of occupation

670. Some countries permit coders to use information about the educational and vocational qualifications of respondents as ancillary information to determine the appropriate occupational code. That approach should be based on detailed knowledge of the relationships between training and qualifications, on the one hand, and the corresponding occupations, on the other. In all countries the relationship varies between occupations, and in most countries the relationship is close for only a limited number of occupations. Even when the relationship is close, the fact that a person has a particular qualification does not mean that his or her job will include the corresponding tasks. For example, a person with a medical degree who is working in a hospital may not have tasks corresponding to his or her degree. That may be because the person has been promoted to a job that consists mainly of management tasks, or it may be because that person could not get a job corresponding to the type and level of his or her technical training (for example, owing to a lack of necessary language skills).

671. For the preceding reasons, education is most useful in an exclusionary function. A person working in a hospital without a medical degree can reasonably be excluded from being a medical doctor. The exclusionary function also can be employed as an edit to resolve problems associated with distinguishing between levels of the same discipline. The use of information on qualifications or educational attainment as ancillary information should therefore be very carefully controlled and probably restricted to query resolution by expert coders or to computerized edits of coded responses.

672. The use of ancillary information of any kind in an occupation coding process may have a tendency to bias the relationship between occupation and other census variables collected and may also have a negative impact on the efficiency of the coding process. Coders should not, therefore, be allowed to use responses to questions on age, income, hours worked or other socio-economic variables to determine an occupation code.

4. Inadequate occupation responses and queries

673. Some responses simply cannot be coded to a detailed occupational category. That will normally be for one of the following reasons:

(*a*) The response may be vague: it does not contain enough information to be coded according to the coding index and coding rules;

(*b*) The response may be precise, but may use a title and/or indicate types of tasks or combinations of tasks that do not correspond to any of the index entries. Supplementary material should be used by experts to code such precise responses and then make appropriate changes to the index.

674. Unfortunately, the number of cases of type (*a*) is likely to be substantial, even with well-formulated occupational questions and well-trained enumerators. In order to keep to manageable proportions the number of queries that the supervisors and expert coders must handle, the coding index and the coding instructions should be designed to guide the coders with respect to the most common of such cases.

675. The simplest solution will be to specify that the response should be coded to a default category, as in the examples of "researcher", "restaurateur" and "secretary" in the above-mentioned example from Australia (see para. 657). The default category may in some cases be a specific detailed category since it reflects the dominant usage of the

terms found in the response: for example, "1503 Restaurateur" and "5101 Secretary" as default categories. However, the default category will often have to be one of the aggregate categories in the classification, since it is not possible to identify one particular detailed category as dominant within the aggregate category indicated. In the example from Australia above, "2000 Researcher" indicates that a response giving only "researcher" as information could be coded validly only to major group 2. Similarly, a response like "manager, running a business" would normally have to be coded to the aggregate category for "managers", unless the industry response gave very clear information.

676.   There is a real danger that coders may use default categories as dump categories for difficult-to-code responses before they have tried to find a precise code. Some countries have therefore tried to avoid their use by first-line coders, preferring that such codes be used only by supervisors. However, that strategy may create a very large query burden for the supervisors. It may therefore be better to allow and monitor carefully the use of default codes by first-line coders.

677.   The fully specified responses that are not adequately covered by the coding index should always be handled by expert coders and recorded carefully. This is done because it is important to ensure consistent treatment of similar cases and because those cases represent an important source of information for the updating of both the coding index and the classification. During the coding operation, such cases can be handled either by using priority rules specified for the classification or by assigning them to one or several categories (or supplementary codes) for occupations not adequately covered by the classification.

678.   Priority rules can be applied to some of the responses that indicate task combinations that cut across the categories defined in the classification, for example, "baker, baking/selling/managing shop". Most classifications based on ISCO-88 or on ISCO-08 will specify priority rules in terms of tasks performed for the allocation of such jobs to occupational groups. In ISCO-88 it is specified that priority should first be given to the tasks that require the highest skill level, and secondly that production-oriented tasks should be given priority over managerial or administrative tasks. "Main tasks", in terms of time spent, for example, are not to be given priority unless they completely dominate, both because an employer is likely to be concerned that a worker can carry out the most skilled tasks required, even if they are only seldom activated, as in emergencies, and because information about time allocation of tasks is normally not available. Thus, in the above example, the code to be specified in the coding index should be the code for "baker" according to the priority rules of ISCO-88. However, this may change in ISCO-08.

679.   Precise responses that cannot be resolved by priority rules should be assigned special "inadequately described" or "not defined" codes placed in categories created for coding purposes within the aggregate categories to which the jobs clearly belong. Steps should also be taken to ensure that these cases can be closely examined outside the coding operation to determine what, if any, contribution such cases can make to the updating of the coding index and to the classification. Note that these categories are not the same as the "not elsewhere classified" categories of the classification. Great care must be taken to avoid confusion between the two types of category.

D.   Developing and updating the industry coding index

1.   Types of industry coding index

680.   Most census coding operations will find it useful to have two coding indices for the coding of industry, as follows:

(*a*)   A list of as many as possible of the establishments which, at the time of the census, are/were operational in the geographical region covered by the coding operation, where each establishment has been given the correct industry code by those who are specialists in establishment surveys and in the coding of establishment activity. In practice, such lists (or business registers) may often cover only large, formal sector establishments as they have been created from lists kept in tax offices, licensing offices and/or chambers of industry and commerce. They may nevertheless cover significant proportions of the workforce, and their use for census coding will eliminate one possible source of inconsistency in employment statistics between the census results and the results of establishment surveys;

(*b*)   A list of significant word combinations reflecting the answers given in response to industry questions, that is, an index of the same type as that created for the coding of occupation.

681.   The industry coding process will therefore usually attempt first to match the name and address of the person's employer with those in the list or register of establishments. If a match cannot be made using the register of establishments, then an attempt is made to match the description of the industry with the index of type (*b*) above.

682.   The process of updating the coding indices for industry responses should be viewed as part of the general process required to maintain the industry classification. As new ways of organizing work between or within enterprises or new technologies are introduced, new functions, end products and services with separately identifiable units (separate establishments) will appear. For example, with the introduction of movies on videotape, units of production providing rentals of the tapes were created. New industries will be created and some old industries may disappear. However, those responsible for census preparations will frequently find that while work to maintain establishment lists of type (*a*) above may have been carried out by those responsible for an establishment register or for establishment surveys, that work has not been undertaken with respect to lists of type (*b*), neither by the custodian of the national standard industry classification, nor by anyone else, and that in fact not only the updating but also the creation of this type of industry coding index must be undertaken from scratch for the census.

2.   Lists or registers of establishments

683.   Where lists of establishments or business registers already exist for use in economic statistics or are used for administrative purposes, they should be adapted for use in census processing. Since information provided by households in censuses and surveys may be less detailed and precise than that available in economic collections, it may be necessary to make adjustments to the list to reflect the type of information typically provided in response to questions on census forms on name and address of place of work or employer. If no such list exists it may be useful to compile a list of known large establishments and pre-code and assign an industry code to each establishment.

684.   For each unit, the list of establishments should give both a name and the physical location, indicated by a street address if possible, otherwise by naming the (smallest) district in which the unit is located or is operating as one unit. If alternative forms of the name, such as abbreviations, initials or old names, are used or have recently been in use, they should also be included in the list as separate entries, owing to the possibility that persons working for them may use those variants in their answers. The entries in the list should be organized alphabetically, with clear rules for where to find entries consisting of initials and abbreviations.

685.   Notices and advertisements in newspapers, journals and on the Internet about their products and services, creation and expansion, or about vacancies, will be used by

many formal sector businesses, and they may provide a useful source for constructing or updating information on establishments and their activities. Telephone directories, trade directories and similar publications are most useful. They often classify industries (the yellow pages, for example) and in some cases provide short descriptions for some entries. The main disadvantage is that small, informal sector establishments are not normally well covered by such reference materials.

686.   Experience suggests that in most countries only a minority of census responses will be coded successfully using a list of establishments. Some developed countries, for example, report that 30 or 40 per cent of industry responses can be coded in that way. Since assignment of the code is based on known characteristics of the establishment, the codes assigned will have a high level of accuracy.

3.   Indices reflecting the answers given in response to industry questions

687.   Most industry coding is likely undertaken using indices of type (*b*). It is important to note that indices of that type for use in population censuses and other household or person-based data collection activities are very different from those used in economic surveys of establishments. The words used in responses provided by individuals to questions about the kind of industry of their employers are likely to be very different from those provided in establishment-based collections. In countries where information on industry is regularly collected in household surveys, such as labour force surveys, it is possible that an index of type (*b*) may be maintained for that purpose.

688.   A full-scale establishment monitoring exercise cannot be used for census preparations, as that would be too time-consuming and costly. The most realistic alternatives would be to carry out (*a*) post-coding reviews of recent household survey operations; (*b*) reviews of advertisements and notices in newspapers and other media; and (*c*) consultations with tax authorities, chambers of commerce and the like. In addition, as much help and information as possible need to be obtained from those responsible for establishment surveys.

689.   Procedure (*a*) makes use of the tools and members of the coding team(s) used for recent survey(s). Coders are a good source of information on the adequacy of the index and other tools that they used. Their suggestions for improving the index and for additional or revised entries must be recorded and investigated. Ideally, preparations for the use of post-coding review procedures for the development of index material should have been an integral part of the design of the data processing operation, as it is essential that information that may be useful in updating a classification and index be carefully collated and retained. Such information typically consists of records of problems encountered, queries raised and decisions and amendments to the working instructions adopted in the course of coding. Such procedures should be made part of the normal routine for continuous or regular surveys, such as labour force surveys. They should also be applied to the registration of establishments and activities, which takes place in the local tax offices, licensing offices or chambers of commerce, and which is also carried out by those responsible for establishment surveys.

690.   When organizing the material for the coding index of activities (industries) the first issue concerns the structure of the index itself. Basically, the choice is between two different approaches. In some statistical agencies the approach has been that the index should be all-inclusive, that is, every distinct type of response found in the process of coding should, in theory, have an entry in the index, although allowance may be made for misspellings and inversions of words that are without consequence for the meaning of the response. An advantage of that approach is that it may be possible for coders, when

faced with an obscure type of activity or product, to find those terms listed in the index. The main disadvantage is that the index may become very large and its sheer size may slow down the process of searching for the desired entry in the index, and thereby slow down the coding, whether the coding is done manually or is computer-assisted.

691.   Large verbatim indices also create the impression that coding is a simple task, involving a straightforward matching between a response and an index entry. However, no matter how large the index, it will always be the case that a significant proportion of responses fail to match the index entries exactly, and one has to use rules and/or judgement to make the best match. If an exact match is not found, the chance that an incorrect entry will be selected will tend to increase with the number of entries available. For those reasons, an alternative approach may be to develop a structured index.

692.   A structured index does not try to reflect every possible response. Instead, it is accompanied by instructions to the coder on how to break down the available response into keywords and qualifying nouns or adjectives. The primary entries in the index are the keywords. If a keyword in itself is not sufficient to uniquely identify the category, an appropriate qualifying word (or phrase) must be added to distinguish between the possible alternatives having the same keyword. If that is not sufficient to resolve all ambiguities, second or higher-order qualifying words should be used. The following examples may illustrate the system for transforming an industry (type of activity) response into an entry in a structured coding index according to the following format:

Response: → Keyword/first qualifying word/second qualifying word:

Sheep farm: → sheep/farm
Car rental agency: → car/rental
Youth club: → club/youth
Tax assessment office: → tax/assessment office
Cleaning service: → cleaning/services
Cleaning products production: → cleaning/products/production

693.   The following examples are from the coding index used for coding industry to the United Kingdom Standard Industrial Classification of Industrial Activities, 1992:

15.11/1 Abattoir
74.40   Advertising/agency
74.40   Advertising/agent
74.40   Advertising/campaigns/creation
74.40   Advertising/campaigns/realization
74.40   Advertising/consultant
74.40   Advertising/contractor
92.11   Advertising/film/production
31.50   Advertising/lights/manufacturing
74.40   Advertising/material/design
22.22   Advertising/material/printing
22.22   Advertising/newspaper/printing
22.12   Advertising/newspaper/publishing
74.40   Advertising (*no further information*)

694.   In a structured index for coding industry, the keyword is the word in the relevant response that alone can serve as a designation of a service, a product or a function, however imprecise. The qualifying words will usually indicate some special form or variety

and/or the type of activity associated with the product or service. That sequence has been chosen because the number of different designations for activities is much smaller than the number of designations for different products, services and functions. Sometimes the keyword may be precise and in itself suffice as an index entry, such as "abattoir" in the example above. However, the keyword may also be very ambiguous, as in the advertising examples above.

695.   If the coding of industry and occupation is to be done by the same staff, the coding rules and structure of the coding index for industry must not be in conflict with the rules to be used for assigning occupation codes. Industry codes are assigned on the basis of the responses to the questions on industry, using, when necessary, permissible ancillary information given in other responses and indicated in the index as qualifying words. Consequently, the index should be organized alphabetically, first in terms of keywords and then in terms of the first qualifying word, with those entries that also have a second qualifying word listed before those that do not. The "except above" or "no further information" instructions should follow the entries with qualifying words. The keywords listed in the index must reflect those that can be selected from the permissible parts of the responses, and the qualifying words must reflect those that can permissibly be selected, in the order in which they should be selected. Rules for identifying the key or functional words for industry may need to be quite different from those for occupation, however.

696.   The advantages of having a structured coding index are threefold. First, it causes the coder to search for index entries in a way that is consistent with the coding rules. Second, it speeds up the task of coding by restricting the coder's search through the index owing to the smaller number of entries. Third, when the index is searched either by computer or by human it reduces the risk of finding matches with words that are irrelevant for the coding decision and therefore reduces error.

E.   Using the industry coding indices

1.   Using the industry response

697.   Coding can be seen as a process in which the task of the coder is to translate the information provided by the recorded responses to the appropriate code in the industrial classification structure. The main tools for the translation are the coding indices and the coding instructions—including instructions on when a response should be treated as a query to be resolved by supervisors or expert staff. The instructions should specify the following:

(*a*)   How the translation process should be carried out;

(*b*)   What items to look for in the industry response and in what order;

(*c*)   What type of ancillary information to use from other responses;

(*d*)   When such ancillary information can permissibly be used and how to use it.

Ideally, the coding indices should have been constructed to reflect and support the use of these instructions.

698.   The industry response is the information given in response to a question or questions about the name and address of the person's employer in a particular job, and the main type of industrial activity undertaken by the employer at that location. The starting point should always be the response given to the first part of the industry question(s)—the question(s) on the name and geographical location, for example, the street address of the place of work. If the name list provides an exact match on both name and location, then the industry code given to that unit in the name list can be given to the response. If there is no exact match, then the coder should make use of the regular coding list for

industry by selecting a word from the response that provides information about the type of products, services or function the unit produces or provides. If that procedure is not sufficient to determine a code, as with "advertising" in the example above, then the coder should identify supplementary words or qualifiers that may give more precise information about the product and/or the type of process involved. If the industry response does not contain sufficient information to determine a detailed industry category, then the coder should choose one of the following three alternatives:

(a) Look at the form for recorded ancillary information of a specified type for further clarification;

(b) Use an appropriate code for inadequately specified responses;

(c) Refer the case to supervisors as a query.

The coders should be given clear instructions on the proper alternative to choose.

2. Using ancillary information on occupation

699. Most modern industry classifications, including the International Standard Industrial Classification, Revision 4, are designed on the principle that "industry", meaning a particular set of productive activities resulting in one or more products or services produced by an economic unit, should be kept conceptually separate from "occupation", meaning the type of work performed by a person working in the establishment. Since many different occupations may be represented in the same establishment, one cannot normally draw valid conclusions about the industry of the workplace from one person's occupation, even if it happens to be an occupation that tends to cluster in a particular industry, such as bus drivers. However, there are a few exceptions to this, for example, "university teachers" are found only in the education industry and "taxi drivers" only in transportation. For some "own-account" workers there will be a direct link to a particular industry, for example, own-account plumbers should logically work only in the construction industry. Such cases can be identified and incorporated into the industry coding index.

700. Thus, since some interrelationships between occupation and industry are inherent in the industrial structure of the economy, they can be made use of to improve industry coding. However, there are costs as well as benefits in that practice. In the first place, there is always some danger that inferences from occupation to industry will be based on incorrect or out-of-date assumptions about the uniqueness of an occupation to a particular industry. In the second place, when coding is being done on a large scale, coding work rates and inter-coder consistency are important considerations. Coding rates are likely to be slowed if the coder routinely considers extra sources of information to arrive at a code, particularly if the extra information is itself hard to interpret. In such circumstances there is also a danger of increasing inter-coder variability, since different coders will tend to interpret the information in different ways. The latter two problems are minimized if:

(a) Occupation is coded in advance of or at the same time as industry, so that no further interpretation of verbatim responses is required for the industry code;

(b) Coders are allowed to use information on occupation only where the responses to the specific questions on name and address of place of work and activities are inadequate to determine an industry code;

(c) The choices that an industry coder can make on the basis of the occupation information are exhaustively specified through index-referenced instructions.

701. A simplified example may clarify the above. A coder encountering the establishment name "Institute for Marketing Studies" might be instructed, in the coding index under "institute/marketing", to choose between the code for "marketing" and the code for "education". If the occupation is given as "manager", "secretary" or "janitor" there

would be no basis in the occupation information to make the choice, as all three types of occupations may exist in either industry. However, if the occupation is given as "account executive" then it is likely that the establishment is a marketing firm (with a fancy name), and if the occupation is given as "professor" or "lecturer" then it is likely that the establishment is a training institution.

3.   Use of other ancillary information for industry coding

702.   Among the information usually collected on census questionnaires, only occupation is generally relevant as ancillary information for industry coding. The only exception is that in the case of one-company locations such as mining camps, where all employed residents are employed in the same establishment. In such cases, it may be possible to use information about the geographical location of the person's place of residence to assign an industry code.

4.   Inadequate industry responses and queries

703.   Some responses simply cannot be coded to a detailed industry category. That will normally be for one of the following reasons:

(*a*)   The response may be vague: it does not contain enough information to be coded according to the coding index and coding rules;

(*b*)   The response may be precise, but may use a title and/or indicate types or combinations of products, services or functions that do not correspond to any of the index entries.

704.   Unfortunately, the number of cases of type (*a*) is likely to be substantial, even with well-formulated industry questions and well-trained enumerators. To keep to manageable proportions the number of queries that the supervisors and expert coders must handle, the coding index and the coding instructions should be designed to guide the coders with respect to the most common of such cases. The simplest solution will be to specify that the response should be coded to a default category. The default category may in some cases be a specific detailed category since it reflects the dominant usage of the terms found in the response, such as "74.40 Advertising" as the default category in the example above from the United Kingdom. However, the default category will often have to be one of the aggregate categories in the classification, since it is not possible to identify one particular detailed category as dominant within the aggregate category indicated.

705.   As stated above, there is a real danger that coders may use default categories as dump categories for difficult-to-code responses before they have tried to find a precise code. Some countries have therefore tried to avoid their use by first line coders, preferring that such codes be used only by supervisors. However, that strategy may create a morale problem among the first line coders and place a very large query burden on the supervisors.

706.   The fully specified responses that are not adequately covered by the coding index should always be handled by expert coders and recorded carefully. That is done because it is important to ensure consistent treatment of equal cases and because those cases represent an important source of information for the updating of the coding index. During the coding operation, such cases can be handled either by using the priority rules specified for the classification or by assigning them to one or several categories for industrial activities "not adequately covered" by the classification.

707.   Priority rules are difficult to apply to the responses that indicate activity combinations that cut across the categories defined in the classification, for example, "repairing cars and selling petrol", since most industry classifications based on ISIC, Revi-

sions 3 and 4, will specify priority rules in terms of contribution to value added of the enterprise, or number of persons employed. That information will not be available to the census coders or their supervisors. One solution may be to refer such cases to the classification specialists, who may be able to identify the establishment and on that basis determine a code.

708.   Precise responses that cannot be resolved by priority rules should be placed in special "not adequately covered" categories created for coding purposes within the aggregate categories to which the activity clearly belongs. Steps should also be taken to ensure that those cases can be closely examined outside the coding operation itself to determine what contribution, if any, such cases can make to the updating of the coding index and of the classification itself. Note that the above categories are not the same as the "not elsewhere classified" categories of the classification, which include clearly defined activities that are not important enough in scale to be given a separate identity within the larger category to which they belong. Great care must be taken not to confuse the two types of category.

Part Six

Using population censuses to improve labour force and related statistics

Chapter XII

Types of data collection on the labour force and other economic characteristics

A. Organization of part six

709.   The previous parts of the *Handbook* have reviewed the ways in which a population census may collect information on the labour force and other economic characteristics, either through complete coverage or from a sample. As indicated, other sources of such information for national level statistics also exist, in particular household surveys, establishment surveys, economic and agricultural censuses and administrative sources. Part six is concerned with the ways in which the results from the population census may be used to develop, evaluate, improve and better utilize those other sources, specifically labour force and other household and establishment sample surveys. An important underlying issue is how the population census design and procedures may be chosen so as to maximize its usefulness for this purpose.

710.   Part six is divided into seven chapters. Chapter XII describes the diverse types of sources (sample surveys and administrative sources) that provide information on economic characteristics of the population. Chapter XIII presents a more detailed description of the main source of data on economic characteristics and activity of the population, namely the labour force survey. The remaining chapters describe ways in which the population census may be utilized to develop, evaluate and improve such surveys. Chapters XIV and XV aim to clarify the diverse manners in which sampling may be used in conjunction with the population census for surveys of households and establishments (including farms). They explain the concepts of sampling frames and master samples, two of the main uses of the population census, and describe how the population census provides the basis for constructing them. Chapters XVI to XVIII describe the use of population census results for weighting and preparing estimates in surveys, the combined uses of census and survey data and ways to ensure consistency between them, and the use of census information for the evaluation of content and coverage of post-censal sample surveys as well as for small area estimation.

B. Diversity of sources

711.   The number of questions used in a population census to investigate the various items has to be limited, often to only a single question per item (activity status, occupation, industry, status in employment, working time, income, sector of employment and place of work), even though actual practices vary significantly between countries. In contrast, household sample surveys, by virtue of their smaller size, can use more than one question for each item, allowing more precise measurement of a wider variety of data for different kinds of analyses.

712.   Apart from the population census, data on the labour force and its economic characteristics come from a variety of sources, as follows:

(a)   The main source is the labour force survey (LFS). The labour force survey is a household-based survey that is designed to provide estimates on the labour force and its characteristics. It is often the largest official survey, and it is conducted continuously or repeated at regular intervals. It is normally based on large probability samples, with wide coverage of the general population. Other household surveys may be based on the labour force survey, and sometimes it is designed so that modules on special subjects can be attached to the core labour force survey on an ad hoc or regular basis;

(b)   There are other population-based surveys providing information on economic characteristics and activity of the population, including income and expenditure surveys, fertility surveys, surveys of living conditions, longitudinal surveys and, more generally, surveys focused on special populations, particular economic categories or particular areas of the country. In such surveys, information on labour force and other economic characteristics is frequently collected to provide background variables or as part of a larger set of social variables;

(c)   There can also be large-scale, one-time census-like operations, covering variables that cannot and/or need not be collected on a census basis. Such surveys are normally closely linked to the population census in terms of objectives and operations, such as intercensal demographic surveys (sometimes referred to as sample censuses). They differ from typical sample surveys in that the infrastructure and impetus developed during the census permit exceptionally large sample sizes in the period closely following the census;

(d)   Information on economic characteristics of the population is also obtained from surveys for which establishments (rather than households) are the reporting units. In some countries, the coverage of units and workers in such surveys may be limited (for example, to large establishments and regular workers with written contracts), usually excluding workers in the informal sector. Sometimes those surveys also pose challenging sampling issues as the units involved may be small scale and numerous, scattered but unevenly distributed in the population, heterogeneous in type, and so on. Establishment-based surveys can obtain information on the economic characteristics of the population as their main variables or as background variables. Some examples of such surveys are production or industrial surveys;

(e)   Sources also include industrial and agricultural "censuses", which, despite the name, are often conducted on a sample basis—albeit on samples of large size.

713.   The International Conference of Labour Statisticians has stated the following:

> Population censuses and sample surveys of households or individuals generally constitute a comprehensive means of collection of data on the economically active population, which can be linked with data on other related topics. Establishment surveys and administrative records may also serve as sources for obtaining in some cases more precise, more frequent and more detailed statistics on particular components of the economically active population. The different sources of information should be regarded as complementary and may be used in combination for deriving, where necessary, integrated sets of statistics.[1]

In some countries, such integrated sets of statistics on employment are prepared as supplements to the national accounts.[2]

C.   Household-based enquiries

714.   Population censuses and household-based surveys cover, in principle, the same population and employ the same type of measurement units (households and individuals). The differences between the uses to which population censuses and household surveys may be put arise primarily from the differences in the scale of the operations involved

[1] International Labour Office (1983), resolution 1, para. 3.

[2] Some of the methodological issues that have to be addressed and some relevant national experiences are reviewed in E. Hoffmann (2000).

(complete enumeration versus sampling), which lead to differences in the methodology, practical conditions of implementation, timing and complexity of the data collected. As stated above, the number of questions used in a census to investigate any of the issues has to be limited, often to a single question per item. In contrast, a household labour force sample survey, by virtue of its smaller size, can be designed to obtain a wide variety of data for different kinds of analyses. It can be tailored more flexibly to fit a variety of user needs and methods of data collection. Household surveys, though by no means inexpensive, are less costly than complete censuses. They can be repeated more frequently and thus provide information on changes over a period of time. Owing to their smaller size, sample surveys may also permit better control of response and other non-sampling errors, and the results can be produced more speedily.

715.   Owing to the limited size of the samples, the major limitation of household surveys is their inability to provide sufficient detail for small areas or sub-groups in the population. In addition, samples of moderate size, while often capable of providing good estimates of proportions and other ratios, tend to be less satisfactory for estimating population aggregates (such as the total number of unemployed persons in a particular group) and changes in such aggregates, that is, for items that may be of particular interest to the user. To obtain reliable estimates of population aggregates, it is usually necessary to supplement or adjust data from samples by using information from other sources, often from the population census or population registers, where available.

716.   Mutually supportive roles and combined uses of population censuses and sample surveys have been described in many publications.[3] The sampling method can profitably be used to facilitate the planning, testing, controlling, evaluating, processing and supplementing of census data collection. In return, the census experience provides the infrastructure, sampling frame, benchmark statistics and so forth that are needed to conduct household sample surveys, together with a general impetus to the development of statistical capability. Those contributions are particularly important for large-scale sample surveys with wide coverage, such as national labour force surveys.

717.   Combinations of the two also exist. In many countries the population census has been designed to have two components: (*a*) an enumeration of the population and its basic demographic and related characteristics on a 100 per cent basis, supplemented by (*b*) a large sample, attached to the census, covering a broader range of items. Such a design can considerably enhance the role of the census as a source of statistics on the economically active population and related topics.

718.   It should further be noted that data from population censuses and household-based surveys are increasingly being used in conjunction with suitable statistical techniques to yield post-census and current estimates for local areas and small domains.

719.   Labour force and other household surveys allow for the joint measurement of the employed, unemployed and economically inactive. They can be designed to cover virtually the entire population of a country, all branches of economic activity, all sectors of the economy and all categories of workers, including own-account workers, contributing family workers and persons engaged in casual work or marginal economic activity. Such surveys therefore have a unique advantage for obtaining information on the total labour force and its structure. As concepts, definitions and subject details can easily be adapted to particular data requirements, different degrees of labour force attachment among various groups of the population can be measured. There is also considerable flexibility with regard to the data items that can be covered. Since in household surveys households or individuals are reached directly, relevant supplementary information on demographic and socio-economic characteristics of individuals and households can be obtained at relatively low additional cost along with information on labour force characteristics (see

3 See, for example, United Nations (1984), paras. 1.7-1.11, and Kish and Verma (1986).

United Nations, 1984, para. 11.10), offering many possibilities for data analysis. With appropriate design and rules of association, household-based surveys can also provide a means to collect information on household-based and other small-scale establishments.

D.   Establishment-based enquiries from information on the population's economic characteristics and activities

720.   In addition to the population census and labour force and other household surveys, the main sources of labour force data include (*a*) censuses and sample surveys of establishments; and (*b*) administrative records of different types. Those sources differ in coverage, scope, units of measurement and methods of data collection. Each has advantages and limitations in terms of the cost, quality and type of information yielded. Generally, one approach tends to be stronger where another is weaker, and vice versa. The various sources tend, therefore, to be complementary rather than competitive or mutually exclusive. Their results can be combined to some extent, depending on the degree to which concepts, definitions, coverage, reference periods, classifications and the like agree.

721.   Compared with household-based surveys, censuses and sample surveys that use establishments as reporting units exhibit the following characteristics:

(*a*)   They are generally more focused on fewer topics (and therefore more limited in content);

(*b*)   They provide estimates of regular paid employment, hours paid and earnings (but usually at aggregate levels), which are more consistent in time since the reporting by establishments is generally based on payroll records and does not depend on recall, as household-based estimates tend to do;

(*c*)   They provide more reliable estimates of industrial distributions, since establishments are better at describing what their activities are;

(*d*)   They are more economical in respect to collection since the number of sampling units is smaller and they may use postal methods and other methods that do not require face-to-face interviews.

However, establishment-based censuses and surveys (*a*) collect data on jobs, not persons; and (*b*) are usually more limited in industrial, establishment and worker coverage.

722.   In relation to coverage, two types of establishments may be distinguished: (*a*) establishments that by law are required to register with a government body (such as the ministry of industry and commerce or the local municipality), belong to the more organized sector of the economy and employ more than a certain number of persons; and (*b*) smaller establishments in the relatively unorganized sector, which may be either non-household establishments or purely household-based operations run by households on a proprietary or partnership basis. Many establishment surveys cover only establishments belonging to group (*a*) since official records, which constitute the easiest sampling frame to maintain, usually contain no information about establishments in group (*b*). An alternative sampling frame is the industrial census, which has full establishment coverage, but the high birth and death rates of small, unregistered establishments makes it quickly outdated. The employment statistics obtained from such surveys tend therefore to be restricted to regular paid employees in the more organized sector of the economy. Often, such surveys cover only those branches of economic activity that are of particular importance, such as mining and quarrying, manufacturing and construction. They are normally based on existing list frames, and tend to suffer from the more or less serious deficiencies of coverage common to such frames.

723.   Furthermore, rules of association between different types of survey units can be complex. Ideally, an establishment survey seeks information about establishments, specifically, an enterprise or part of an enterprise that engages in one, or predominantly one, kind of economic activity at or from one location or within one geographical area. However, as establishments provide information that is found in their registers, and as those registers respond to administrative needs related to their accounting and operational activities, it may be that some enterprises are not able to provide information separately for the various establishments that constitute it. Thus, the sampling units used for selection, the responding units that provide the information, and the units of enumeration and analysis on which information is sought in the survey may be different and difficult to associate. As an example, a gas station that also sells groceries constitutes two distinct establishments, but such enterprises are often unable to provide production and employment information separately. Thus, all workers engaged in that enterprise (and their corresponding sales and operation costs) will be assigned to the dominating activity, fuel or food. Another problem relating to the use of registers to report information is that on the one hand, payrolls may exclude persons with a job but temporarily away without pay and thus not report them among their employed staff. On the other hand, if a worker has two jobs in different establishments, she may appear on both payrolls and would thus be counted twice among the employed. Such problems do not arise in household surveys.

724.   The strength of establishment surveys lies in their greater specificity, in terms of both coverage and content. When the interest is in specific industries, establishment surveys, given an adequate sampling frame, can achieve a more efficient sample design (a smaller sampling error) than household surveys. More reliable and detailed information on certain topics that are by their nature sensitive, such as earnings by components, including bonuses, payments in kind and social security contributions, and hours paid for, can be obtained in establishment surveys, especially where the respondents can draw upon payroll and other available records. Other information is unique to establishment surveys, including labour turnover and labour cost. This type of survey can provide an opportunity to collect information on many other economic variables such as output, costs, investment and technological and organizational factors, which can then be directly related to information on employment, earnings and productivity and can form a much more comprehensive basis for the analysis of economic activity.

725.   Experience varies but establishment surveys are generally more economical than household surveys. There tend to be fewer respondents in establishment surveys, especially if only larger units are covered, and they tend to be congregated together and easier to contact. Moreover, cheaper methods of enumeration such as mailed questionnaires or telephone interviews can sometimes be used in place of the more expensive face-to-face interviewing. Another cost-reducing factor is that a single or a few respondents can provide summary information on all the persons employed in a large establishment.

726.   In the case of small establishments, the distinction between establishment and household surveys is less clear-cut. List frames are generally not available for small establishments, which are characterized by high fluctuation and often lack recognizable features. Therefore, the only feasible approach is the usual household survey, based on multi-stage sampling of areas with special listings of units at the last stage. However, unlike households, small establishments tend to be unevenly distributed geographically in the population, often in pockets of considerable concentration by type of economic activity. Information on the pattern of distribution from censuses or other sources is helpful and often necessary to improve the economy and efficiency of survey design and implementation (M. N. Murthy and A. S. Roy, 1970).

E.   Administrative sources

727.   Statistics based on administrative records are by-products of administrative processes. Consequently, administrative records can be a very economical source of statistical information. They are often based on continuous operations, and the statistics can therefore, in principle, be produced at almost any frequency or for any reference period. Other administrative records such as the payrolls and files of civil service organizations, government enterprises and other public institutions may also be usefully exploited to obtain information on some parts of the employment in the public sector. Where those sources are available and tabulated at frequent and regular intervals, they can be used to good account in particular analyses if careful attention is paid to their coverage and other limitations. Administrative sources can also be used in compiling and updating sampling frames for employment and related surveys based on samples of establishments, at least in respect of the larger establishments in the more organized sectors of the economy.

728.   However, administrative sources can also suffer from various shortcomings such as limited coverage and content, inflexible concepts and definitions, incompleteness, lack of timeliness, inconsistencies and restricted access due to legal or administrative constraints. In developing countries with unorganized labour markets, administrative sources such as unemployment insurance and employment exchange records often do not exist at all or are limited to certain narrowly defined categories of workers (see United Nations, 1984, paras. 1.13-1.14). More comprehensive remarks on the possibilities of using administrative records can be found in the 1997 report of the ILO East Asia Multidisciplinary Team[4] and in Pember (1998).

[4] ILO East Asia Multidisciplinary Team, "Labour statistics based on administrative records: guidelines on compilation and presentation" (Bangkok, ILO, 1997).

Chapter XIII

Labour force survey structure and arrangements

A. Purpose of the labour force survey

729.   To appreciate the relationship between the population census and sample surveys of households on economic characteristics of the population, and how the census may be designed and used to aid the latter, it is useful to describe the main features of the structure and arrangements of labour force surveys. That is the objective of chapter XIII.

730.   The labour force survey is meant to be a comprehensive source of information on economic activity of the entire population. According to the International Conference of Labour Statisticians, the programme of statistics of the economically active population should cover all branches of economic activity, all sectors of the economy and all categories of workers, and should be developed to the fullest extent possible in harmony with other economic and social statistics, as follows:

> The programme [of statistics of the economically active population] should specifically provide for both short-term and long-term needs, i.e., statistics for current purposes compiled frequently on a recurrent basis and statistics compiled at longer intervals for structural in-depth analysis and as benchmark data:
>
> (*a*)   The current statistics programme should encompass statistics of the currently active population and its components in such a way that trends and seasonal variations can be adequately monitored;
>
> (*b*)   The non-current statistics programme, which may include censuses and surveys, should provide:
>
> > (i)   Comprehensive data on the economically active population;
> >
> > (ii)   In-depth statistics on the activity pattern of the economically active population over the year and the relationships between employment, income and other social and economic characteristics;
> >
> > (iii)   Data on other particular topics (for example, children and youth, women, households) as determined by the long-term and continuing needs.[1]

B. Common characteristics of surveys of the economically active population

731.   National practices in conducting labour force surveys or, more generally, surveys of the economically active population, vary greatly, depending on specific data requirements and survey conditions and facilities. However, a number of commonly encountered features can be identified on the basis of available empirical information, as follows:

(*a*)   Labour force surveys tend to be relatively large-scale surveys of the whole population; they are often national in scope and have some sort of official status. Only the

[1] See International Labour Office (1983), para. 2. Minimum requirements concerning the frequency, item coverage and classifications of the statistics are set forth in Recommendation No. 170 concerning Labour Statistics, adopted by the International Labour Conference in 1985 to supplement Convention No. 160 concerning Labour Statistics.

national statistical office or some other major public agency engaged in statistical work can generally undertake such major operations. In many instances, the primary users of the information are also major public agencies;

(*b*)   Many countries undertake labour force surveys on a regular basis with the objective of measuring current levels of employment and unemployment for important groups in the population and the corresponding changes. Monthly or quarterly releases of survey results constitute the predominant pattern, particularly in developed countries, while in many other countries, such surveys are conducted at quarterly, yearly or less frequent intervals. A number of countries also undertake surveys that aim at providing more detailed information on the structure of the labour market. Such surveys are of longer-term interest, owing to their value in providing essential information for planning and policy formulation at the national level. Greater emphasis, however, needs to be put on the production of structural information of longer-term interest, as compared to the attention given to the production of statistics on current levels;[2]

(*c*)   In most countries, information pertaining to households and individuals in labour force surveys is collected through face-to-face interviewing by field staff visiting survey respondents in private households. Other methods of data collection, such as mail or telephone interviewing, are largely unfeasible in many countries, though telephone interviewing is used increasingly at least in some, mostly developed, countries for the second and subsequent interviews with the same household.

732.   The common characteristics of labour force surveys sketched above have a number of consequences for survey design and execution. First of all, owing to their official status and national scope, labour force surveys are likely to be subject to fairly stringent requirements of timing, data accuracy and internal consistency, especially consistency of the time series generated by regular surveys. Those requirements can be met only if the sample is representative of the whole population, not only geographically but also over a period of time owing to seasonal and other variations and the need to provide good estimates of changes. Thus, samples have to be selected randomly, be of fairly large size and be drawn from an up-to-date frame covering the whole population. The survey estimates have to be as consistent as possible with estimates obtained from other sources such as the population census, other household surveys, establishment surveys and administrative data. Many users require estimates of aggregates or population totals (as distinct from, and in addition to, estimates of means, proportions and rates), and estimates of changes in such aggregates. Chapter XVI provides more information on the use of population census benchmarks in ratio estimation.[3]

733.   Second, in many situations, the executing agency of a labour force survey is also involved in various other surveys and statistical operations. That increases the importance of coordination and integration in survey planning, design and execution.

734.   Third, where personal interviewing is involved—which is mostly the case for labour force surveys in developing countries and for at least the first enumeration in developed country surveys—the time and cost of travel for fieldwork are often a major component of the total survey cost. The sample therefore has to be selected in such a way that the cost of travel is minimized. That could be achieved by clustering households at the last sampling stage (through multi-stage area-based designs).

735.   Fourth, interviewing, coding and data entry operations may have to be decentralized, especially for large countries, resulting in a need for thorough interviewer training, supervision and quality control procedures and measures.

736.   At the same time, however, there can be a tendency to avoid experimentation and innovation owing to the necessary size, regularity and repetitive nature of the survey operations. Proper attention therefore needs to be paid to the continuous evaluation and

2 For a documentation of labour force surveys from about 90 countries, see ILO (2005).

3 See also para. 921 concerning problems relating to the back revision of population estimates and the consequent revision of survey estimates for compatibility.

periodic redesign of survey methods and procedures. It is by no means an easy task. The problem can be serious when, as is often the case, the available resources are limited and there is pressure to increase the quantity (variety, volume) of statistics routinely generated, at the expense of evaluating and improving the quality of the statistics. Rigid timetables for data collection can further accentuate the quality control problem.

C.   Diversity of labour force survey structure and designs

737.   Among household surveys of the economically active population, a variety of designs and arrangements are possible. The primary determining factors are the substantive objectives of the survey, that is to say, the content, complexity, timeliness and periodicity of the information sought. The survey may be designed to obtain regular time series on current levels and trends for a few basic characteristics of the labour force, such as the levels of employment and unemployment. Alternatively, or in addition, it may focus on less frequent information of a more structural nature and longer-term interest. Such substantive considerations will also determine the appropriate timing, frequency, reference period, sampling arrangements and other aspects of the survey structure. The requirement of appropriate linkages with other surveys, in terms of both subject matter and field operations, can be another important factor determining the survey structure and arrangements.[4]

738.   The following sections consider several dimensions of the diversity described above. First, there is a major distinction between regular versus occasional labour force surveys. Second, the labour force survey may be linked in different ways to other household surveys in the country, from a separate survey to a system of closely integrated surveys. There can also be other types of surveys, different from the typical labour force survey, providing information on economic characteristics and activity of the population.

D.   Survey frequency

1.   Regular surveys for current data

739.   Regular surveys are primarily conducted to generate a time series of statistics on current levels and trends. When the survey is designed to follow individuals from one survey round to the next, the survey may also provide estimates of gross changes and flows of individuals between different activity statuses and types of economic activity. Such statistics are yet to be published on a regular basis, even when the requisite information has been collected, owing to difficulties in separating observed changes from noise. Typically, such a survey consists of an ongoing series of survey rounds, each round being designed to produce separate estimates covering a specified time period. Regular surveys are used to monitor the performance of the economy; to obtain indicators of changes in current rates of labour force participation, employment, unemployment and underemployment; and to measure trends, as well as cyclic and seasonal variations in the rates.

740.   As indicated in paragraph 43, conducting a population census may disrupt a regular labour force survey programme. It is recommended that the programme not be interrupted if resources permit, to avoid any break in the time series of estimates from the surveys and resulting impact on many users. However, in some countries it may not be possible to conduct both the population census and the regular labour force survey during the same period.

1.1.   Regular surveys conducted on a continuous basis

741.   Among regular surveys, two types of arrangement are commonly found. The first arrangement is to conduct the survey on a continuous basis so that the fieldwork is car-

4 For further information, see R. Hussmanns, F. Mehran and V. Verma (1990).

ried out with no interruption. Typically, the information is obtained by using a moving reference period that relates to a period of a specified duration immediately prior to the interview, which varies among respondents depending on when the interview is held. If the survey round is of a sufficiently long duration, it may be divided into sub-rounds, each covering a representative sample during a part of the whole period for the survey round. Division into sub-rounds permits more representative coverage of the sample over shorter time segments during the round (such as over months or quarters during the year). Results from the sub-rounds can therefore be used to study seasonal and other variations during the round. The system also permits more frequent and timely release of the results.

742.   The distinction between rounds and sub-rounds is mainly one of sample size. In the case of each round, with a larger sample size obtained by the accumulation of sub-rounds, it may be possible to tabulate and analyse the data in fuller detail. In contrast, in the case of individual sub-rounds the sample size may only be large enough to provide main estimates with adequate precision, with less geographical and other disaggregation. With the fieldwork divided into representative sub-samples for each time segment (sub-round), the aggregated results for the whole round are themselves improved, since seasonal and other temporal variations during the round are covered or averaged out in a more balanced way. Division into sub-rounds can also have important operational advantages: the fieldwork can be better controlled and distributed more evenly over a period of time. In that system, the increased travel required to cover well-dispersed sampled households separately during each sub-round will represent an additional cost.

1.2.   Regular surveys conducted periodically

743.   The second arrangement is that of regular surveys with intermittent fieldwork concentrated over relatively short intervals. There can be certain advantages in concentrating fieldwork, as follows:

(*a*)   It can make it easier to control and implement field operations;

(*b*)   It becomes easier to obtain information with a fixed reference period, so that all respondents have the same reference period in terms of fixed calendar dates.

744.   While it may be considered necessary or even preferable for certain purposes to use such a fixed reference period, there can be some disadvantages in concentrating fieldwork, as follows:

(*a*)   The average conditions over a period such as a year may not be as well represented in the periodic surveys as in continuous surveys with fieldwork evenly distributed throughout;

(*b*)   Periodic surveys do not provide field enumerators with a continuous and evenly distributed workload.

745.   A choice has thus to be made between (*a*) employing permanent fieldworkers and letting them remain idle for part of the time; (*b*) resorting to the use of temporary staff during the periods of fieldwork; and (*c*) having a permanent field staff and using them for other survey operations during the slack period, such as sample updating, editing and coding, or for other statistical work.

746.   The first option, (*a*), is obviously wasteful, though apparently it has been seriously considered in some situations where the use of permanent, well-trained enumerators is considered essential to ensure data quality and where personnel field costs are relatively low owing to low wages. Option (*b*), the use of temporary staff, can provide more flexibility but can result in lower interviewing quality and hinder the development of the permanent survey capability of the organization. Option (*c*), combining field operations across surveys, is in common use, particularly in countries with relatively well-established

statistical infrastructures, where permanent field staff stationed in different parts of the country can handle a variety of data collection tasks, including household survey interviewing. In the intervals between periods of fieldwork for the labour force survey, the field staff can continue with data collection for other surveys. A distinct alternative would be to use the staff between rounds of the survey, not for other fieldwork and data collection, but for editing and coding the labour force survey data collected during the preceding round of fieldwork. That approach is becoming more feasible with the increasing decentralization of data processing operations and with the development of communications and computer technology. In some countries, the national statistical office has subcontracted the actual field operations and processing for the labour force survey to commercial survey organizations, subject to appropriate quality control and confidentiality mechanisms.

747.   Given that the main purpose of ongoing surveys, whether through continuous or periodic fieldwork, is to generate a regular sequence of data, it is imperative that in such surveys the results be released regularly and opportunely and that the volume and complexity of the information collected not overwhelm the organization's data processing and reporting capacity. Otherwise, unprocessed data will pile up in increasing quantities, and the whole objective of providing current statistics will be defeated.

2.   Occasional surveys for more structural information

748.   Comprehensive surveys of the economically active population may be conducted less frequently to obtain benchmark data and detailed structural information at the national level. Such ad hoc surveys are carried out in many countries that cannot afford to carry out regular surveys. Those surveys would obtain, for example, detailed information on the economically active population by industry, occupation and status in employment; on activity patterns over the year, work experience, multiple job-holding, education and training, hours worked and income from employment; and so on. Similarly, they can provide information on the population not economically active classified by type and by various socio-economic and demographic characteristics. In national development planning such surveys are needed for analyses of the employment conditions at the beginning of the planning period and for fixing targets and goals.

749.   While ad hoc surveys of the type described above are not designed to yield a continuous flow of current statistics or information on changes over short periods, they are well suited to provide less frequently needed information on essential structural characteristics and on longer-term changes, mostly at the national level. As those characteristics do not change rapidly, it is not necessary to undertake such surveys more than once every few years. For example, they may be conducted every five years as post-censal or intercensal surveys in countries with decennial population censuses. In any case, it is often simply not feasible to carry out such detailed surveys more frequently because of resource constraints.

750.   As with the timing of the census (see paras. 43-44), the timing of the survey has to be determined carefully. The results should be available when structural and benchmark data are most needed, as for example to provide a basis for the formulation of development plans. To be of long-term value, survey timing should not coincide with periods of abnormal or transient employment conditions unless the measurement of that abnormality itself constitutes a main objective of undertaking a special survey. The survey period should also take into account seasonal and other short-term variations. For the preceding reasons, it may be useful to spread out fieldwork over a whole year, covering all seasons; it may be done on a continuous basis with uninterrupted fieldwork throughout the year, or in the form of more concentrated periods of fieldwork spaced out over the year.

E.   Linkages of the labour force survey with other surveys

751.   In many countries, household surveys are increasingly used as a source of a wide variety of statistical information. As a consequence, it is often necessary to undertake surveys, whether on the economically active population or on other topics, as part of a common survey system or programme, or at least to ensure that individual surveys are closely coordinated. Those requirements can be particularly important in the case of surveys on the economically active population, which often tend to be comprehensive in coverage, national in scope and relatively large in size. The need for linkages becomes even greater when a survey involves regular operations repeated periodically or continuously, for they can greatly affect and be affected by other operations in which the statistical organization is simultaneously involved.

752.   Linkages between surveys involve the following two broad aspects:

(*a*)   Coordination at the design and operational level, where common procedures, arrangements and facilities are used to increase the economy and flexibility of the operations;

(*b*)   Integration at the substantive or subject-matter level, where a number of topics are covered in conjunction with each other, using similar definitions and classifications, to permit the production of interrelated statistics that can be analysed jointly.

753.   Studies of linkages between censuses and surveys can also be an important and useful way of studying non-response bias in surveys. In those studies, survey records are matched with census records so that the characteristics of survey non-responders can be analysed using census data (see also chap. XVII).

754.   "Coordination" implies that individual surveys are designed and undertaken in proper operational relationships to one another, utilizing common procedures and infrastructures, including organizational arrangements, sampling frames and other materials, technical and supervisory staff, field and office personnel, as well as transport, data processing, printing and other facilities. The degree of coordination and sharing of facilities may vary, depending on the type of organization involved, the nature of its operations, special requirements, funding and other arrangements for the surveys, and so on. While smaller one-time surveys can often be carried out on the basis of more or less special or ad hoc arrangements, it is much less likely to be the case for such major undertakings as national labour force surveys. Often, labour force surveys are carried out as part of the regular operations of a national statistical agency, and their planning requires careful consideration of operational links with other undertakings, an evaluation of the possible constraints and problems (such as increased pressure on available data processing facilities and technical staff) and as of the opportunities and flexibility which those linkages can offer.

755.   "Integration" at the substantive level implies the use of a common study population; common concepts and definitions; a common system of classifications, and possibly standard survey questions for such frequently used classifiers as age, sex, ethnic group, education and economic activity status; and common or overlapping samples of respondents. Sometimes the term "complete integration" is used to indicate coverage of multiple topics in a single survey, over a common sample, and possibly during a single interview with the respondent. By contrast, the term "partial integration" implies a situation where the various topics are covered using the same sample of areas but with different samples of households within each area. Only complete integration permits data linkages at the microlevel.

756.   In practice, various patterns can be found in the manner and degree of linkages of labour force surveys with surveys on other topics, including the following:

(*a*)   A labour force survey may be organized as an operation more or less separate from other surveys;

(*b*)   A more comprehensive labour force survey may serve as a vehicle for covering other related topics as well, in the form of modules attached to it;

(*c*)   A labour force survey with limited content may be incorporated into some other ongoing survey as a module;

(*d*)   A labour force survey may form part of a multi-purpose survey covering a range of topics, or of surveys specially designed to enumerate population groups of special interest;

(*e*)   A labour force survey may be undertaken as one round of an ongoing survey system that focuses on different topics in different rounds.

1.   Separate labour force surveys

757.   A number of developing and developed countries carry out surveys that are primarily or exclusively concerned with labour force topics. In the present context, they may best be described as "separate" labour force surveys. Their single-subject focus does not preclude operational coordination and the use of common facilities and arrangements with other surveys or the use of common coverage, concepts, definitions and classifications. What is implied by "separate" is their single-subject focus and a considerable degree of separation in design and execution. Such separation can sometimes be helpful in providing better control and supervision and greater flexibility in the design and operation of the survey.

2.   Labour force survey as a vehicle for other data

758.   The establishment of a regular labour force survey can be a major and relatively expensive undertaking. Once in place, the survey can be usefully exploited as a vehicle for covering additional topics and for supporting various household surveys in other areas. Indeed, diverse household surveys may be developed using the labour force survey capacity and be integrated with it in design and operations to varying degrees, and conducted as supplements to the labour force survey. As an alternative, or in addition, more independent surveys may use the same sample areas as the labour force survey, but different sets of households and different survey periods.

3.   Labour force module attached to other surveys

759.   In contrast to the above, a labour force survey with limited content may itself be attached as a module to surveys focused on other topics. It is most effective when the survey serving as the base for the labour force survey is a regular survey, so that it can be used for generating a regular time series of labour force statistics. Inclusion of basic items in other ongoing surveys, where they are useful as background variables for the main items of the surveys, can be an extremely economical way of obtaining some essential information on the labour force and its characteristics. Using the module can be particularly helpful when a full-fledged labour force survey cannot be undertaken owing to limited resources or competing priorities. A module on labour force items may be included in other surveys on an occasional basis, or on a more sustained basis, depending on the requirements.

760.   However, it is also necessary to recognize some limitations of the approach. First, there are limits to the number and detail of labour force items that can reasonably be inserted into operations concerned primarily with other topics. In a population census, for example, each of the few labour force items that may be included generally has to be

confined to a single simple question, or to a simple sequence, similar to that used in the previous population census. Care must be taken to ensure that such additions do not adversely affect the overall quality of the information obtained in a census or survey due to excessive respondent burden, delays in data processing or other consequences of the increased size and complexity of the operation.

761.   Second, to ensure data quality and usefulness of the results, it is necessary that the various topics included in the same survey are compatible in terms of concepts, definitions, survey methods, reference periods, coverage and other design requirements. It is not always possible or easy to achieve such compatibility. At the same time, the requirement of compatibility with other topics covered in the same survey may itself limit the type of information that can be requested in the labour force module. Furthermore, the surveys may be confined to specific groups in the population, so that the labour force data can be used only as explanatory variables for other topics.

4.   Omnibus multi-purpose and specialized surveys

762.   Integration can also take a more extreme form in which a large number of detailed topics are combined in a single omnibus undertaking. In principle, the main advantages of comprehensive multi-purpose surveys are the possible economies of scale and the potential for cross-checking and for combined analysis of detailed data on various topics. Multi-purpose surveys can thus yield a wealth of information on labour force variables in conjunction with other topics, provided that the samples are sufficiently large, given that they have to be able to provide precise estimates for a multitude of variables. However, complex multi-subject surveys can, and often do, suffer from serious disadvantages, especially in the more difficult circumstances of some developing countries. Such disadvantages include the increased length and complexity of the interview, increased respondent burden, possible increase in non-sampling errors, reduced efficiency of the questionnaire design, which needs to accommodate diverse requirements, and of the sample design, which needs to provide precise estimates for many topics; and the danger of delays and failures at the data processing stage owing to the increased volume and complexity of the data collected (see United Nations, 1984, paras. 1.20-1.28). Some of those problems may be reduced in scope, for example, by using different sub-samples for different sets of topics in addition to certain core topics obtained from the overall sample (in which case, not all the topics can be linked at the microlevel), or by organizing data processing separately for different sets of topics. Nevertheless, caution is needed to avoid making any survey system too complex or overburdened.

763.   Occasionally, more specialized surveys may be undertaken to investigate in depth certain relationships, special phenomena or population groups of special interest. Possible examples are relationships between labour input, training and experience on the one hand, and income from employment, family income, welfare and the like on the other. However, such surveys tend to be fairly complex in content and involve special arrangements, specialized staff and relatively heavy cost and effort. Consequently, they are usually undertaken on a one-time or infrequent basis, often with less than national coverage and with smaller and less well-dispersed samples. Insofar as a specialized survey is research oriented (that is to say, it aims at providing information of longer-term interest to enhance the general understanding of issues and problems, rather than at meeting some more immediate data needs), there may be considerable flexibility in its timing. For the same reason, comprehensiveness of content and high quality of data may be more important considerations than extensive coverage and quick release of data.

764.   Sometimes, survey objectives require the collection of additional information on particular population groups of special interest such as the handicapped, migrants,

female household heads, unemployed young persons or underemployed workers. Where such groups are small, special arrangements such as multiphase sampling with screening may be required to include sufficient numbers of respondents. In regular surveys, there is also the possibility of accumulating such cases from several rounds. If the additional information required on groups of special interest is too detailed or complex, it may be necessary to organize its collection as an operation separate from the main labour force survey (employing for instance special questionnaires and/or special interviewers), though coordinated with it in an appropriate manner.

765.   Population censuses, demographic surveys, household income and expenditure surveys and other surveys often collect basic information on the economic characteristics and activities of the population. Often the objective is to provide explanatory variables that are useful for cross-classification and analysis of the characteristics that are the primary focus of the survey. Although in that case the survey is not aimed at providing estimates for labour force variables per se, it can be useful in enhancing the understanding of relationships between labour force and other characteristics such as fertility, childcare, health, income, consumption behaviour and so on. In fact, some of those topics are so closely related to labour force characteristics that any survey on them may require a fairly comprehensive coverage of the labour force characteristics as well. For example, a survey on migration may need to include items such as activity status and status in employment, occupation, industry, institutional sector of employment and income to help understand the causes and consequences of migration.

5.   Labour force survey as a periodic round of a survey system with varying focus

766.   Another pattern of integration, which may be particularly suited to conditions and requirements in some developing countries, is to establish an ongoing survey system with a varying substantive focus from round to round. Each round may cover a specified period such as a year and a separate representative sample. The survey system can use common organization, personnel and other facilities, but the subject matter changes from round to round, with the possible exception of some core items common to all rounds. Among the common elements are likely to be some on the employment situation of the respondents, in addition to the basic demographic variables. Such a system can periodically include comprehensive surveys of the economically active population as the main focus of the rounds. The comprehensive survey can provide structural or in-depth information of longer-term interest. Such an arrangement has many potential advantages for survey work in developing countries,[5] and is a model that was promoted through the National Household Survey Capability Programme.[6]

[5] The National Sample Survey of India provides a well-known example of such a survey system; see, for example, Murthy and Roy (1970 and 1975).

[6] United Nations, "National Household Survey Capability Programme: prospectus", DP/UN/ INT-79-020/1 (1980).

Chapter XIV

Sampling and related uses of census information

A.   Sampling in conjunction with the population census

767.   In most countries today the decennial census of population is the primary source of geographically detailed information on the basic demographic, economic and related characteristics of the population. Many official statisticians agree that a population census need not gather all demographic and housing information on a 100 per cent basis. The issue of the use of sampling in conjunction with the population census is pertinent in the context of the development of a system of labour force surveys. The labour force survey is often the largest and most important population-based survey undertaken regularly in many countries.

768.   According to the United Nations, the use of sampling actually saves a good deal of time and money; and, furthermore, in certain circumstances, only the sampling approach ensures data of acceptable accuracy (see United Nations, 1971; and 2008b, chap. IV). The practice of governments is increasingly to collect information on the size, age and sex composition, geographical distribution, and certain other basic demographic and socio-economic characteristics of their populations on the basis of a complete (100 per cent) enumeration, and to supplement that basic information by collecting information on a larger range of variables on a sample basis. The additional information, when it is gathered as part of the census operation, may be related to employment, income, migration, mortality, fertility and health.

769.   Sampling may be used in conjunction with the population census in diverse ways and for different purposes, such as in census operations (planning, design, testing, evaluating, tabulating), and for additional data collection. Additional data may be obtained in the census in a number of ways, including (*a*) from enumeration of more detailed items on a sample basis as part of the census; (*b*) through sample surveys attached to the census or through large-scale surveys closely following the census; (*c*) through the development of a system of sample surveys based on the census frame for sample design and selection; and (*d*) by using the census data in the production of improved estimates from the surveys. By combining the census with a sample survey it is possible to exploit the advantages of each.

770.   Combining data collection on a 100 per cent and a sample basis during the census is just one way (albeit the most important) that sampling is used in conjunction with the census. Complete census and sampling can be used in combination or in related ways with the objective of capturing the advantages of each (see L. Kish and V. Verma, 1986; M. N. Murthy, 1980; and V. Verma, 1989). Those applications include the following:

(*a*)   Using sampling in the design and control of census operations, such as in planning, testing, controlling and evaluating the census;

(*b*)   Using sample enumeration to supplement the items covered in the complete census. In certain circumstances it is also possible to consider substituting complete enumeration with one or a series of sample enumerations;

(*c*)   Sampling the census results for processing, with the objective of making the results available more quickly and at lower cost; extracting samples of microdata files of detailed census data so as to facilitate dissemination of primary data more widely for analysis by other users;

(*d*)   Using the census as a basis for sample surveys by enhancing its statistical capability and resources and by providing baseline data, population controls for estimation and sampling frames for surveys in the post-censal period;

(*e*)   Using the census and post-censal sample survey data in combination to provide estimates for local areas and small domains.

B.   Census as a source for sampling frames

771.   In most countries, sampling frames for household surveys are based on the most recent available census of population, though the census information is often supplemented or updated from other sources for this purpose. It is much less common to have alternative sources entirely independent of the population census.

772.   While the census of population is obviously the primary source of sampling frames for household surveys, with foresight and careful planning it can also serve to create area frames for surveys and censuses of establishments, particularly small-scale establishments, in the agricultural and non-agricultural sectors.

1.   Survey population

773.   The definition of the population to which the sample results are to be generalized is a fundamental aspect of survey planning and design. While basic decisions about the nature and scope of the population to be covered are taken early in the survey planning process, the content and extent of the population has to be specified more precisely at the stage of technical design. This specification is in terms of the following:

(*a*)   Population content: the definition of the type and characteristics of the elementary units comprising it;

(*b*)   Population extent in space: the boundaries of its geographical coverage;

(*c*)   Its extent in time: the time period to which it refers.

774.   For example, in the labour force and other household surveys, the elementary or analysis units are generally households and persons. Many surveys are aimed at covering the general population, but may nevertheless exclude, for substantive or practical reasons, certain types of households or persons, such as collective, institutional, foreign, diplomatic or non-civilian households; homeless persons or families; persons temporarily abroad; or persons above or below certain age limits. Various considerations may also result in the exclusion of certain areas from geographical coverage of the survey, for example, remote and sparsely populated areas in many countries. Units and their characteristics change with time, and the population covered in any survey is also bounded in time, even though those boundaries may not be as sharply defined as geographical boundaries.

775.   Population content and its extent are often defined more narrowly in surveys and censuses in the non-household sector. For example, many surveys of establishments cover only those units that, in relation to a certain specified period in the past, existed and satisfied certain criteria in terms of the size and type of activity (sector or industry) of their operation. In agricultural censuses and surveys as well, it is common to limit the coverage to holdings that meet certain minimum size criteria, which may differ according to the type of activity.

776.   Three important practical points should be noted. First, in any survey, rules of population inclusion and exclusion must be defined in clear operational terms. Otherwise confusion and errors result at the implementation stage. Second, the limitations in the population covered must be kept in view in drawing inferences from the survey results, and in comparing results from different sources. It is important to document the extent of exclusions, since they define the limits within which inferences can be scientifically drawn from the survey. The effect of exclusions depends on the type of statistics produced. For instance, exclusions can affect estimates of population aggregates more seriously than rates and ratios. The latter can, however, still be affected to the extent exclusions are selective with respect to the characteristics measured. Third, it is important to reassess periodically the assumptions behind excluding certain parts of the population that should ideally be covered and seek clear justification for the continued exclusion in each case. Sometimes, it may be possible to adopt special arrangements or take reduced samples from the more difficult and expensive parts of the target population, rather than exclude them altogether.

777.   Apart from deliberate and explicit exclusions, surveys also suffer from coverage errors, which are less easily identified and measured. Painstaking work is usually required to control those errors and assess their effect on the survey results. Their magnitude depends on the quality of the sampling frame and sample implementation.

2.   Sampling frame

778.   The population to be surveyed has to be represented in a physical form from which samples of the required type can be selected. A sampling frame is such a representation. In the simplest case, the frame is simply an explicit list of all units in the population, from which a sample of the units concerned can be selected directly. With more complex designs, the representation in the frame may be partly implicit, but there is still a need to account for all the units.

779.   In practice, the required frame is defined in relation to the required structure of the sample and the procedure for selecting it. Hence, to understand the concept and requirements of the sampling frame, it is necessary to understand some basic aspects of the structure of the samples as typically encountered in large-scale household surveys. The simplest, but generally quite impractical, design is to select a simple random sample of the elements of interest such as households or individual persons. In such a design, each element and each combination of elements in the population has the same constant probability of appearing in the sample. The most common aspects in which samples in practice depart from simple random sampling are stratification (when the population is divided into non-overlapping relatively large groups such that the groups are homogeneous with respect to one or more characteristics and then a random sample of units is selected within each group); cluster sampling (when the population is divided into non-overlapping relatively small groups and then a random sample of groups is selected); and, in some situations, selection of the elements with unequal probabilities. The most popular method is multi-stage sampling. In this approach, the sample is selected by stages: the sampling units at each stage are sub-sampled from the larger units chosen at the previous stage, with the final units being the households or persons.

780.   A basic distinction should be made between list frames, from which the units of interest can be directly selected, and the more general multi-stage frames, which are usually area-based. In the latter, the frame for the first stage of sampling (called the primary sampling frame) is a list of geographical areas, usually drawn from the census, and it covers the entire population. Following the first stage of selection, a list frame of units at any lower stage is required only within the larger units selected at the preceding stage. Area frames are generally preferred when the list frame is drawn mainly from population censuses, as under-

enumeration of units is very likely: there may be persons or households that exist but were overlooked when compiling the list frame, for example, because they were not enumerated in the population census. Area frames guarantee, if they are exhaustive of the country, that all units found in the areas selected in the first stage of sampling will be covered.

781.   Several variations and combinations are possible. A frame may be constructed from a single source or may have to be compiled by combining information from a number of sources. Different types and/or sources of frames may be used for different parts of the population. It is also possible to use more than one frame in combination to represent the same population more adequately.

2.1. Area-based frames

782.   In the context of household surveys, especially in developing countries, the frame generally consists of one or more stages of area units, followed by lists of households or dwellings within the selected ultimate area units. It is useful to distinguish three components of such a multi-stage frame, as follows:

(*a*)   The primary sampling frame, which is a frame of the first or primary sampling units and must cover the entire population exhaustively and without overlaps. Primary sampling frames are usually geographical or administrative areas in the country. Following the first stage of selection, the list of units at any lower stage is required only within the larger units selected at the preceding stage;

(*b*)   Possibly, a hierarchy of secondary area-based frames, consisting at each stage of units selected at the preceding stage, until a frame of the lowest or ultimate area units is obtained. Below the ultimate area units, the sampling process moves from areas to the listing and selection of individual dwellings, households or persons;

(*c*)   Explicit lists of the ultimate sampling units such as dwellings or households within the selected ultimate area units. The elements for data collection and analysis in the survey may be the ultimate sampling units themselves, or may be other units uniquely identifiable from the ultimate sampling units through definite rules of association. For example, persons (elements) may be associated with selected households (ultimate area units) on the basis of a "population present" or "usual resident population" coverage definition.

783.   The three components have major differences in terms of the cost and durability of the frame. The durability of the frame declines in moving down the hierarchy of the units. The lists of ultimate units require frequent updating. In most surveys, it is necessary to prepare fresh lists of ultimate sampling units shortly before the survey enumeration. It is a major advantage in a survey if it can utilize lists prepared for some other recent survey or census. Lists of structural units such as dwellings are usually more durable than lists of social units such as households; pre-existing lists of individual persons are rarely useful.

784.   By contrast, the primary sampling frame (and to a lesser extent, frames of intermediate level units) usually represents a major investment for long-term use. A most important decision in the development of the primary sampling frame concerns the choice of the type of units to be used as the primary sampling units. Several practical considerations may be noted. Larger areas generally provide more stable and clearly demarcated units, but that choice can also increase the segmentation, listing and other work required to complete lower stages of the frame construction and sample selection. The main source of the primary sampling frame is the most recent census of population.

785.   Similar considerations apply also to secondary frames of area units, but to a much lesser extent. For one thing, secondary frames are required only for the primary sampling units selected in the first stage, and therefore it is not impossible to consider their construction for use over a limited duration, or even for the purpose of a single survey. In any case,

many surveys are based on a design involving a single area stage (for example, a sample of census enumeration areas as the primary sampling units, followed by direct listing and selection of households within selected enumeration areas), and no secondary area frames are involved except on a selective basis where particularly large original primary sampling units need to be segmented.

786.   Area-based frames are also used for surveys of small-scale agricultural and non-agricultural establishments, which are too numerous and dispersed to be covered by list frames and for which no registers or lists are available in most cases. Special procedures are required in constructing such frames since the units to be covered are often very unevenly distributed in the population. Compared with households in the general population, they also tend to be more heterogeneous in size and characteristics and are divided into distinct types, each requiring separate coverage.

2.2.   Census enumeration areas as the primary sampling units

787.   Careful consideration must be given during the planning and execution of the census to its function as a source of area sampling frames for diverse surveys. The enumeration areas of the census have the following multiple functions:

    (*a*)   To partition the population into geographical areas with clear, stable, and identifiable boundaries, which can be mapped and described;

    (*b*)   To facilitate complete and unique coverage of the units in the population; to create reasonably equitable and feasible workloads;

    (*c*)   To facilitate the organization and control of census operations; to provide a flexible basis for the production of areal statistics at various levels and types of aggregation; and to provide a basis for scientific and efficient sample selection for subsequent surveys.

788.   Although the preceding requirements cannot all be satisfied simultaneously to the same degree, they indicate desirable characteristics of census enumeration areas.

789.   Enumeration areas should be small (containing a few hundred people on the average) as well as reasonably uniform in population size. However, the requirement of clear boundaries is more important than uniformity in size. Enumeration areas should be real units covering the entire country exhaustively and should be mapped and described for clear identification of boundaries; they should not usually cut across administrative subdivisions or natural boundaries. They should be geographically ordered with proper identification systems, to facilitate the production of results at different levels of aggregation and for different types of geographical units, for example, administrative, commuting and water catchments. Information on their size and other basic characteristics should be collected, coded and tabulated for individual areas.

790.   The census generally provides sampling frames or master samples of areal units for subsequent surveys. Lists of housing units, households, and individuals from the census are usually not allowed outside the organization responsible for the census owing to confidentiality concerns. For many countries, such lists are also too difficult to arrange, and individual households are too mobile to be used for subsequent surveys, except perhaps in surveys attached to or conducted soon after the census. In some countries, the lists are generated and found to be useful.

2.3.   List frames

791.   Completeness of the frame is the most critical requirement of list frames. It is also important that the list contain pertinent and accurate information on the size and other characteristics of individual units so as to permit efficient stratification and control of the selection process.

792.   When the units to be surveyed are relatively large in size and limited in number, they are best selected in a single stage from a list frame. Often good lists exist or can be compiled with reasonable effort for large units. However, moving to smaller and more numerous units generally increases coverage and other problems, especially when there is no system for registration of the units or for updating existing lists. Several situations may be distinguished for separate treatment, depending on the size, number and distribution of the units involved, including the following:

(*a*)   Large units that are few in numbers for which good lists may be available;

(*b*)   Medium-sized units, which can also be covered by only a list frame, but for which such frames are more difficult to construct and maintain;

(*c*)   Medium to small units that may require a combination of list and area-based frames;

(*d*)   Small or very small units that can only be covered with area-based frames.

793.   The last category may itself require differentiation into several types. For instance:

(*a*)   Numerous and relatively well-dispersed units may be adequately covered by an area-based frame of the general population;

(*b*)   Relatively numerous but unevenly distributed units would require specially constructed area frames, taking account of the patterns of concentration;

(*c*)   Some special populations (such as fishing households or others engaged in very specialized activity) are confined to pockets of concentration, which may be numerous but vary greatly in size. They are not adequately captured by the general frame, and it can be extremely difficult to identify and list them all. The best that can be done in some circumstances is to concentrate the effort on covering the larger and more important concentrations, with an attempt to improve the coverage over time, as knowledge on the patterns of distribution of the units accumulates;

(*d*)   Rare but dispersed populations also require special methods and arrangements, as do mobile populations.

2.4.   Multiple frames

794.   When two or more overlapping frames are used simultaneously, it is important to ensure that the probabilities of selection of the units remain definite and known. Various approaches are possible in selecting samples from multiple frames. Perhaps the simplest option is to make the frames non-overlapping, if possible. That may be done by eliminating duplicate listings of any unit from all but one frame. Another way is to impose an order of priority among the frames such that any unit can be selected from only one of the frames. For instance, if a list and an area frame are used in combination, any unit selected from the latter may be automatically excluded if it is also present (whether or not selected) in the list; in other words, any unit present in the list frame is taken as eliminated from (treated as a "blank" in) the area frame.

3.   Common problems with sampling frames

3.1.   Problems with area frames

795.   Area frames have the advantage over list frames in that areas as units are larger, more easily identified and more stable than dwellings, households, establishments or persons which appear as units in typical lists. Nevertheless, area-based frames also suffer from coverage and related errors. The errors usually arise from failures to define and identify the physical boundaries of the area units correctly and from the poor quality of the lists of the ultimate units.

796.   For population-based surveys, it is usually possible to assess whether large under-coverage has occurred (which appears to be an extremely common problem) by comparing estimates of population aggregates obtained directly from the survey (the so-called "simple unbiased estimates") with some more reliable external information on population size (for example, demographic estimates based on birth and death administrative statistics). For establishment-based surveys, however, external information of good quality about the number of establishments may not always be available.

797.   Common imperfections of area frames include the following:

(*a*)   **The failure to cover the population of interest exhaustively.** In a number of countries with inadequate cartographic work, the available frames are actually composed of lists of localities rather than of proper areal units; scattered populations outside the listed localities may not be covered in such frames;

(*b*)   **Under-coverage due to outdated frame of area units.** The problem of under-coverage can become much more serious as the frame becomes outdated with time.

(*c*)   **Errors and changes in area boundaries.** They may arise from errors in identification of the boundaries and from boundary changes after the frame was prepared. The unit boundaries as defined in maps or descriptions may differ from the boundaries of the units for which other relevant information is available in the frame (information such as the size and density of the population) or from the boundaries of the actual sampling units;

(*d*)   **Inappropriate type and size of units.** The available units may be too large, too small or too variable in size to serve as efficient sampling units;

(*e*)   **Lack of auxiliary information.** Information on size and other characteristics of the units, required for efficient sample selection, may be inaccurate or simply unavailable. Area frames for household surveys based on the population census often suffer from outdated information on the population size of the area units. The requirements and problems of auxiliary information can be much more serious in the case of area frames for surveys of small-scale establishments and other economic units;

(*f*)   **High cost.** Area frames are generally expensive to create. Usually, the investment is justifiable only when the frame is to be used repeatedly over a relatively long period of time.

798.   Serious coverage errors occur in the use of area frames when the boundaries of the enumeration areas cannot be clearly identified during the fieldwork for the survey. The problem is very common in surveys based on outdated area frames, or when the available maps and descriptions are of inadequate quality. To understand the problem one needs to look at the precise way in which enumeration areas are identified. The following is an illustration of what may happen in practice; in theory, an enumeration area is identified simply by delineating its boundaries on a map. Those boundaries are supposed to follow natural features such as streams or roads. However, such features are not always available, so the mapper often draws straight lines on the map that have no basis on the ground. He or she may very well draw the line through an inhabited area, reasoning that since no one appears to live there, it does not really matter exactly where the boundary lies. For greater clarity, the mapper is normally instructed to insert on the map the location and name of all settlements (rural) or streets (urban) lying within the enumeration area. The household listers for the survey, who may arrive on the scene a few years later, soon learn that the map boundaries, often straight lines, cannot be located on the ground, and so turn their attention to the names written on the map. Very quickly they come to see the job as that of listing the households in the named places and streets. Clearly, wherever the population is growing, that practice will lead to error. New settlements and new streets will tend to be omitted by the listers. What can be done to reduce this problem?

799.   First of all, it can be helpful to try to isolate the problem. It often happens that the problem with area boundaries is serious only in some domains. For instance, census enumeration areas in rural areas may correspond to individual localities, and it may therefore be relatively easy to allocate them to the correct areal unit even when the precise boundaries are not easily identifiable; in the other domains, the problem may exist only in the more densely populated (and more rapidly changing) urban areas.

800.   Second, it can help to draw on all the information available, not just the maps. Often census maps are accompanied by helpful descriptions. More importantly, census household lists, even if outdated for actually selecting a sample of households, can be extremely useful in defining and identifying the original boundaries of census areas.

801.   Next, it is possible to look for different, larger, area units as primary sampling units in place of individual enumeration areas, units that are likely to have less serious boundary problems. From a given frame of basic area units such as enumeration areas, larger units can be created in two ways: by selecting enumeration areas in contiguous groups rather than individually from the list of enumeration areas; or by introducing a higher stage of selection, such as taking whole towns or villages, each composed of a number of enumeration areas, as the primary sampling units. The larger areas can then be sub-sampled as appropriate. For instance, the selected units may be segmented and one or more segments selected per area in an additional sampling stage, or the whole area may be listed for the direct selection of households.

3.2.   Problems with list frames

802.   As stated above, completeness of the frame is the most critical requirement (and perhaps also the most common problem) of list frames. In addition, to permit efficient stratification and control of the selection process, the list should contain pertinent and accurate information on the size and other characteristics of individual units.

803.   Problems can arise in the absence of a one-to-one correspondence between the listings (which are the units actually subject to the selection process) and the elementary units (the actual objective is to obtain a probability sample of these units). The lack of correspondence can manifest itself in a number of different ways, as follows:

(*a*)   Presence of blanks in the list, meaning that some listings represent no real units;

(*b*)   Clustering of elements, meaning that the same listing may represent more than one unit;

(*c*)   Duplications, meaning that more than one listing represent the same unit. Sometimes the problem arises from the nature of the frame: as for example in the selection of households from an electoral roll (listing all eligible voters in each household) or from telephone directories; selecting parents from a list of school children; or selecting clients or patients from records of visits to the facility. Much more difficult is the problem of unsystematic duplications in the list, usually resulting from the failure to identify the fact that different listings actually represent the same unit. This can happen, for example, if the same unit is recorded in the list several times with slight differences in name, address or description. In such cases painstaking work to eliminate all duplications in the list may be the only solution;

(*d*)   Failure to locate units, in other words, the failure to identify which unit(s) a selected listing represents. It is a common problem in the absence of a clear and complete description in the frame for identifying units in the field. It can also be caused by insufficient effort during survey fieldwork. The problem is often confused with that of blanks—units not located being indiscriminately reported as non-existing—which, among other things, causes difficulty in correctly computing the response rates actually achieved;

(*e*)   Change in units and unit characteristics, that is, the unit itself or the characteristics of the unit associated with the listing have changed. To overcome the problem resulting from mobility of the ultimate units, many surveys use the "population present" coverage definition. That means taking the sample selected to be a sample of addresses or locations, rather than of particular households or establishments; whoever is found to be present at the selected location is enumerated in the survey. The problem can be more serious in establishment surveys where the sample may have to be finely tuned to produce estimates for each of a very large number of branches of activity (industries), with possibly only a small sample allocated to any one of them. Special procedures are often needed to deal with reclassification of units across industry or economic sector boundaries. Special procedures are also required for dealing with changes in boundaries of units in area frames;

(*f*)   Under-coverage, that is to say, units not represented in the frame. It is the most serious and difficult problem and biases the results of many surveys. No simple or cheap solutions to the problem of under-coverage exist.

804.   The lack of a one-to-one correspondence between listings in the frame and survey units to be selected affects the probabilities of selection of the units. In more serious situations, it can damage the scientific (probability) nature of the samples that can be drawn.[1]

805.   The problems with frames increase with the interval between the creation of the frame and the conduct of fieldwork for the survey. List frames are generally much less durable than area frames. Consequently, for sample surveys in the following intercensal period, the census is normally the source for area-based frames but not for list frames.

4.   Physical representation and maintenance of the frame

806.   Physical representation of the frame should meet the objectives of easy access, use and manipulation, including the production of summary statistics that can help in sample design and estimation. Those objectives are best served if the frame of the area units is computerized. Each unit must have a unique identifier and information that can be used for its classification, ordering and reordering, and so on, as required for different purposes. It may or may not be reasonable to computerize the lists of households or similar ultimate units. Often a household listing is made just prior to and specifically for the purpose of a particular survey. Computerizing the list in such cases would almost certainly not be useful; indeed, it may do harm by distracting from more important tasks in the survey.

807.   Maintenance and updating of the frame is a difficult task and cannot be covered here in any detail. The following general principles may, however, be noted:

(*a*)   There is a clear distinction between the task of updating a frame of the more stable area units and of the lists of households, establishments or similar small and less stable units; the two differ in the amount of resources and frequency of updates required;

(*b*)   A basic requirement is to develop plans and procedures for frame updating and allocate the resources necessary for the purpose;

(*c*)   It is a good principle to control the frame updating operations centrally and to standardize the procedures to the extent possible. That applies in particular to decisions about urban/rural and other reclassification of area units, adjustment of area boundaries, or the addition, deletion, merging and segmentation of area units, and so on. As far as possible, final decisions on such matters should not be left to the fieldworkers involved;

(*d*)   Full use should be made of the opportunity offered by the population census to undertake major periodic updating of the frames for surveys following the census.

808.   Updating the frame of area units, covering the whole country, is clearly a major undertaking. Its main source is the decennial population census. Creating the basis for

[1] The presence of blanks in the list is an exception to this in that it does not affect the selection probabilities of the units. However, it can be inconvenient and inefficient.

updating area frames for intercensal household surveys (and possibly also for surveys of small establishments based on area samples) should be viewed as a major objective of the census. Between censuses, updating of area frames is best kept to a minimum—for example, once at the mid-censal period—and confined to some especially fast-growing peri-urban areas or other well-defined strata where the updating is both important and feasible. In addition, there are certain improvements that can be confined to only the selected units, such as improving the area maps, clarifying the boundaries prior to the listing of households within sample areas, segmentation of sample primary sampling units to permit further stages of sampling, and compiling information that can be useful in the planning and control of field operations. The work may be done in the areas selected for a particular survey, or on a larger master sample that serves a number of surveys.

809.   By contrast, lists of households (or other ultimate units) need to be updated or redone much more frequently, usually specifically for and immediately prior to each major survey. Often that activity is feasible since the listing is confined to the ultimate area units actually selected into the sample. It is highly desirable that operationally the listing be kept separate from the final sample selection and the main field interviewing, otherwise serious biases can result.

810.   Listing is an expensive operation, and often it is tempting to use pre-existing lists (available from other sources) or old lists (prepared for earlier surveys), or to use the same lists for a number of surveys. The third option mentioned is clearly desirable, but its feasibility depends upon the relationship in design and timing of the surveys concerned. Usable pre-existing lists from administrative and other sources are rarely available in developing countries, at least for surveys aimed at covering the general population.

811.   Using old lists prepared for earlier surveys or censuses raises a crucial question: how old can the lists be before the sampling biases become unacceptably large? In the absence of sufficient empirical information and with the diversity of conditions encountered, only rough indications can be mentioned. When the subject matter is closely linked to family formation and current economic activity, as in many demographic and labour force surveys, it is highly desirable to keep the time lag between listing and interviewing to a minimum, not exceeding six months in any case. For some other surveys, the maximum tolerable gap can be somewhat larger; less so in urban areas where the population tends to be more mobile, and possibly longer in rural areas; less so for lists of households; and somewhat longer when more stable structural units like dwellings or addresses are used.

5.   Master samples

812.   In any survey with a multi-stage design, each stage of the sampling process involves the task of frame preparation and sample selection, until finally a sample of the required ultimate, or lowest stage, units is obtained. For economy and convenience, one or more stages of the task may be combined or shared among several surveys.

813.   The sample resulting from the shared stages is called a master sample. For example, in a three-stage design with enumeration areas, there are segments (smaller areas within enumeration areas), and households as the first-, second-, and third-stage units respectively. A large sample of enumeration areas may be selected, with each sample enumeration area divided into segments and with some or all of the segments retained in the sample. The resulting large sample of segments can serve as a master sample from which smaller samples of segments and households can be drawn for individual surveys. Thus,

the task of frame preparation and sample selection up to the stage of segments is shared among different surveys using the master sample.

814.   The master sample therefore provides a common sample of units up to a certain stage, from which further sampling can be done to serve the needs of individual surveys. Its purposes are to minimize the costs of developing sampling frames and materials and of sample design and selection, by combining those operations for different surveys; to facilitate substantive as well as operational linkages between different surveys and survey rounds; and to facilitate, as well as to restrict and control when necessary, the drawing of multiple samples for various surveys from the same frame.

815.   At a minimum, the concept of the master sample involves the construction of a common, and sufficiently large, sample of primary sampling units that have been selected for use for several surveys. For a particular survey, some or all of those primary sampling units may be taken to continue the sampling process to lower stages as required for the survey. The primary units initially selected may be subject to further sub-sampling for individual surveys, and the sub-samples drawn for different surveys may be independent or related, distinct or overlapping. In the present example, the task of constructing the primary frame and selecting an initial sample of primary units is common to all surveys using the master sample. As noted, the initial sample of primary sampling units for multiple uses is the minimum requirement implied in the concept of a master sample. It may, however, be useful to carry on the shared sampling process to lower stages, such as to ultimate area units or even to the ultimate sampling units. Sampling for individual surveys will be confined within the lowest stage of units obtained in the master sample.

816.   The objectives of using the master sample approach include the following:

(a)   To economize, by sharing between different surveys, on costs of developing and maintaining sampling frames and materials and on costs of sample design and selection;

(b)   To simplify the technical process of drawing individual samples;

(c)   To reduce respondent burden by making it easier to avoid multiple selection of the same household for different surveys;

(d)   To facilitate substantive as well as operational linkages between different surveys, in particular successive rounds of a regular survey;

(e)   To facilitate, as well as restrict and control as necessary, the drawing of multiple samples for various surveys from the same frame.

817.   The savings in costs are achieved to the extent that the office and field operations necessary to obtain samples for different surveys are combined in the form of a master sample. The savings are the greatest when the common sample of primary sampling units (and other higher stage units) is used repeatedly and over an extended period of time, and especially if, after the areas are selected, considerable fieldwork is required to update their maps, descriptions and boundaries or to segment them to create more suitable units.

818.   In practice, master samples also have their limitations or disadvantages, as follows:

(a)   The savings in costs can be negligible when the master sample concept cannot be extended to lower stages of sampling, where the units involved are less stable and the corresponding frames or lists need frequent updating. A case in point is the list of households that often needs to be prepared or updated for individual surveys;

(b)   Reasonable saving can be obtained only if the master sample is used for more than one, preferably several surveys. Effective use of a master sample requires long-range planning and determination of the size and other basic design parameters of the surveys planned for its use. Such planning is not achieved in the circumstances of many countries. In certain cases, the investment made in constructing the master sample can actu-

ally be wasteful, if for instance the planned surveys that were to use it fail to materialize, or are radically changed, or if the master sample turns out to be unsuitable or inadequate for other reasons;

(*c*)   Another basic problem can be the lack of flexibility in designing individual surveys to fit a common master sample. The constraints and requirements (concerning sample size, clustering, stratification, survey timing, and so forth) imposed by the available master sample can be difficult to accommodate and can result in loss of efficiency of the individual surveys;

(*d*)   Drawing sub-samples from a master sample and correctly computing the selection probabilities of the units selected is not always a straightforward operation. It requires continuity and detailed and accurate maintenance of documentation on master sample design and use.

819.   In conclusion, although the use of master samples is not necessarily advantageous or easy in all circumstances, in many situations, a properly planned and maintained master sample can have significant advantages in terms of convenience, economy and control. The instrument can be applied flexibly: it is not necessary for a single master sample to meet all requirements of all surveys or for a uniform sampling scheme to be used for diverse surveys.

5.1.   Possible relationship to a large sample attached to the population census

820.   Often there is a close relationship between the large sample attached to the census for supplementary data collection and the master sample required for subsequent surveys. Hence, a close relationship can exist between the designs of the two. For instance, if special attention and resources have been devoted to staff deployment and training or to the development of maps, other materials and data for the areas included in the attached sample, it will be efficient to ensure that the master sample overlaps with the attached sample as much as possible. The design of the attached sample will have to take that requirement into consideration.

821.   Sometimes the attached sample can itself serve as a master sample for subsequent surveys, in which case it should be designed as such, with replications and other features that will facilitate sub-sampling from it at a later stage. In general, however, owing to their differing objectives, the two sample systems may not fully coincide. For example, to provide acceptable estimates for the whole country, as well as for numerous large and small census domains, the attached sample may have to involve greatly varying sampling rates among domains. For ordinary sample surveys, however, the reporting domains will be typically larger and fewer, and will probably have less extreme variations in size. Consequently, more uniform sampling rates will be appropriate for the master sample. For similar reasons, it is likely that a master sample may not have to be as large as the attached sample, and therefore it may not be necessary to undertake the taxing and expensive operations of maintaining and updating the full attached sample for future use. Both of the preceding examples imply that in many situations the master sample itself requires sampling, often at variable rates, of areas in the attached sample; and that requirement too has to be considered in the design of the attached sample.

822.   Conversely, the master sample may be especially intensive or large (for instance, to serve intensive cost-of-living surveys in a few urban centres or surveys to monitor the effects of intensive programmes confined to particular areas) and therefore samples have to extend beyond the areas included in the attached sample. Whether either or both of the sample systems should, or even can, be modified to minimize their incompatibility depends on circumstances.

C. Census-based area frames for labour force and other household surveys

823.   The nature and requirements of area frames can differ considerably between population-based surveys, such as those on the labour force, and surveys of other types of units such as economic establishments. The present section describes the more specific technical aspects of using the population census as the source of the frame for the labour force survey and other population-based household surveys. Special considerations in the case of economic surveys of small-scale establishments will be taken up in chapter XV.

824.   The construction of sampling frames for household surveys from the population census is determined by the type, size and structure of the required samples in the post-censal period. In particular, in the case of continuous surveys or those periodically repeated such as the labour force survey, it is also necessary to take into account the sample patterns over time (the relation between samples over survey rounds and across related surveys panel, designs and so on).

825.   The use of the population census as a frame for population-based surveys involves at least four crucial steps, as follows:

(*a*)   The construction of an area frame of census enumeration areas covering the whole country and summarizing the census information to the enumeration area level;

(*b*)   Possible reconstruction of area units that can more appropriately serve as the primary sampling units for the surveys, insofar as they differ from the original census enumeration areas;

(*c*)   Stratification of the area units, specifically the primary sampling units, making use of the information available in the frame;

(*d*)   The selection of a master sample of the area units, which can meet the common requirements of diverse surveys for a number of years and can provide a framework for the integration of survey designs and operations.

826.   Additional steps are necessarily involved in the use of the frame of census enumeration areas or other types of area units constructed from those for individual surveys, such as area sampling for specific surveys; the listing of households or other units in the areas selected for specific survey or surveys; and the selection of the final sample of households or other units for enumeration.

1.   The frame of census enumeration areas

827.   Of the four steps mentioned above, the construction of an area frame of census enumeration units is the most critical. It comprises the first step in the processing of census data and is therefore crucial in its own right. Summarizing the census information to the enumeration area level will give marginal distributions of characteristics enumerated in the census, classified by geographical and/or administrative divisions, against which more detailed tabulations from the census can be checked. In addition, it will provide an integrated frame for household—and possibly other types of—surveys. This frame involves a complete list of population census enumeration areas covering the whole country, with the various population characteristics aggregated to the enumeration area level. Ideally, most of the items in the census questionnaire should also be aggregated to the enumeration area level. At a minimum, the following types of items will be useful for constructing a sampling frame:

(*a*)   Geographical and administrative information: province, district, locality and so on;

(*b*)   Characteristics of the locality: apart from urban/rural classification, this item may include information on the size and type of locality in which the enumeration area lies, subdivision or sector within large cities, identification of rural areas with urban characteristics, and so on. In principle, some of that information may come from sources other than the census, such as administrative records;

(*c*)   Number of households and number of persons disaggregated by sex in the enumeration area;

(*d*)   Number of persons by level of education, disaggregated by sex, in the enumeration area;

(*e*)   Breakdown of the population by activity status and sex, separating out persons below working age;

(*f*)   If available, classification of the working population: the number of persons (women and men) by main sector (such as agriculture versus non-agriculture; or agriculture, industry and services) and subsectors;

(*g*)   Classification of the working population by status in employment and by occupation.

2.   Defining primary sampling units

828.   The second step may involve the combining or segmentation of original census enumeration areas to create area units more appropriate for efficient sample design and implementation. Normally, the first choice will be to use existing census areas as they are or at least to confine their restructuring to a small part of the population, insofar as possible, owing to the high cost involved in such an operation for the whole population.

829.   The types of units chosen to serve as the primary sampling units can have a great influence on the quality and cost of the survey. For an area sample, the units need to be well defined, with clear boundaries, accurate maps and descriptions for identification and demarcation, and up-to-date information on size and characteristics. The areas should cover the survey population exhaustively and without overlaps. Their stability over time is another important requirement, especially if their use must extend over a long period of time. The primary sampling units should be of an appropriate size, in line with the arrangement and cost-structure of the survey's data collection operation. On the one hand, if the units are too large, it may not be possible to include a sufficient number of them to obtain a good spread of the sample; furthermore, the costs of listing, subsampling and data collection within big units may also become excessive. On the other hand, if the units are too small and compact, it may be difficult to ensure a sufficient spread within the units to obtain an efficient sample. Small area units also tend to lack clear boundaries and stability over time.

830.   To determine the appropriate size of the units to serve as the primary sampling units, it is necessary to consider the circumstances and objectives of the survey. A variety of practical considerations, including cost, quality control, administration, the availability of the frame for sample selection and the efficiency of the resulting design, need to be addressed simultaneously. A thorough understanding of the diverse considerations can be obtained only on the basis of a good knowledge of sampling theory and plenty of practice. Several patterns can be identified from national survey practices in different countries, in particular in relation to labour force surveys, as follows:

(*a*)   At one extreme, some surveys (mostly in developed countries) use very large primary sampling units. The total number of units in the population is often not large and only a small number of them are selected for the sample. Each unit represents a substantial investment in mapping and listing, and in recruiting and stationing field staff.

Often one or more enumerators are assigned to work exclusively in each area. Therefore, while little or no travel between primary sampling units may be involved, extensive travel and mobility is often necessary within each (relatively large) area, requiring cheap and convenient transport facilities (such as a personal automobile) at the local level. For example, some surveys in the United States use fewer than 100 counties as primary sampling units, a county being a large unit extending from 50 to 100 kilometres across. Within each selected primary sampling unit, the sample of smaller areas and households is, of course, widely dispersed;

(b)   At another level, many samples are based on numerous very small primary sampling units. Each primary sampling unit may be a small cluster of households (for example, from 5 to 15 households), some or all of which are taken into the sample. Such a system may be suitable in densely populated urban areas where lists of housing units and/or very detailed maps of small area segments are available and where, owing to well-developed transport facilities and the short distances involved, travel between units presents no particular problems. Several examples of this type exist in household surveys in developed countries, as well in urban areas in some developing countries;

(c)   At the other extreme is the use of direct sampling of addresses, households or persons for labour force surveys (an approach that a growing number of countries are adopting). The trend is primarily the result of the increasing use of telephone interviewing, which does not require in-person visits to households. The increasing cost of interviewer time and the drastic reduction in transport costs (virtually zero with telephone interviewing) also reinforce that tendency.

831.   In many developing countries neither of the two extreme variants is common in surveys, since the basic conditions for their efficient operation are not met. In many developing countries, while there is a sufficiently developed network of major trunk roads to facilitate travel between areas, local travel can be extremely difficult and time-consuming, even if the physical distances involved are short. That problem diminishes the potential cost-saving benefits of employing a very large and extensive primary sampling unit, since locally recruited enumerators would have a more difficult time covering the units in their area, relative to those from other areas served by the more developed trunk road network.

832.   However, the use of numerous small primary sampling units (and even more so, of completely unclustered samples) is also precluded in the absence of good maps and other materials needed to define suitable small area units. In addition, the number of primary sampling units has to be kept limited to control travel and supervision costs. Indeed, a common requirement in the choice of the design is to ensure that each primary sampling unit yields a large enough sample to keep the enumerators occupied for a sufficient length of time (for example, a few days) in each area.

833.   It is for such reasons that so many surveys in developing countries use census enumeration areas as the primary sampling units, an enumeration area typically consisting of between 100 and 300 households. Nevertheless, there are examples in developing countries of surveys using a small number of very large units (such as whole provinces, counties or cities) as primary sampling units. Often, the reason is some perceived administrative convenience of this type of design. It is doubtful whether the choice of such large units as primary sampling units is actually justified or efficient. Examples include the use of provinces in some surveys in Thailand and the use of counties as primary sampling units in surveys in China. Of course, there are situations in which administrative, supervision or travel considerations dictate the use of fewer and larger primary sampling units than typical census enumeration areas, but still not as few nor as large as whole provinces, districts or counties. Examples are the use of such intermediate level

units as communes or townships in China, *tehsils* in India, *kecamatan* (subdistricts) in Indonesia or *amphoe* (communes) in Thailand in surveys requiring particularly intensive data collection and supervision activities.

834.   It is important to understand the possible effect of using very large areas as primary sampling units. Apart from considerations of sampling efficiency, there is an added reason for the undesirability of using a small number of very large areas (especially large administrative areas such as provinces or whole districts or counties) as primary sampling units. Often, such large units represent collections of different types of areas (or other types of units) that need to be separately represented in the sample proportionately or in some other controlled way. However, the use of large areas that mix different types of units as primary sampling units does not readily permit control over the representation of the different types of units within the sample. An important example is the mixing of urban and rural areas in big administrative units. If such units are used as primary sampling units and the total number in the sample is small, and if stratification by the urban/rural composition of the areas cannot be used in the sample selection, then it cannot be ensured that the urban/rural composition of the resulting sample reflects that of the population. This issue can have serious implications for any survey, but especially in a situation where there are likely to be very large urban/rural differentials in many of the indicators to be measured.

835.   The main implication of the above remarks is that in the circumstances of many developing countries, census enumeration areas constitute basically appropriate primary sampling units for labour force and similar household surveys. However, the provision of suitable units for sample surveys is by no means the only consideration in the choice and construction of census enumeration areas. The first objective in a census has always been to provide complete coverage of the population and facilitate practical implementation and control of census operations. Hence, it can be expected that some restructuring of the census enumeration areas may be required to create more suitable primary sampling units. Grouping of census enumeration areas when larger units are required is relatively easy since such operations can often be carried out in the office, without involving fieldwork. Segmentation of enumeration areas that are too large to serve as primary sampling units is usually more costly and time-consuming, and therefore has to be undertaken more selectively.

3.   Construction of strata

836.   The third step—stratification—is important for the efficiency of the designs based on the population census enumeration area frame. Stratification is generally a very effective way of improving efficiency in multi-stage (clustered) designs of the type in many surveys, which may be subject to large design effects. A common system of stratification for all household surveys will be convenient and facilitate their integration. Each different type of survey (household, economic, agricultural) has its own requirements in relation to stratification. However, that does not preclude the use of a common framework, at least in part.

3.1.   Primary stratification

837.   At the primary (highest) level, stratification must correspond to the reporting domains, such as regions or provinces, with appropriate urban/rural classification. An urban/rural classification should correspond to administrative (legal) definitions wherever possible, owing to reporting requirements. It may be more than a simple dichotomy, distinguishing big cities from other urban areas, for instance, and rural localities by size category.

838.   Often a common system of primary stratification can be used for all types of surveys and censuses, whether household, economic or agricultural. Hence, up to this point at least, a common enumeration area frame can be constructed for general use. It may become necessary to distinguish between different types of uses in further, more detailed, stratification.

3.2.   Further stratification

839.   It is often necessary to go beyond simple geographical stratification. Within each primary stratum as defined above, further stratification is desirable on the basis of the characteristics of the primary sampling units. In creating strata, a number of decisions have to be taken. What characteristics should be used for stratification? How many strata should be created? How should boundaries of the strata be determined? Apart from explicit strata, what characteristic or characteristics should be used for ordering the areas to provide implicit stratification when systematic sampling is used?

840.   The important point to keep in mind in relation to the preceding questions is that the answer may differ from one part (primary stratum) to another. It is neither necessary nor efficient to seek a uniform system for the whole country or even for a whole province.

3.3.   Stratification variables

841.   When a range of enumeration area characteristics are available as potential variables for further stratification, from among those a choice has to be made of the most useful characteristics (for example, household income, the incidence of informal sector units, characteristics of dwellings, occupation of household head).[2] The first step will be to examine the range of variation (in relative terms) or the coefficient of variation of each characteristic, separately for each primary stratum. That involves examining the frequency distributions and coefficients of variation where appropriate (the unit of analysis being the census enumeration areas or other type of areas serving as the primary sampling units), for each important characteristic within each primary stratum. Characteristics with high coefficients of variation or ranges of variation are candidates for stratification. It is generally more effective to use a number of relevant characteristics simultaneously, each classified coarsely, than to use fine classes of a single variable. However, it is not useful to include all characteristics if they are highly correlated. For that reason, correlation between all variables with high ranges or coefficients of variation should be examined.

842.   In general, the variables applied are dependent upon the focus of the survey to be undertaken. Therefore, in addition to geographical or administrative primary stratification, the following criteria are likely to be useful for further stratification of household survey frames:

(a)   In urban areas, more detailed information on type and size, apart from geographical location (administrative subdivisions);

(b)   In rural areas, percentage of employed persons in non-agricultural activities, and/or percentage of females who are literate.

843.   More sophisticated approaches than the above rather simple one of examining potential stratification variables one by one or in pairs or in other combinations are possible, such as using multivariate analyses. However, the suggested simple approach should suffice for many purposes since the primary stratification is already likely to be quite detailed, and the scope for further stratification in the case of household survey frames is limited. Economic and agricultural surveys have their own requirements in regard to controlling the sector of activity in detail.

2 If the same sample frame is used for different surveys, especially on non-economic topics, then one must be careful in choosing the stratification variables to ensure that the frame remains appropriate and sufficiently flexible to undertake those surveys.

3.4.   Number and size of strata

844.   Within each primary stratum defined above, the number of further (sub) strata to be created depends on (*a*) the number and distribution of the units (enumeration areas) in the population, and (*b*) the likely number of units to be taken into the (master) sample.

845.   The basic rule is to try to create strata roughly equal in population size. Hence, the strata boundaries will depend on the distribution of units according to the stratification variable(s). However, that basic rule can be modified to take into consideration the likely differences in the sampling rates, which normally will be applied in different domains. It is useful to create smaller (that is, more numerous) strata in domains, which are likely to be sampled at higher rates.

D.   Selecting a master sample of enumeration areas for household surveys

846.   The fourth step—the construction of a master sample of areas (enumeration areas)—has a number of objectives. It helps to (*a*) reduce the cost of frame maintenance; (*b*) facilitate sample design and selection for individual surveys; (*c*) control sample rotation over time for regular surveys; and (*d*) permit proper integration and separation between the different surveys as required.

847.   The idea of a master sample is to select a large sample of areas. The sample should be large enough, and it should be distributed and structured so as to permit easy drawing of representative subsamples as required for particular surveys or survey rounds. Owing to all the cost, operational and statistical advantages, the construction of a master sample, especially for household surveys, is highly recommended. Separate master samples may be constructed for economic and agricultural surveys and censuses.

848.   The guiding principles in constructing a master sample are the following:

        (*a*)   In size and distribution, the master sample must be able to meet known and unexpected sampling requirements of all the surveys to be conducted using it;

        (*b*)   The master sample should be constructed such that samples for individual surveys can be drawn from it quickly and cheaply in the required relationship with each other (such as sample independence, overlap, rotation and so forth).

1.   Size of the master sample for household surveys

849.   As noted above, the overall size (the number of primary sampling units selected) of the master sample should be large enough to meet the requirements of all the samples that may be drawn from it. It is better to err on the liberal side than to fall short: in practical samples it is much easier to reduce the sample size by dropping units than to increase it by selecting supplements from the population that has already been subjected to previous sample selection.

2.   Allocation across domains

850.   The master sample should be allocated among the reporting domains, regions, provinces and so on, to reflect the allocation of the main surveys, which will be based on it. In large countries and also in other situations requiring subnational results, that procedure may involve disproportionate allocation across regions or provinces and, increasingly, even among districts in view of the growing need in many countries for decentralized reporting.

3. Selection of master sample areas

851.   Within strata, master sample areas should normally be selected with probabilities proportional to the population of the area (number of persons or households). That is the common scheme in household surveys, and it is consistent with the proportionate allocation within provinces. The actual selection probabilities will necessarily vary in line with the provincial allocation.

4. Replications

852.   It is common and convenient to select the master sample in the form of independent replications, each representative of the whole population. Normally, the replications will be identical in design (stratification, clustering and so forth) and size. The purposes of such replicated design include the following:

   (*a*)   To facilitate the selection of samples for individual surveys;

   (*b*)   To permit the required patterns of sample rotation in regular surveys;

   (*c*)   To permit the desired level of overlap or separation between different surveys;

   (*d*)   To reduce or enhance the sample size at a short notice;

   (*e*)   To permit flexibility in changing the allocation of the sample across domains.

853.   Occasionally, replicated designs have been used to produce quick, approximate estimates of variances, which is possible when a reasonable number of replications is available. In practice, replications can be created simply by systematically dividing a single systematic sample into the required number of parts. Taking, for instance, the Kth unit systematically from the master sample will give one of the K possible replications.

854.   The basic idea is that the sample for any particular survey can be constructed simply by taking one or more replications; that the rotation of a continuous survey can be obtained simply by replacing one or more replications from one round to the next; or that a sample allocation different from that in the master sample can be achieved by taking different numbers of replications in different domains; and so forth. In short, in place of individual primary sampling units, whole replications—each representative of the entire population or the domain of interest—become the basic "units" for the construction of individual samples of the required size and structure.

Chapter XV

The population census as a frame for household-based censuses and surveys of agricultural and small-scale economic units

A. Introduction

855. One of the main uses of the population census for sample surveys is to provide sampling frames for different types of surveys following the census. That application is widely recognized in the case of labour force and other household or population-based surveys. However, it is less commonly appreciated that it applies equally to economic surveys, in particular of small-scale establishments, in both the agricultural and non-agricultural sectors. Indeed, that is also true of what are often referred to as agricultural and economic censuses. Outside the population field the term "census" or "sample census" is commonly used to refer to operations that are not in fact a complete enumeration of all units in the study population but include a probability sample. For example, most agricultural censuses are conducted on a sample basis, albeit often on a large scale. Similarly, although censuses of economic establishments cover large and medium-sized units on a 100 per cent basis, it is often reasonable as well as practically unavoidable for them to cover the numerous small establishments only on a sample basis.

856. Surveys and censuses of small-scale economic units have special characteristics and requirements. In most situations, the samples have to be selected in multiple stages using area-based designs. To be efficient, the sampling designs have to take into account the patterns of geographical distribution of the units, which is often very uneven. Special information is required for the construction of such frames. The census of population can be a major source of required information for constructing sampling frames for what may be termed "economic surveys", as distinct from population-based surveys such as the labour force surveys, provided its provision is carefully taken into account in the design and planning of the census.

857. Economic surveys refer to sample-based enquiries (whether termed "surveys" or "censuses") concerned with the study of characteristics of economic units, such as agricultural holdings, household enterprises, own-account businesses or other establishments covering diverse types of units, products, activities or sectors of the economy. The reference here is to sample surveys of economic units which, like households, are small scale, numerous and widely dispersed in the population, and which may be unstable as well as informal (see paras. 395-400). Units that are medium to large in size, few in number or are not widely dispersed may require different approaches, often based on list frames obtained from administrative sources, such as operating permits.

858. A variety of technical considerations are involved in the design and use of the population census as frame for economic and agricultural surveys and censuses. It is useful to clarify the role of establishment surveys in the provision of information on

economic characteristics and activity of the population and to identify the basic similarities and differences of sampling for small-scale economic establishments and typical household surveys such as the labour force survey. It is important to appreciate those similarities and differences in order to understand how the results from the population census may be used for the construction of sampling frames for economic and agricultural surveys and censuses, and what steps must be taken in the design and implementation of the census to ensure that it yields the information required for the construction of such sampling frames.

B.   Censuses and sample surveys of establishments: similarities with household survey design

859.   The type of sample designs used in typical household surveys provides the point of departure in this discussion of sampling of other small-scale units. Indeed, there may often be a one-to-one correspondence between such economic units and households, and households rather than the economic units as such may directly serve as the ultimate sampling units. Nevertheless, despite much common ground with sampling for population-based household surveys, sampling small-scale economic units involves a number of different and additional considerations.

860.   National or otherwise large-scale household surveys are typically based on multi-stage sampling designs. Great variations are of course possible in the types of designs used in different circumstances and for different surveys. A commonly used type of design is considered briefly below.

861.   In a commonly used type of design, the first step is to select, in one or more stages, a sample of area units and at the last stage to select within each sample area, a sample of ultimate units (for example, dwellings, households or persons). Increasingly—including, and especially in, developing countries—a more or less standard two-stage design is becoming common. In that design, the first stage consists of the selection of area units with probability proportional to some measure of size, M_k, such as the estimated number of households or persons in area k from some past source providing such information for all areas in the sampling frame. At the second stage, ultimate units are selected within each sample area with probability inversely proportional to size. The overall probability of selection of a unit in area k is:

$$f_k = \left(\frac{aM_k}{M} \right) \left(\frac{b}{N_k} \right) = f \left(\frac{M_k}{N_k} \right),$$ [1.1]

where (a, b, M and f) are constants. Here, a is the number of areas selected, when M is the total of M_k values in the population; b is the expected number of ultimate units selected per sample area; hence ($a\,b$) = n is the expected sample size; and f is a constant defined as:

$$f = \frac{ab}{M} = \frac{n}{M} .$$ [1.2]

862.   It is common in national household surveys to aim at self-weighting or approximately self-weighting designs. This often applies at least within major geographical domains such as urban/rural or regions of the country. The denominator N_k may be the same as M_k (the measure of size used at the first stage), in which case the sample is self-weighting with $f_k = f = const$. Alternatively, N_k may be the actual size of the area, which gives a "constant take" design, that is, with a constant number "b" of ultimate units selected from each sample area irrespective of the size of the area. It is also possible to have N_k as some alternative measure of size, for instance representing a compromise between

the above two designs. In any case, M_k and N_k are usually closely related to—and are meant to approximate—the actual size of the area.[1]

863. The selection of ultimate units within each sample area requires a listing of those units. As mentioned above, existing lists may have to be updated or new lists prepared for the purpose of capturing the current situation. No such lists are required for areas not selected at the first stage. The absence of up-to-date lists of ultimate units for the whole population is a major reason for using area-based multi-stage designs.

864. Surveys of small-scale economic units, such as agricultural holdings or other types of household enterprises, share many of those features. As is the case with households, such units tend to be numerous and dispersed in the population. In fact, households themselves may form the ultimate sampling units in such surveys, the economic units of interest coming into the sample through their association with households. In a situation similar to that of household surveys, up-to-date lists of small-scale economic units are not typically available for the entire population. As a result, it is necessary to resort to an area-based multi-stage design, just as in the case of a typical household survey.

C.   Special features of sampling for economic surveys

865. Despite the similarities noted above, certain major differences exist in the design requirements of population-based household surveys and surveys of small-scale (often household-based) economic units. They arise from differences in the type and distribution of the units and in the reporting requirements. The main difference is that economic surveys generally require major departures from self-weighting designs.

1.   Heterogeneity

866. Household surveys are generally designed to cover the entire population uniformly. Different sub-groups (such as households by size and type, age and sex groups in the population and social classes) are often important analysis and reporting categories but, except possibly for geographical classes, are rarely distinct design domains. By contrast, economic units are characterized by their heterogeneity with respect to characteristics, such as size in terms of number of workers and economic activity performed, with a much more uneven spatial distribution. The population comprises multiple sectors—broad aggregations of economic activity, often with great differences in the number, distribution, size and other characteristics of the units in the different sectors—representing different types of economic activities to be captured in the survey, possibly using different questionnaires and even different data collection methodologies. Separate and detailed reporting by sector tends to be a much more fundamental requirement than is the case for different population sub-groups in household surveys. The economic sectors can, and often do, differ greatly in size (number of units in the population) and in sample size (precision) requirements, and hence in the required sampling rates. Therefore, it is necessary to treat them not only as separate analysis and reporting categories, but also as distinct design domains.

2.   Uneven distribution

867. The heterogeneity of economic units as described in the preceding paragraph is accentuated by the uneven geographical distribution of economic units of different types, in terms of relevant characteristics. Normally, different economic sectors to be covered in the same survey are distributed very differently across the population. Their distribution can vary, with (*a*) a few areas having high concentration of some sectors; (*b*) most areas having a representation of some sectors (in other words, the sectors are widely dispersed

across the geographical areas); or (*c*) many sectors in the intermediate or mixed situation, which are neither highly concentrated nor widely and uniformly dispersed. Those patterns of geographical distribution have to be captured in the sampling design. Although population sub-groups of interest in household surveys can also differ in their distribution (as in the typology of geographical, "cross" and "mixed" subclasses defined by Leslie Kish, R. M. Groves and K. Krotki (1976) in the analysis of design effects), in such surveys normally type (*b*) rather than type (*c*) predominates. By contrast, often situation (*c*) predominates in economic surveys and, as noted above, such mixed sectors need to be treated as distinct design domains.[2]

2 It is worth noting that what has been described above assumes some sort of stability over time in the geographical and industry orientation of the economic unit. (See also para. 886.)

3.   Sampling versus survey units

868.   A number of other factors make the design of economic surveys more complex than that of household surveys. Complexity arises from the possibility that the ultimate units used in sample selection may not be of the same type as the units involved in data collection and analysis. The two types of units may lack one-to-one correspondence. For instance, the ultimate sampling units may be, and often are, households, each of which may represent none, one or more than one type of economic activity of interest. Furthermore, the same household may undertake different types of economic activities, for example agriculture as well as the processing and trade of the resulting products. Hence, seen in terms of the ultimate sampling units (households), different sectors (substantive domains) are not disjoint but overlapping. This gives rise to two possible design strategies: (*a*) an integrated design based on a common sample of households, in which all sectors of activity in which a selected unit is engaged would be covered simultaneously, or (*b*) separate sectoral designs, in which the sector populations (in terms of the sampled units, households) and hence the samples may overlap. In each sectoral survey, activity of the selected households pertaining only to the sector concerned would be enumerated.

869.   Separate surveys are generally more costly and difficult to implement. The overlap between the sectoral samples may be removed by characterizing the sampling units (households) in terms of their predominating sector. That characterization helps to make the sampling process more manageable. However, it precludes separate sectoral surveys: insofar as the sample for any particular sector is restricted to only the households in which that sector predominates over all other sectors, the coverage of the sector remains incomplete.

4.   Sampling different types of units

870.   In practice, the process of identifying and separating out the ultimate survey units into different sectors and of applying different sampling procedures or rates by sector is often costly, difficult and error-prone. Hence it is desirable, so far as possible, to absorb any differences in the sampling requirements by sector at preceding area stage(s) of sampling, to avoid having to treat different types of units differently at the ultimate stage of sampling. The cost of such (very desirable) operational simplification is, however, the increased complexity of the design that this may entail.

D.   Data required for sample design and selection

871.   It follows from the above that the frame for sampling small-scale economic units has similar but more complex requirements. The required sampling frame will typically consist of area units (which form the primary sampling units in a two-stage design or the ultimate area units in a design with multiple area stages). The units can be similar or even the same as those for household surveys, typically the population census enumeration areas described in previous chapters.

872.　However, in addition to the information on simple population counts by census enumeration area at the time of the population census, information is also required on the economic characteristics of each area unit, for instance on the number of economic units of different types and sectors present in the area, the number of persons employed in different sectors and occupations and so on.

873.　The information requirement by economic sector of activity is more elaborate than a single measure of population size of the type normally required for probability proportional to size sampling in household surveys. It is important, especially in the context of countries with limited administrative sources, that potential sources such as population, agricultural and economic censuses be designed to yield such information, which is required for the efficient design of surveys of small-scale economic units.

1.　Information on economic characteristics of area units and corresponding measures of size

874.　National surveys of small-scale economic units in developing countries typically have to be based on multi-stage sampling designs. Great variations are of course possible in the type of design used in different circumstances and for different surveys. One of the most commonly used types of design is briefly explained below. The objective is to point out the characteristic differences between economic surveys and household surveys such as the labour force survey.

875.　The population of units in economic surveys comprises a number of "sectors", such as different types of holdings, agricultural, processing and trading activities, or products. Sample size requirements generally have to be met separately for each sector. The available sampling frame has to consist of area units (which form the primary sampling units in a two-stage design or the ultimate area units in a design with multiple area stages); for each area unit, information is required on the expected number of economic units of different types.

876.　In essence, the overall selection equation (following equation [1.1] above) is of the form:

$$f_k = f\left(\frac{M_{.k}}{N_k}\right);$$
[1.3]

where M_k is some measure of size assigned to the area in its selection by probability proportional to size, and the selection of ultimate units within the area is with probability inversely proportional to the size, that is to say, measure N_k in the frame, assumed to estimate the actual size of the area.

877.　The design weight to be applied at the estimation stage is inversely proportional to the overall selection probability:

$$w_k = \frac{1}{f_k}\alpha\left(\frac{N_{.k}}{M_k}\right).$$
[1.4]

878.　The expected number of units contributed to the sample by area k is

$$f_k N_{.k} = f M_{.k}.$$
[1.5]

Summing over all areas k in the population in [1.5] gives the total expected sample size

$$n = f M.$$
[1.5a]

The expected number of units of a particular sector i contributed to the sample by area k is

$$f_k \, N_{i.k} = f M_k \left(\frac{N_{i.k}}{N_k} \right) = f M_k \, P_{i.k}. \qquad [1.6]$$

The basic design problem is to determine the modified size measures M_k such that the sample size requirements

$$n_i = f \sum_k \left(M_k \, P_{i.k} \right) \qquad [1.7]$$

or, in terms of relative quantities more convenient for numerical work,

$$\frac{n_{i.}}{n} = \sum_k \left(\frac{M_{.k}}{M} \right) P_{i.k} \qquad [1.7a]$$

are satisfied for all sectors i simultaneously in the most efficient way.

879.   In [1.7] the sum is over all areas in the population and not merely the sample. Note also that the above formulation assumes that at the ultimate sampling stage, units within a sample area are selected at a uniform rate, inversely proportional to $N_{.k}$, irrespective of the sector. This is a very desirable feature of the design in practice. The process of identifying and separating out the ultimate survey units into different sectors and of applying different sampling procedures or rates by sector is often costly, difficult and error-prone. The preceding sampling stages are assumed to absorb any difference in the required sampling rates by sector by incorporating them in the definition of the size measures $M_{.k}$.

880.   The most convenient but also the most unlikely situation in the application of [1.7] occurs when units of different types (sectors) are geographically completely segregated; in other words, when each area contains units belonging to only one particular sector. In reality, the situation is more complex since areas generally contain a mixture of units of different types (sectors). Clearly, it is necessary to inflate the size measure (and hence inflate the selection probabilities) for areas with proportionately more units from sectors that need to be oversampled, and vice versa.

2.   Collecting information on economic characteristics of population census enumeration areas

881.   To apply a design like the one described above, it is necessary to have certain types of information in the sampling frame. It must include not only some single (and simple) measure of size such as the number of persons or households in each sample area at some time in the past (such as at the time of the population census), which is related to the actual size of the area at the time of sample selection, but also, for surveys of small-scale economic units, some information on the number of units of different types in each area.

882.   At a minimum, the population census frame involves a complete list of enumeration areas with the various population characteristics aggregated to the enumeration area level. They include geographical, administrative and other information on characteristics of the area as well as some demographic and social information on the population in the area. The usefulness of the population census to serve as a frame of area units for economic surveys is greatly enhanced if, in addition, information is collected and compiled to the enumeration area level on some basic economic characteristics of the population.

883.   For instance, the information collected could include the following:

(*a*)   Information on distribution of the population according to broad sectors of activity, which is useful for constructing frames and selecting samples for censuses and

surveys on these topics. Even if only this type of information is available, it can be of great value in increasing the efficiency of the sample designs for economic surveys, which may be constructed from the frame of census enumeration areas;

(*b*)   Information on the number of economic units by main sector of activity at the time of census in each enumeration area. That type of information is normally not obtainable from the population census without some very careful preparation for the data collection and processing procedures. In particular, questions on either place of work or branch of activity that include the name and address of the employing establishment (see sample questions in chap. VII) should have been asked in the census and the relevant information processed and tabulated as indicated above. Questions have been included in the population censuses of some countries, for example Belize, 2000 ("Do you own any land for farming?"), that might be useful for this purpose in an agricultural census;

(*c*)   Information from past economic and agricultural censuses for the areas that were selected for both types of censuses. Sometimes it may also be possible to incorporate into the population census relevant information from both of them. A crucial requirement is that the economic and agricultural censuses be based on the same system of areas as the population census enumeration areas. A well-planned census system will include that objective, which has been carried out successfully in a number of countries, notably in India and Indonesia among the largest developing countries.[2]

884.   The information on economic characteristics for individual areas (such as census enumeration areas) is required to determine the sampling rates needed to obtain the required sample sizes for different sectors of economic activity, to allow efficient stratification of the area units and also to determine the probabilities of selection of the areas to meet the requirements for reporting by sector in economic surveys. Such information can also be useful when preparing estimates based on data from the survey. Stratification and variations in sampling rates have to be targeted, that is to say, determined on the basis of the composition of the area units in terms of the relative size of the different economic sectors present in the area. Such targeting is required to ensure efficiency and control in the sample design.

E.   Primary sampling units and stratification

885.   Owing to the great variety of surveys of establishments and other economic units, little can be said in general about the appropriate choice of area units to serve as the primary sampling units. Census enumeration areas are often suitable units for the purpose of such surveys. Sometimes it is necessary to create larger areas to serve as primary sampling units so as to capture sufficient numbers of units of interest. That is normally not a major problem since existing area units can be easily combined to define new units for sampling.

886.   With respect to stratification, economic surveys tend to involve units greatly varied in type and size, and hence usually require much more detailed stratification than household surveys. The main stratification required is in terms of the predominating sector of activity in the area. For the purpose of stratification, it is not crucial that such information be highly accurate or up-to-date, as long as it is reasonably well correlated with the area's current economic characteristics. If that is the case, information on the number and distribution of establishments by sector of activity collected during census listing can continue to serve a useful purpose for a number of years following the census.

[2] See for example M. N. Murthy and A. S. Roy (1970); and V. Verma (1989).

Chapter XVI

Use of census data in the production of survey estimates

A. Introduction

887.   Several requirements need to be taken into account when population census data are used in the production of estimates from sample surveys, including the following:

(*a*)   A comprehensive survey such as a continuous labour force survey is used to produce a wide range of estimates for different populations at different levels of aggregation, and over different time periods;

(*b*)   Several types of variables may be involved, such as the rates of labour force participation, unemployment rates, status in employment, occupation and industry. The frequency and precision requirements may differ for different types of variables;

(*c*)   Estimates are normally required for different types of units, such as individuals, households, families and communities;

(*d*)   Estimates are required not only at the total (national) level but also for separate geographical and administrative areas, for urban and rural classification, and for different subpopulations such as demographic classifications;

(*e*)   Estimates are required for proportions, rates, means and other ratios (where the numerator and the denominator are obtained from the same source), as well as for population aggregates (where the survey results have to be inflated to the population total using information internal or external to the sample survey). Often, surveys of moderate size are able to yield estimates of population totals (such as the total number unemployed) but with much less accuracy, in terms of both variance and bias, than estimates of various types of ratios (such as the unemployment rate).

888.   Furthermore, various types of estimates can be distinguished in terms of the time dimension. They include the following:

(*a*)   Estimates of current levels for each reporting period (month, quarter, year) of the survey;

(*b*)   Estimates averaged or aggregated over longer periods, such as annually from a quarterly survey;

(*c*)   Estimates of trends or net change from one period to another;

(*d*)   Possibly, estimates of gross change at the level of individual person or household, such as flows among categories of economic activity status (although such statistics are still rather uncommon);

(*e*)   More complex "composite" estimates.

889.   In the production of any of the preceding types of estimates of population parameters, the sample data need to be appropriately weighted. Most common estimates from a survey, for example on the labour force, are in the form of weighted ratios. The most basic forms in which the sample weights appear in estimation are as follows:

$$r = \frac{\sum_i w_i y_i}{\sum_i w_i x_i} = \frac{y}{x} \qquad\qquad\qquad [2.1]$$

where r is the ratio of two weighted sample aggregates y and x, with each of those aggregates formed by summing up the individual sample values (on two survey variables, for instance, y_i and x_i), for each sample unit i weighted by its associated weight w_i.

890.   The estimate r may be the ratio of two substantive variables, such as wage rate (r) as the ratio of (weighted aggregate of) wages received (y) to the (weighted aggregate of) number of hours worked (x). Many labour force survey statistics, such as means, proportions or percentages, are merely simpler forms of ratios. For instance, average hours worked (r) is the ratio of the (weighted) aggregate of hours worked to the weighted number of workers; the denominator in this case being simply the weighted count of cases (meaning that x_i in the denominator is identically equal to 1 for all sample cases involved in the estimation). Similarly, in computing the unemployment rate, y_i is a simple dichotomy (unemployed=1, employed=0), with x_i=1 as simply the count for all persons in the labour force for which the unemployment rate is being computed.

891.   Labour force and other surveys are also often used to estimate population totals, such as the total number of persons unemployed. The estimations generally take the form:

$$\hat{Y} = r X = \frac{y}{x} X = \left(\frac{\sum_i w_i y_i}{\sum_i w_i x_i} \right) X, \qquad\qquad\qquad [2.2]$$

where X is an estimate of the denominator (x) from some more reliable source external to the survey. In principle, the sample quantity y also provides an estimate of the above, but a ratio estimate in the above form is generally more precise (see below).

892.   The population census is often a valuable basis for the information required to compute appropriate sample weights w_i, and even more so for obtaining the external inflation factor X. The following sections cover some of the technical issues involved in weighting and estimation from surveys using external information from the census and other sources.

B.   Weighting of sample data

893.   The objective of weighting sample data is to improve the representativeness of the sample in terms of the size, distribution and characteristics of the study population. For example, when sample units have been selected with differing probabilities, it is common to weight the results inversely proportional to the unit selection probabilities so as to reflect the actual situation in the population. In a survey selected from a good frame and well implemented with high response rates, application of the above mentioned design weights is all that may be required.

894.   In practice, however, the situation is usually more complicated owing to shortcomings in the selection and implementation of the sample (which introduce biases in the results) as well as to the need for (and possibilities of) introducing improved estimation procedures to reduce variances. The need for more complicated estimation procedures tends to be greater in surveys with high non-response rates and coverage errors; inconsistencies in the definitions of units used at different stages in the survey operation; departures from representative (probability) sampling; and small sample sizes. It is also the case when estimates for many separate subpopulations have to be produced. The need (and the opportunity) is also greater in the presence of more extensive and more reliable external information for the purpose.

895. Issues involved in sample weighting can be complex, and the best solution may be situation-specific, depending on the nature of the data at hand, the sources of error that need to be controlled, knowledge of the specific circumstances and limitations of the survey, and the existing practices and preferences of the survey organization. Nevertheless, there are major advantages in following certain basic standards and a systematic approach, as outlined below.

1. Source of information

896. In applying the weights, the best use has to be made of the information available, both internal to the sample and from external sources. The primary role is given to information internal to the survey; external information is introduced to the extent judged useful for further improving the representativeness of the sample.

897. Five types of information sources may be noted, which can be used in a systematic manner to apply weights in a step-by-step procedure, as follows:

(*a*) Sample design: the design probabilities of selection of each household;

(*b*) The sampling frame, which may provide additional information on sample areas and on all responding and non-responding households;

(*c*) Sample implementation: response rates and information on non-respondents;

(*d*) Other, significantly larger survey(s) with better coverage, higher response rates and more reliable information on certain characteristics of households and/or persons. In fact, the labour force survey itself is often a good source of information for weighting in order to improve the representativeness of more complex and difficult surveys with smaller samples and/or poorer response rates;

(*e*) Population projections from the census (and/or current registers where available) providing information on the characteristics of the population and their distribution.

898. When the same or similar information is available from more than one source, priority should first be given to the source internal to the survey; for instance, weighting to compensate for differences in selection probabilities and known incidence of non-response should always be applied before corrections on the basis of external data are introduced. In using external information, it is necessary to ensure that the information is significantly more reliable than the available information internal to the survey, that the items of information used are comparably defined and measured, and that the coverage and scope are the same. For instance, if the survey is confined to the population residing in private households, external information on the population used for sample weighting should be similarly restricted.

2. A step-by-step procedure

899. For achieving common standards, as well as for clarity and convenience, it is desirable to adopt a step-by-step procedure that separates out the different aspects of weighting. As a rule, each step should be applied separately so that its contribution to the final weights can be identified. For a household survey of the general population, for instance, the basic steps involve the application of the following weighting factors in sequence, which are described in turn in the next sections:

(*a*) Design weights;

(*b*) Non-response weights;

(*c*)   Weights correcting the distribution of units in the sample, to agree with more reliable external information;

(*d*)   Where applicable, appropriate inflation of the sample estimates to agree with external control totals, such as the current size of the total population.

C.   Design weights

900.   Each household or person in the sample is weighted in inverse proportion to the probability with which it was selected. The design weights are introduced to compensate for differences in the probabilities of selection into the sample, the weight given to each sample unit being inversely proportional to its probability of selection.

901.   For instance, with p_i as the unit's overall probability of selection into the sample, the design weight may be computed as:

$$w_i = \frac{1}{p_i} \left(\frac{n}{\sum \left(1/p_i\right)} \right), \qquad\qquad\qquad [2.3]$$

where the summation is over n units successfully enumerated in the survey. The above form has the convenient property $\sum w_i = n$, meaning that the weights have been scaled to average 1.0 per unit enumerated in the sample. It is the relative rather than the absolute values of the selection probabilities that matter. With multi-stage sampling design, the reference is to the overall selection probabilities of the ultimate units; other aspects of the sample structure such as clustering and stratification do not enter the equation.

902.   With probability sampling (units selected with known non-zero probabilities) the design weights can be computed, in principle and mostly also in practice, from information on sample design and outcome internal to the survey, without requiring external information from census or other sources. However, in certain circumstances it is useful, and necessary, to incorporate into the design weights a correction for known exclusion or gross under-coverage of some parts of the study population, which may have occurred owing to defects in the sampling frame or for other reasons relating to the procedures of sample selection and implementation. One way to apply such a correction would be to reduce the design probabilities of selection, that is, to increase the weights correspondingly, in proportion to the coverage rates in the affected domains. Another would be to incorporate compensation in other covered domains similar to the one(s) excluded. For instance, in countries where certain remote islands or areas have been excluded, the weights for certain other, perhaps similarly remote but still included, areas may be increased in proportion to the populations involved.

D.   Non-response weights

903.   Non-response weights are introduced to reduce the effect of differences in response rates achieved in different parts of the sample. The weights are based on characteristics that are known for both responding and non-responding units. Weighting for non-response is particularly important when rates of non-response are high and generally variable from one part of the population to another. Such defects make it necessary to weight the data for non-response to correct for at least gross distortions in the distribution of the achieved sample according to various characteristics of the analysis units. Of course, weighting for non-response corrects only for distortions resulting from differential non-response in terms of the variables used in the classification. It cannot take into account the effect of the absolute levels of non-response, only that of the relative levels across the classes.

1.   Defining weighting classes

904.   Essentially, weighting for non-response involves estimating the units' propensity to respond as a function of unit characteristics and weighting the responding units inversely to that propensity. In a simpler and commonly used form, weighting for non-response involves the division of the sample into certain appropriate weighting classes, and within each weighting class, the weighting-up of the responding units in inverse proportion to the response rate, so as to make up for the non-responding cases in that class. Specifically, the following steps are involved:

(*a*)   Division of the sample units into groups or weighting classes, showing the number of units (households, persons and so on) selected and the number of interviews completed in each class. Depending on the number of control variables involved, several parallel divisions may be involved, each providing marginal distribution of the whole sample according to some classification variable or some cross-classification of a set of variables, such as marginal distributions by geographical location or stratum, by some other classification of sample areas, by tenure, household type and size, or other characteristics of the households;

(*b*)   Computation of the response rate for each category in the classification;

(*c*)   Assignment of a uniform weight to all units in a category, in inverse proportion to the response rate in the category.

905.   The effectiveness of the procedure depends on the extent to which non-responding units within each class are similar to the responding units in that class in terms of all the other variables not used in the classification. Differences in unit characteristics and in response rates should be maximized across the weighting classes chosen.

906.   Given the preceding requirement, it is still necessary to choose the appropriate number and size of classes to be used. On the one hand, the use of many weighting classes has the possible advantage of reducing non-response bias by creating relatively small and homogeneous weighting categories within which characteristics of respondents and non-respondents can be assumed to be similar. On the other hand, the use of many small weighting classes can result in the application of large and variable weights that can greatly increase the variance of the sample estimates. A compromise is therefore required. The choice will depend on the variability of the response rates across different parts of the sample and how the variations are related to the characteristics of units.

907.   When a number of classification variables are available, each variable, or each cross-classification of two or more variables, may be used to divide the whole sample into parallel sets of weighting classes. Each set will provide the distribution, according to the classification variable(s) concerned, of the numbers of units selected and the numbers successfully interviewed, on the basis of which response rates can be computed in each category of the classification.

2.   Data sources

908.   Clearly, weighting classes can be defined on the basis of only those characteristics that are available for both the responding and non-responding units. While the non-responding cases are identified from the survey records, normally only limited information is available from the survey on non-responding units. External sources, including the population census, are often the primary source of information on those units.

909.   In a multi-stage sample of households, for instance, some of the characteristics involved may pertain to sample areas and others to individual households and persons, including both responding and non-responding units.

910. Area-level characteristics refer to characteristics relating to areas or other aggregates, such as geographical location (administrative divisions), type of place of residence (for example, an urban/rural classification) and a variety of socio-economic characteristics of the areas. Some information of that nature is always available from the sampling frame itself, specifically, the geographical location and other information used for stratification of sampling areas. Additional information may also come from local area statistics from the census or administrative sources.

911. Other characteristics relate to individual households or persons, such as household size and type, tenure of accommodation, income, socio-economic status and other characteristics of the household head or reference person, and to demographic characteristics, activity status and other characteristics of individual household members. Several sources of information on such characteristics are possible, as follows:

(*a*) Where survey records can be linked at the individual level with administrative sources or the population census, both sources can provide information on characteristics of non-responding units in the survey. When such external information is used for non-responding units, the same source should be used for the classification of both non-responding and responding households, even if for the latter the same information is also available from the survey itself. This is necessary to retain consistency in the classification;

(*b*) In situations where the sample is drawn from lists that include relevant information for the classification of the selected units, the information in the frame can be used to provide common classifications for responding and non-responding units. At a minimum, geographical location and other information used for stratification of the sample areas must be used to appropriately define weighting classes for non-response adjustment in household surveys. Other classifications based on additional information of the types described below should be created where possible;

(*c*) In cases where the current sample is based on some previous survey or the census, they can provide information on non-respondents to the current survey;

(*d*) In complex surveys prone to high rates of non-response in particular, it is desirable to make a special effort to collect at least a few basic items of information on each unit selected into the sample, regardless of whether the unit is successfully enumerated in the main survey.

E. Correcting the sample distribution with more reliable external information

912. After the sample data have been adjusted for differential sampling probabilities and response rates, the distribution of the sample according to the number and characteristics of the units will usually still differ from the same distributions available from more reliable external sources such as the population census, projections, registers or other large-scale surveys. Normally, the precision of the estimates is improved by further weighting the sample data so as to make the sample distributions agree with external information on, for example, population estimates classified by age, sex, ethnicity, nationality, urban/rural and so on.

913. In a household survey, the distributions in various categories may involve numbers of households and/or aggregates of some variables measured on each household. With a survey involving households as well as persons as the units of data collection and analysis, it is necessary to control the sample distributions in terms of both types of units.

914. Several considerations are highly relevant in using auxiliary information from a past census in the production of current estimates from surveys. Obviously, the external

information has to be sufficiently accurate and up-to-date, so that its imposition on the sample actually results in improving of the latter. Another basic requirement is that the characteristic(s) used for matching the sample distribution to the external control distribution be the same in the two sources: they should be defined and measured in exactly the same way. This cannot always be taken for granted in using the census data in the weighting of sample surveys. Furthermore, when more than one external source is used, as is normally the case in practice, the information from the various sources must be consistent. Re-weighting on the basis of external control distributions does not require matching of the sample and the external source at the level of individual households or persons. The weighting adjustments are made on the basis of comparison of sample and external distributions at the aggregate level.

915.   The following two points are relevant in relation to the use of population census data for the purpose:

(*a*)   While tabulated census data often provide distributions by one or, at the most, by a small number of individual characteristics at a time rather than by a wide range of variables, in most situations the distribution by only a single characteristic is not sufficient. However, although it is desirable to control all important characteristics simultaneously, having too many controls results in small adjustment cells and large variations in the resulting weights. It is also preferable to control for a number of marginal distributions simultaneously, rather than to use a single variable with a very detailed classification, or a full post-stratification involving many cells in a cross-tabulation of several variables. In addition, the census data can be out of date; however, that tends to affect actual levels more than the underlying distributions;

(*b*)   Another point to note is that external information on the relative distribution of units by characteristics is often more readily available than information on the actual numbers of units in different categories in the population. Further, the relative distributions tend to be less sensitive than absolute numbers to any differences between the survey and other sources in terms of the definitions used.[1]

F.   Estimating totals

916.   It is statistically convenient to scale the weights computed at each of the preceding steps (see paras. 899 ff.) to average 1.0. Alternatively, they can be expressed to any arbitrary scale. The scaling from sample to population aggregates can then be specified in terms of an overall inflation factor, common to all sample units, that inflates the sample numbers or estimated totals to the corresponding aggregates in the population. It does not affect the survey estimation of proportions, means or other ratios, but only the estimation of totals or aggregates. In principle, different inflation factors may be involved in the estimation of aggregates for different types of quantities and for different types of units such as households and persons. However, as explained below, additional considerations are involved in estimating aggregate population values from sample data.

917.   Simple unbiased estimation (with sample data weighted simply by inverse of sample selection probabilities), while often suitable for means and other ratios, usually needs modification when the objective is to estimate population aggregates. It is especially the case for surveys with a multi-stage design and small sample size. With multi-stage sampling design, the resulting sample size varies at random, and therefore aggregates directly estimated from the survey can have a large sampling variance. The problem is even more serious when estimates are required for population subclasses, the selection of which is not explicitly controlled in the multi-stage design.

[1] A convenient method of adjusting the sample distribution to a number of external controls simultaneously is the classical iterative proportional fitting or raking method originally proposed in W. E. Deming and F. F. Stephan (1940). A more accessible source is W. E. Deming (1964).

918.   An equally important problem arises from the fact that estimates of aggregates are biased in direct proportion to the magnitude of the coverage and related errors. By contrast, the effect on estimates of proportions, means, other ratios, and more analytic statistics is often less marked.

919.   The appropriate procedure for estimating population aggregates is generally as follows. In place of simple inflation of the form

$$\hat{Y} = Fy, \qquad\qquad\qquad\qquad [2.4]$$

that is, inflating the sample aggregate y by F, the inverse of the overall sampling fraction, the required aggregate may be expressed in the form of a ratio-type estimate, shown in equation [2.2] above, namely:

$$\hat{Y}_r = \frac{y}{x} X, \qquad\qquad\qquad\qquad [2.5]$$

where y and x are estimated totals from the sample; y being the variable of interest, and x an auxiliary variable for which a more reliable population aggregate value X is available from some external source.

920.   The value and applicability of the above procedure depends on several factors and on considerations that are highly relevant in using auxiliary information from a past census in the production of current estimates from surveys. First, the correlation coefficient between y and x must be positive and preferably large, for example, greater than 0.6 or 0.7 at least. Second, X should be available with higher precision than the simple estimate of the population aggregate that can be produced directly from the sample itself. Third, X in the population and x in the sample should be based on essentially similar measurement on the same population; a difference between the two would introduce a bias into the estimate. The procedure often requires that values of the variable x for individual units—unless they are simply a count of the cases, as in the case of an ordinary mean—be taken from the external source rather than directly from the measurements in the survey, though of course those values must be for the actual units included in the sample.

G.   Back revision of estimates

921.   In many countries, the availability of results from a population census leads to a back revision of intercensal population estimates, and hence to the back revision of estimates from the labour force survey programme (if any) that uses those population estimates as a benchmark. When doing the back revision, the opportunity is also taken to generate revised survey estimates classified by occupation and industry (if those classifications have changed over time). This task can become complex and resource-demanding.

Chapter XVII

The census as a basis for evaluating survey data

922.   Another potential use of census data is to serve as a benchmark for the evaluation of results from post-censal sample surveys. Owing to differences of timing, content, methodology, practical conditions of data collection and so on, the comparability between the two types of sources are limited, and hence a past census cannot fully serve as a benchmark for more current sample surveys. If different classifications and definitions of concepts are used, then comparability of information between censuses and surveys and the capacity to use census data to evaluate survey data will be limited. Nevertheless, the population census can be useful in a number of ways for the above purpose. The questions to be addressed are whether and to what extent data on economic characteristics and activity of the population from the labour force and other sample surveys can be evaluated on the basis of information from the population census and projections. It is useful to distinguish between content and coverage aspects of the evaluation.

A.   Survey content

923.   Surveys such as the labour force survey, often contain information on economic characteristics and activity of the population that is much more accurate than the information it is possible (or desirable) to collect in a population census. The extent to which census data can be used to evaluate the results of such surveys is therefore limited. It is important to begin by recognizing the limitations of census data on activity status and other economic characteristics.

1.   Limitations of census data on labour force characteristics

924.   The main limitation of the census data results from its large size, which allows the census to include only a short series of questions on the labour force. Consequently, individuals, especially those at the margins of the labour force categories—who may be wrongly classified as unemployed or inactive when they are employed, or vice versa—may not be classified with the same degree of accuracy as is normally possible in a labour force sample survey.

925.   In relation to the correct identification of persons in employment, the main problems are likely to be the treatment of persons temporarily absent from work and the correct application of the concept of "work". To determine whether a person is employed or is unemployed or economically inactive, a labour force survey normally includes the necessary questions on duration of and reasons for absence from work, whether payment is received, formal job attachment and so forth. Typically, a census cannot incorporate such a detailed series of questions; consequently, persons away from work are much more likely to be automatically classified as employed.[1] Similarly, detailed prompts and probes are needed to ensure that all the different types of economic activities are included in the

[1] See, for instance, J. Carew, R. Woods and B. Brady (1999).

concept of work (including, for example, volunteer work for profit-making enterprises, water carrying, wood gathering and the making of handicrafts). Such detailed efforts are not always possible in a population census, and persons solely engaged in such activities may be classified incorrectly as not employed.

926.   Similar considerations apply to the measurement of unemployment. To measure this variable properly, taking fully into account the recommendations of the International Labour Office, requires a fairly elaborate series of well-tested and well-administered questions. A labour force survey of limited size can be expected to incorporate such a series fully, but a population census will normally need to rely on an appropriately simplified version. Future starts (those waiting to start a job) tend to be especially difficult to classify correctly.

927.   Furthermore, the interviewers in a survey can be expected to be better trained to administer the questions in the required form and are usually allowed more time to probe in depth when complicated situations so require. By contrast, censuses have to be conducted with more limited resources per unit contacted, and in any case may have to give priority to other aspects, such as getting the basic population counts right. Hence, it is not entirely unexpected that censuses often produce different (generally lower) participation, employment and unemployment rates than do many labour force surveys. As a general rule, it can be expected that, despite being subject to sampling error, a labour force survey will produce more reliable and comparable estimates of activity status and other labour force–related characteristics than a population census, at least at the national level and major subnational domains.

928.   In comparing census and labour force survey results, it is also important to identify, and adjust for where possible, differences in scope, coverage, treatment of non-response and other differences in the mode and conditions of data collection. For a variety of such reasons, differences observed between the census and the surveys on nominally the same items may be wholly anticipated and legitimate, without one necessarily being better than the other. An important consequence of the limits in comparability between survey and census data, even for items that are nominally the same, should be noted in the context of the weighting of survey data on the basis of the census, as presented in chapter XVI. Insofar as the data are not comparable, it would be a mistake to force—as is sometimes done—the survey data to conform to distributions from the census.

2.   Potential for comparison with and evaluation of labour force survey data

929.   Many censuses nevertheless do cover broad aspects of the population's economic characteristics and activity, such as activity status (employment, unemployment, inactivity), status in employment, industry, possibly also occupation, institutional sector of employment, working time, working schedules (full-time, part-time) and so forth.

930.   A good census can provide a solid basis against which the overall plausibility of the survey results can be judged and at least any gross errors identified. It is always worth checking the results of a survey against the census when the latter is available. Furthermore, some of the differences between the two sources have less impact on differentials between population groups than on overall levels. That happens to the extent that different population groups in each of the two sources are subject to similar measurement biases, thus reducing their impact on the observed differentials between groups. Examples include differences in participation or unemployment rates by age, sex, geographical region, urban/rural classification and the like.

B. Survey coverage and representativeness

931.　In contrast to the limited scope for the evaluation of substantive results of sample surveys against the census, information on the size, composition and distribution of the population from the census and projections based on it can provide a solid basis for the evaluation of the coverage and representativeness of the sample.

932.　It has been commonly found that, owing to the special conditions and political context of their implementation, censuses are often able to achieve more complete coverage of the population than sample surveys. That applies to both the size and the spatial distribution of the population (by geographical and administrative location, urban/ rural and so on). Normally, it also applies to the distributions by basic demographic characteristics such as age, sex, and household type and size. The use of census-survey linkage studies (as indicated in para. 756) may also be considered.

933.　Sample survey data, inflated by the inverse of the selection probabilities of the units in the sample, should always be compared where possible with the totals and distribution in the sampling frame and with the census and census projections. That process is known as preparing simple unbiased estimates from the survey. The estimates are a valuable means to evaluate the coverage and representativeness of samples, as explained below.

C. The importance of preparing simple unbiased estimates

934.　It is very important to be able to prepare simple unbiased estimates from the survey data, even though they may be refined and modified subsequently in the production of the final estimates. In practical survey work, the term "simple unbiased estimates" is used in the sense (*a*) that the estimates are produced directly from the survey results without recourse to data external to the survey, by weighting each observation in inverse proportion to its probability of selection into the sample; and (*b*) that at least with moderate to large sample size, the estimates so produced are approximately unbiased in the statistical sense (that is to say, apart from any sample implementation and non-sampling errors).

935.　Such estimates can be prepared only with probability sampling, that is, for samples selected in such a way that each element in the population has a known and non-zero probability of being selected. To prepare good estimates, it is also necessary that problems of sample implementation, such as non-response and under-coverage, do not significantly distort those probabilities. Good simple estimates would also imply that any adjustments, which may have to be made subsequently to improve their precision, will not turn out to be large. In short, being able to produce good simple unbiased estimates indicates that the survey has been designed and implemented properly; hence the importance of such estimates in survey practices.

936.　In a survey with a two-stage sample, for example, involving the selection and household listing of sample areas, followed by the selection and enumeration of households within each sample area, the number of households listed within each sample area would be inflated by the inverse of the area's selection probability and summed over all areas in the sample to provide an estimate of the total population (number of households in this case). This estimate can be directly compared with the expected population size from census projections or other sources. The ratio of the sample estimate of the total population to the expected population size provides a good indication of completeness of coverage of the listing operation. In many surveys it has been found that the ratio falls short of 1.0 by a significant margin, indicating the need to improve the quality of household listing operations and using a more up-to-date list for

2 Note that in estimating population aggregates actual values of the sampling probabilities are required, while for checking on population distribution, relative values of those probabilities suffice.

sample selection. In some countries, the term "slippage" is used to describe the difference between external population counts and population counts from the survey (for example, undercounts of young men).

937.   The above exercise should also be carried out by major geographical classifications of the survey population, by type of place (urban/rural) and other major domains, and also by main demographic classifications such as by age group, sex, household size and type.[2]

Chapter XVIII

Combined uses of census and survey data: current estimates for small domains, including local level estimates

A.   The need for small domain estimates

938.   The census can provide geographically detailed but infrequent and not very current statistics, while statistics based on sample surveys can be more frequent and up-to-date but lack geographical and other small domain details owing to sample size constraints. Sample survey results are widely used to derive reliable estimates for totals and means for large areas and domains. However, despite great developments in survey capability and practice, the usual direct survey estimators for a small domain, based only on observations of sampled units in the domain, are likely to yield unacceptably large sampling errors owing to the limited sample sizes involved.

939.   Nevertheless, demands are growing everywhere for more timely and varied statistics for lower-level administrative units and other small domains. Those needs can be met only by producing estimates that are both more frequent and more up-to-date than can be provided by censuses conducted at long intervals but which can be classified in much greater detail than is possible from sample surveys of limited size. Small area estimation methods, aimed at providing current yet detailed information, have been developed for that purpose. The basic idea of the procedures is to borrow and combine the relative strengths of more than one source in the production of better (more accurate and hence more useful) estimates.

940.   In the past few decades, there have been major and sophisticated developments in the methodology of small domain estimation procedures. The objective of chapter XVIII is not to review or evaluate those procedures, but merely to provide an introduction to the issues and to some basic approaches.[1] Any efforts to undertake small-area estimation should be examined thoroughly as there may be limits to the capacity to produce estimates for a small domain based on the techniques described in the present chapter. However, the approach can maximize the benefits of including questions on economic characteristics in the population census, and it increases the incentive to integrate the labour statistics system.

B.   What are small domains?

941.   First, it should be noted that the terms "domain" or "estimation domains" are used to refer to the population or any part of the population for which separate statistics are required. The terms "small area" or "local area" are commonly used to denote a geographical area that is considered small in some sense, such as provinces, counties, districts, other smaller administrative divisions, localities, or even census divisions or enumeration areas. The term "domain" is more general; it may refer to geographical areas or to other subpopulations of interest such as specific groups by age, sex, nationality, race or other characteristics.

[1] For an appraisal of recent developments in the field, see, for instance: M. Ghosh and J. N. K. Rao (1994); and J. N. K. Rao (1999). See also C. Elbers, J. O. Lanjouw and P. Lanjouw (2002); and J. N. K. Rao (2005).

942.   Secondly, it is important to note that the term "small" does not refer necessarily to smallness of the population size of the domain of interest, but to the smallness of the sample (number of observations) available on it. Small in one context may be very different from small in another context. For example, in statistically more developed countries where large and frequent sample surveys and/or administrative sources are available, the reference may be to very local areas or small population groups. In many less developed countries, the production of useful statistics even for large provinces or districts may require special small domain estimation techniques since it is impossible to expand sufficiently the sample sizes of national surveys for the purpose. The possibilities and suitability of specific small domain estimation procedures may differ greatly between the two situations.

943.   Classification of domains according to size category can be useful in keeping the distinctions clear. Adapting a rough classification proposed by Kish (1980), it is possible to distinguish between the following:

(*a*)   **Major domains.** From approximately 1 to 10 divisions of the population, such as major regions of the country, major groups by occupation, industry, sex, age and so on. Sample surveys are usually designed to provide useful direct estimates for such divisions;

(*b*)   **Minor domains.** From about 10 to 100 divisions of the population, such as individual provinces in Indonesia and Thailand or individual districts in Kenya, or two-way classification by occupation and sex and age groups. In many developing countries, the primary interest is to extend the available statistics to this level of detail;

(*c*)   **Mini domains.** From approximately 100 to 1,000 divisions of the population, such as individual counties in the United States or China. The production of reliable estimates at this level of detail is often beyond the capacity of many countries;

(*d*)   **Rare domains.** These may be used to refer to roughly 1,000 or more subdivisions, such as rare populations or multi-way classifications.

C.   Diversity of methods

944.   Earlier methods of small area estimation focused on demographic methods for population estimation in post-censal years and on estimation of the size of labour force status categories and other such characteristics of interest. Many of those methods used current data from administrative registers in conjunction with related data from the latest population census. Essentially, those symptomatic estimation methods exploited the relationships between symptomatic variables (such as the locally recorded number of births, deaths, school enrolment, housing units and the like) and variables of interest to be estimated (such as the local population size). Generally, such methods were very specific to the situation and depended on the quality and type of administrative data available for the purpose.

945.   A second class of methods, the so-called synthetic estimation methods, borrow the structural relationships between variables available in detail, but which are not current, in the census. They then impose those relationships in some appropriate manner on the less detailed but more current survey data to produce estimates which are current and detailed at the same time. The quality of the resulting estimates depends on the validity (stability) of the imported (postulated) relationships.

946.   It can also be useful to combine indirect estimates with direct estimates from the sample, to construct composite estimates as an appropriately weighted sum of direct and indirect estimates.

947.   More recent and sophisticated techniques include such procedures as the empirical Bayes, hierarchical Bayes and empirical best linear unbiased prediction procedures.[2] Those methods have made a significant impact on small area estimation practice in recent years.

948.   Some aspects and forms of the above-mentioned synthetic estimation techniques appear both feasible and practical for application in statistically less developed countries. The basic idea behind a synthetic estimation approach will be outlined in the next section (paras. 951-956). Beyond that, the procedures mentioned above will not be reviewed here any further. Rather, in this practical *Handbook*, it is useful to emphasize the correct approach to the development and use of small domain estimation procedures. The suitability of any technique is situation-specific and is determined by several criteria, as follows:

(*a*)   Availability of data;

(*b*)   Accuracy of the estimates;

(*c*)   Practicality;

(*d*)   Acceptability among the users.

949.   The following excerpt from Purcell and Kish (1980) summarizes a number of practical lessons in the application of such techniques:

(*a*)   From a variety of available methods several lessons may be learned. First, that one may find among them a better method than the one he is arbitrarily using at present for small domains; this is often the passive "null" method of continued reliance on the last decennial census that may be 12 years out of date;

(*b*)   Second, there is no single method that is best for all situations. Great differences between countries exist in the sources and quality of data available; the scope and quality of its census; the extents, contents and sizes of its sample surveys; and especially the scope and quality of its administrative registers. However, passive and negative attitudes are generally unjustified since every country has some resources, and ingenuity and effort can find unused resources of data. These may be of apparently different origins, but potentially useful because of high correlation with population sizes;

(*c*)   Furthermore, the choice of sources and methods should vary with and depend on the nature of the statistics, on the desired estimates and also on the domains to which they pertain. Note also the effects of the lapse of time since the last census. More generally, the balance between biases of a census and the variance of a large sample survey will move in favour of the latter during the 10 years between decennial censuses. The balance will also move in favour of less accurate but more current registers. The balance of sample surveys versus censuses or registers should depend on the sample size, but a fixed sample size has relative advantages in smaller populations;

(*d*)   Finally, the choice between methods is more difficult because the "best" is often not clear even after the event. Errors in the estimates arise from biases chiefly, and "true values" are usually not available for measuring the biases directly. Tests must be combinations of empirical and model bases, often depending eventually (and uncertainly) on results from decennial censuses. Better methods and criteria must be pursued with several methods and over the long range with an evolutionary approach and with patience.

950.   In the publication *Small Area Statistics: An International Symposium*, Kish (1987) sketches the bare outline of present and future developments, as follows:

(*a*)   "Methods are useful and used now for post-censal estimates for local area statistics;

(b)   "These methods will be used for other statistics also and in other situations;

(c)   "The present methods can and will be improved;

(d)   "The relative strength of different methods is difficult to predict and it depends on specific circumstances; they may be discovered in specific empirical trials;

(e)   "Success depends on first using better data and second on better methods. … Good data sources are the principal means to better statistics … [it is necessary to] work towards the collection of other and better data [and also to consider] strategies for cumulating data from samples for small areas".

1.   Illustration of a procedure

951.   This section provides a simple illustration of a potentially useful synthetic approach. Consider that a sample provides a current estimate of some quantity or count (such as numbers by activity status) y_g by sex and age group (g). From the census, a detailed but less current distribution of the population Y_{gh} by sex-age (g) and small area (h) is known. The latter provides an estimate of the distribution of a sex and age group by small area:

$$\left[Y_{gh}/Y_g \right].$$
[3.1]

952.   On the assumption that the distribution (the structural relationship) is still valid, small area estimates of variable y of interest are given by summing over all sex and age groups g, as follows:

$$y_h = \sum_g \left(\frac{Y_{gh}}{Y_g} \right) y_g.$$
[3.2]

953.   The basic idea of a more general form of the procedure, termed "structure preserving estimation" (SPREE) by its authors (Purcell and Kish, 1980), is illustrated in figure VI.

954.   Suppose that in the census, a detailed but less current distribution X_{igh} of activity status (i) by age group (g) and small area (h) is available. A current but less detailed distribution $Y_{ig.}$ of activity status (i) by age group (g) is available from the sample. This corresponds to the marginal distribution of the full census distribution, summed across small areas in the country. The interest is in estimating the distribution $Y_{i.h}$ of activity status (i) by small area (h), in other words, the other margin of the full census distribution as shown in the diagram. This estimate is given by

$$Y_{i.h} = \sum_g \left(\frac{X_{igh}}{X_{ig}} \right) Y_{ig.}.$$
[3.3]

955.   This formula is based on the assumption that the term in the parentheses, namely the distribution of economic activity status by age group (i by g) across small domains (h), has remained unchanged since the census. The corresponding distribution is imposed on each current (sample) value by i and g, and then summed over all age groups to obtain small area estimates of economic activity status (without classification by age group of course).

956.   Different forms of the procedure can be developed depending on the completeness of the information available from the census and on the types of constraints implied by the information available from the survey.

Figure VI
Illustration of a small area estimation procedure
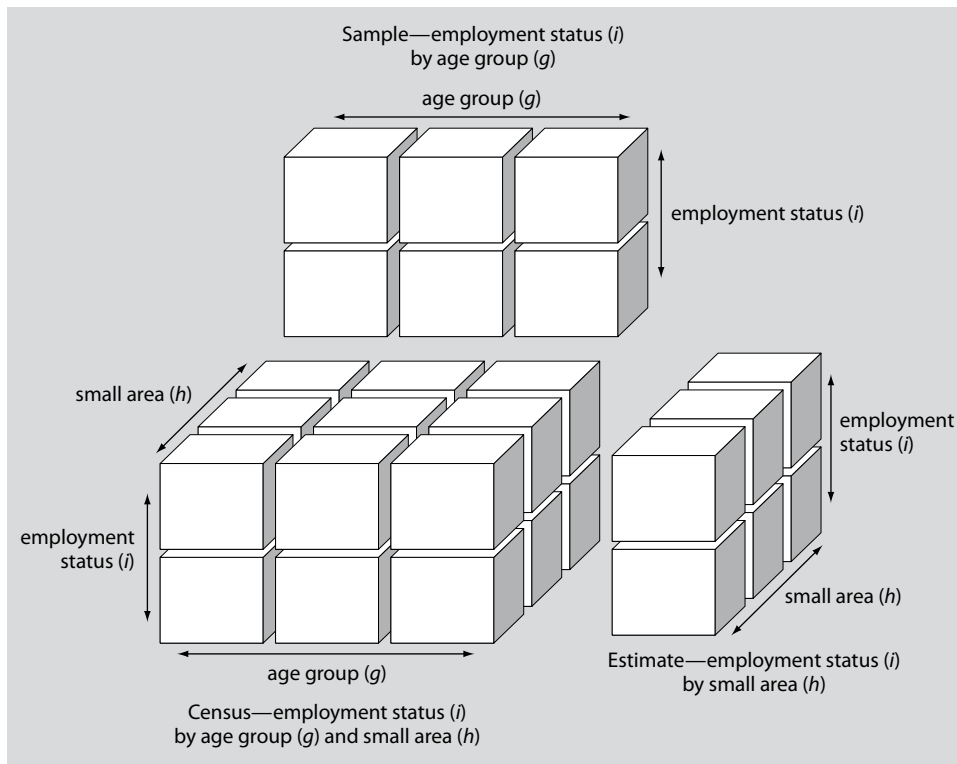
Sample—employment status (*i*)
by age group (*g*)

age group (*g*)

employment status (*i*)

small area (*h*)

employment
status (*i*)

age group (*g*)

Census—employment status (*i*)
by age group (*g*) and small area (*h*)

employment
status (*i*)

small area (*h*)

Estimate—employment status (*i*)
by small area (*h*)

Note: The term "employment status" used in figure VI is referred to as "economic activity status" in current international standards for labour statistics and elsewhere in the present handbook.

Annexes

Annex I

Examples of complete sets of economic activity questions in population censuses

The website of the United Nations Statistics Division shows copies of population census questionnaires where they have been provided by countries and territories.[1] It is difficult to find census questionnaires in which the economic activity questions cover all the topics included in *Principles and Recommendation for Population and Housing Censuses, Revision 2* (United Nations, 2008b). The incomplete coverage of topics arises in part from other competing demands for space on the census questionnaire. Another reason is that some of the topics (and changes in conceptual boundaries) have been added since the 2000 census round and thus few countries have had a chance to even consider adopting them. A few reasonably comprehensive examples of the economic block of questions are presented in annex I to show how they have been structured.

A wide range of examples from developing countries have been included. The examples are presented in three groups (see sections A, B and C below), following the format of the questionnaire from which they were extracted.

In addition, question sets from developed countries are presented separately as a fourth group (section D). The sets of questions from those countries are quite lengthy, and in most cases they will take far too much space for use in the census in countries that represent the major audience for the present *Handbook*. However, they present a good basis for other countries considering a review of their questions and/or the testing of new questions for experimentation with other formulations adapted to their specific contexts.

The four groups of sample questionnaires are organized as follows:

A.   Questionnaires using the landscape format, typically using one or two pages for several individuals within a specific household, with the listing of individuals along the rows and the questions presented in the columns;

B.   Questionnaires using the "type 1" portrait format, typically using one or more pages for several individuals within a specific household, with the listing of individuals along the columns and the questions presented along the rows;

C.   Questionnaires using the "type 2" portrait format, with one or more pages devoted to an individual, and the question set repeated for each individual in the household;

D.   Questionnaires from developed countries. The questionnaires from which the block of questions was extracted follow either the type 1 or the type 2 portrait formats described in B and C.

A.   Questionnaires using the landscape format

The landscape format has been particularly favoured by countries in Africa, which may have very few resources for censuses and may therefore place all questions on one side of a large questionnaire. In most cases, the topics included and the space allowed for

1 Available from http://unstats.un.org/unsd/demographic/sources/census/censusquest.htm.

each topic are severely restricted. Six examples of the block of economic questions in landscape format are presented, as follows:

Botswana, 2001. There are total of six questions, three on activity status (long-term activity, activity in the past seven days and job seeking); and one each on status in employment, occupation and industry.

China, 2000. There are six questions. The first question is used to identify those currently working for pay. The next three questions attempt to measure time worked, industry and occupation. The last two questions are addressed to those not currently working, asking the reasons for not working and source of livelihood. The set of questions is shown in Chinese and in English translation.[2]

Jordan, 2004. There are 10 questions: two on employment, two on industry, two on unemployment and one each on hours worked, occupation, status in employment and type of inactivity. The questionnaire shown is bilingual (Arabic and English).

Malawi, 1998. There are three columns. The first column combines current activity (including active, inactive, unemployment and first-time job seekers) and status in employment. The remaining two columns deal with occupation and industry.

Namibia, 2001. There are four questions, the first dealing with activity status (unemployed and employed) as well as reason for inactivity. The other three questions cover occupation, status in employment and industry.

South Africa, 2001. The questionnaire devotes 10 columns to economic characteristics: the first two deal with activity status and reason for inactivity; the next two deal with unemployment, and the next four cover status in employment, industry (two related questions) and occupation. The last two are on hours worked and place of work.

2 Translated by the Statistics Division of the United Nations Secretariat.

Botswana 2001

ALL PERSONS 12 YEARS AND OVER					
ECONOMIC ACTIVITY			**OCCUPATION**	**INDUSTRY**	**JOB SEEKER**
What has ... been doing mainly since Independence day 1999? **Seasonal work** 01 Paid 02 Unpaid **Non-seasonal work** 03 Paid 04 Unpaid **Other** 05 Job Seeker 06 Home Maker 07 Student 08 Retired 09 Sick ... Other (specify)	What was ... doing mainly in the past 7 days? 01 Worked for payment (Cash or in-kind) 02 Worked in family business (unpaid) 03 Worked at lands/cattle post (unpaid) 04 Job seeker: GO TO A26 05 Home maker 06 Student 07 Retired 08 Sick ... Other (specify)	GO TO A23 **IF FEMALE** GO TO A27 **If male end interview**	What was ... working as during the past 7 days? 01 Employee - Paid cash 02 Employee - Paid in kind 03 Self-employed (no employees) 04 Self-employed (with employees) 05 Members of producers' cooperatives 06 Working in family business (unpaid) 07 Working at land/cattle posts (unpaid) 08 Apprentices	What type of work did ... do in the past 7 days? (Probe as necessary, use two or more works to describe the occupation)	What was the main product, service or activity of ... place of work? (Probe as necessary, use two or more works to describe the industry) GO TO A27 for female	What steps did ... take during the last 30 days to seek work? 1 Register at Labour Office 2 Applied directly 3 Visited work sites 4 Asked friends/ relatives Other (specify)
A21(2)	A22(2)		A23	A24	A25	A26

China 2000

15 周 岁 及 以 上 的 人 填 报					
R17. 是否有工作	R18. 工作时间	R19. 行业	R20. 职业	R21. 未工作者状况	R22. 未工作者主要生活来源
10月25－31日是否从事过一小时以上有收入的工作： 1. 是 2. 在职休假、培训、季节性歇业未工作 →R19 3. 其他原因未工作 →R21	10月25－31日从事有收入的工作时间： 1. 一天 2. 二天 3. 三天 4. 四天 5. 五天 6. 六天 7. 七天	填写单位全称： 主要产品或业务范围：	做什么具体工作： →R23	1. 在校学生 2. 料理家务 3. 离退休 4. 丧失工作能力 5. 从未工作正在找工作 6. 失去工作正在找工作 失去工作前的主要职业： ———— 7. 其他	1. 退休金 2. 领取基本生活费 3. 家庭其他成员供养 4. 财产性收入 5. 保险 6. 其他
10月25－31日是否从事过一小时以上有收入的工作： 1. 是 2. 在职休假、培训、季节性歇业未工作 →R19 3. 其他原因未工作 →R21	10月25－31日从事有收入的工作时间： 1. 一天 2. 二天 3. 三天 4. 四天 5. 五天 6. 六天 7. 七天	填写单位全称： 主要产品或业务范围：	做什么具体工作： →R23	1. 在校学生 2. 料理家务 3. 离退休 4. 丧失工作能力 5. 从未工作正在找工作 6. 失去工作正在找工作 失去工作前的主要职业： ———— 7. 其他	1. 退休金 2. 领取基本生活费 3. 家庭其他成员供养 4. 财产性收入 5. 保险 6. 其他
10月25－31日是否从事过一小时以上有收入的工作： 1. 是 2. 在职休假、培训、季节性歇业未工作 →R19 3. 其他原因未工作 →R21	10月25－31日从事有收入的工作时间： 1. 一天 2. 二天 3. 三天 4. 四天 5. 五天 6. 六天 7. 七天	填写单位全称： 主要产品或业务范围：	做什么具体工作： →R23	1. 在校学生 2. 料理家务 3. 离退休 4. 丧失工作能力 5. 从未工作正在找工作 6. 失去工作正在找工作 失去工作前的主要职业： ———— 7. 其他	1. 退休金 2. 领取基本生活费 3. 家庭其他成员供养 4. 财产性收入 5. 保险 6. 其他

China 2000 (*English translation*)

For persons 15 years of age and over					
R17. Working or not	**R18. Amount of time worked**	**R19. Industry**	**R20. Occupation**	**R21. Reason for not working**	**R22. Primary source of livelihood for non-working persons:**
Indicate whether or not worked with pay for at least 1 hour between October 25 and October 31: 1. Yes 2. No, employed but did not work owing to taking leave, training or seasonal suspension of operations → R19 3. No, did not work for other reasons → R21	Indicate the amount of time worked with pay between October 25 and October 31: 1. 1 day 2. 2 days 3. 3 days 4. 4 days 5. 5 days 6. 6 days 7. 7 days	Please provide the complete name for the work unit: Major products or services:	Describe the specific work → R23	1. Students at school 2. Managing household affairs 3. Retired 4. Lost the ability to work 5. Never worked before, currently seeking employment 6. Lost previous job, currently seeking employment. Please describe main occupation prior to unemployment: _____ _____ 7. Other	1. Pension 2. Drawing basic living expenses 3. Supported by other family members 4. Income from property 5. Insurance 6. Other
Indicate whether or not worked with pay for at least 1 hour between October 25 and October 31: 1. Yes 2. No, employed but did not work owing to taking leave, training or seasonal suspension of operations → R19 3. No, did not work for other reasons → R21	Indicate the amount of time worked with pay between October 25 and October 31: 1. 1 day 2. 2 days 3. 3 days 4. 4 days 5. 5 days 6. 6 days 7. 7 days	Please provide the complete name for the work unit: Major products or services:	Describe the specific work → R23	1. Students at school 2. Managing household affairs 3. Retired 4. Lost the ability to work 5. Never worked before, currently seeking employment 6. Lost previous job, currently seeking employment. Please describe main occupation prior to unemployment: _____ _____ 7. Other	1. Pension 2. Drawing basic living expenses 3. Supported by other family members 4. Income from property 5. Insurance 6. Other
Indicate whether or not worked with pay for at least 1 hour between October 25 and October 31: 1. Yes 2. No, employed but did not work owing to taking leave, training or seasonal suspension of operations → R19 3. No, did not work for other reasons → R21	Indicate the amount of time worked with pay between October 25 and October 31: 1. 1 day 2. 2 days 3. 3 days 4. 4 days 5. 5 days 6. 6 days 7. 7 days	Please provide the complete name for the work unit: Major products or services:	Describe the specific work → R23	1. Students at school 2. Managing household affairs 3. Retired 4. Lost the ability to work 5. Never worked before, currently seeking employment 6. Lost previous job, currently seeking employment. Please describe main occupation prior to unemployment: _____ _____ 7. Other	1. Pension 2. Drawing basic living expenses 3. Supported by other family members 4. Income from property 5. Insurance 6. Other

Jordan 2004

234	233	232	231	230	229	228	227	226	225

2. بيانـــــات أفراد الأسرة

For person 15 years of age and over   للأفـــراد الـــذيـــن أعمارهـــم 15 ســنـــة فأكثـــر

حالة النشاط الاقتصادي خلال السبعة أيام المنتهية بلحظة العد (خلال الأسبوع 9/25 – 2004/10/1)

إذن: ما علاقة (اسم الفرد) بالنشاط الاقتصادي الرئيسي خلال السبعة أيام المنتهية بتاريخ 2004/10/1 ؟ 1.طالب 2.مدبر المنزل 3.لـه"ميما دخـل أو إيراد 4.العرض/العجز/ كبر السن 5.أخرى(حدد)	هل بحث (اسم الفرد) عن عمل بشكل جدي خلال الأسابيع الأربعة المنتهية بتاريخ 2004/10/1 ؟ 1- نعم ← انتقل الى الفرد التالي 2- لا ← استمر	لو توفرت إلى (اسم الفرد) فرصة عمل خلال السبعة أيام المنتهية بتاريخ 2004/10/1 أو خلال الـ (15) يوما اللاحقة لها، هل كنت/ستكون مستعدا للبدء بذلك العمل؟ 1.نعم ← استمر 2.لا ← 234	الحالة العملية للفرد 1. مستخدم بأجر 2. صـاحب عمل مع وجود مستخدمين آخرين 3. يعمل لحسابه دون وجود مستخدمين آخرين 4. يعمل لدى الأسرة دون أجر 5. يعمل دون أجر إلى الفرد التالي	المهنة الرئيسية الحالية؟ العمل الذي يقوم به (اسم الفرد)	النشاط الاقتصادي الرئيسي الحالي للمؤسسة؟ ما أهم المنتجات أو السلع والخدمات التي تنتجها أو تقدمها المؤسسة	اسم المؤسسة التي يعمل فيها (اسم الفرد) حاليا؟	ما مجموع ساعات العمل الفعلية التي عملها (اسم الفرد) في جميع الأعمال خلال الأسبوع المنتهي بتاريخ 2004/10/1	هل كان لدى (اسم الفرد) عمل ولكنه لم يزاوله (أي كان غائبا عنه بشكل مؤقت) بسبب المرض، الإجازة، عطلة سفر، تقلص النشاط في المؤسسة، خلل مؤقت في المؤسسة خلال الأسبوع المنتهي بتاريخ 2004/10/1 1.نعم ← 228 2. لا ← 232	هل عمل (اسم الفرد) خلال الأسبوع المنتهي بتاريخ 2004/10/1 حتى ولو ساعة واحدة : - في أي عمل مقابل أجر؟ أو فـي مـصـلحة خـاصـة بملكها جزء منها؟ أو فـي مـصـلحة للأسرة دون أجر (مثل: كالعمل في مزرعة، بئر...؟ أو فـي أي عمل آخر؟ 1-نعم ← 227 2-لا ← استمر
×	×	×	×	××××	××××	×	××	×	×
The relationship of (Name) with major economic activity during the 7 days ended in October 1st 2004 1. Student 2. Housemaker 3. With means 4. Disabled 5. Other (specify)	Did (Name) actively search for work during the 4 weeks ended in October 1st 2004 1. Yes — next person 2. No — continue	If a job was offered during the 7 days ended in October 1st or in the next 15 days, would (Name) be prepared to start the job 1. Yes — continue 2. No — 234	Employment Status 1. Paid Employee 2. Employer 3. Own-account worker 4. Unpaid family worker 5. Unpaid worker **Skip to next person**	Major current occupation	Major current economic activity of the establishment	Name of establishment in which the person works	Total no. of hours worked by (Name) in all jobs during the week ended in October 1st 2004	Had (Name) any job, but he did not practice it (was temporarily absent) due to illness, vacation, holiday, travel, reduction in economic activity, temporary breakdown in the establishment during the week ended in October 1st 2004? 1. Yes to 228 2. No to 232	Did (Name) work during the week ended in October 1st 2004 even for one hour: - In any paid work - In a business completely or partially owned - In a business owned by the household without payment - In any other business 1. Yes to 227 2. No continue

Malawi 1998

PERSONS 10 YEARS OR MORE		
Economic Activity		
Type of activity last week	**Occupation**	**Industry**
	What is this person's main occupation?	**What is this person's trade or business?**
What was ... doing last 7 days? *Enter code*	*For example: teacher, nurse, driver, mechanic, carpenter, etc.*	*For example: construction, agriculture, transport, manufacturing, community services, etc.*
B17	**B18**	**B19**

ACTIVE
  Mlimi = 01
  Employee = 02
  Family business worker = 03
  Self-employed = 04
  Employer = 05
Unemployed
  Worked before, seeking work = 06
  Worked before, not seeking work = 07
  Never worked before, seeking work = 08

INACTIVE
  Non-worker, never worked before and not seeking work = 09
  Home worker = 10
  Student = 11
  Other = 12

If female, skip to B20. Otherwise go to the next person or to Part C.

Namibia 2001

E	For persons aged 8 years and above				
During the last 7 days prior to census night, did (name) work for at least one hour for pay, profit or family gain? 01 Yes, Worked 02 No, but has a job or a business 03 Unemployed (worked before) 04 Unemployed (first time jobseeker) 05 Student 06 Homemaker 07 Income recipient 08 Severely Disabled 09 Retired 10 Old age 11 Other, specify 99 Don't know	For those coded 01, 02 and 03 in E1				
	What type/kind of work did/does (name) do at his/her place of work? Write the main work done	For official use only	In your main job did you work as 01 Subsistence/ Communal farmer (with paid employees) 02 Subsistance/Communal farmer (without paid employees) 03 Other employer (with paid employees) 04 Other own account worker (without paid employees) 05 Employee (Government or parastatal) 06 Employee (private) 07 Unpaid family worker (Subsistance/ Communual farmer) 08 Other unpaid family worker 09 Other, specify 99 Don't know	What kind of activity is carried out at (name's) workplace ? Describe the activity	For official use only
E1	E2	E3	E4	E5	E6
	— — — — —			— — — — —	
	— — — — —			— — — — —	
	— — — — —			— — — — —	
	— — — — —			— — — — —	
	— — — — —			— — — — —	
	— — — — —			— — — — —	

South Africa 2001

ASK FOR ALL PERSONS AGED 10 YEARS AND OLDER (BORN BEFORE 10 OCTOBER 1991)					
ANY WORK IN THE 7 DAYS BEFORE 10 OCTOBER (P-18)	DID NOT HAVE ANY WORK			HAD WORK	
	REASON WHY NOT WORKING (P-18a)	ACTIVE STEPS (P-18b)	AVAILABILITY (P-18c)	WORK STATUS (P-19)	BUSINESS/COMPANY NAME (P-19a)
In the SEVEN DAYS before 10 October did (the person) do any work for PAY (in cash or in kind), PROFIT or FAMILY GAIN, for one hour or more? 1 = Yes: formal registered (non-farming) 2 = Yes: informal unregistered (non-farming) 3 = Yes: farming 4 = Yes: has work but was temporarily absent 5 = No: did not have work If YES go to P-19	If NO to P-18 What is the main reason why (the person) did not have work in the seven days before 10 October? 1 = Scholar or student 2 = Home-maker or housewife 3 = Pensioner or retired person/ too old to work 4 = Unable to work due to illness or disability 5 = Seasonal worker not working presently 6 = Does not choose to work 7 = Could not find work If more than one reason, write the code of the MAIN (most important) reason.	If NO to P-18 In the PAST FOUR WEEKS before 10 October has (the person) taken active steps to find employment? Y = Yes N = No For example, (the person) went to visit factories or other employment places, placed or answered advertisements, looked for land or a building or equipment to start own business or farm.	If NO to P-18 If offered work, how soon could (the person) start? 1 = Within one week 2 = More than 1 week, up to 2 weeks 3 = More than 2 weeks, up to 4 weeks 4 = Some time after 4 weeks 5 = Does not choose to work Go to P-20	If YES to P-18 How can one best describe (the person's) main activity or work status? 1 = Paid employee 2 = Paid family worker 3 = Self-employed 4 = Employer 5 = Unpaid family worker 6 = Other (specify)	If YES to P-18 What is the FULL name of the business/company or organisation for whom (the person) works? If the person works for him/herself, and the business does not have a name, write SELF in the appropriate row. If doing PAID domestic work in a private household, write DOMESTIC SERVICE. Use CAPITAL LETTERS only.

South Africa 2001 (*continued*)

ASK FOR ALL PERSONS AGED 10 YEARS AND OLDER (BORN BEFORE 10 OCTOBER 1991) WHO HAD WORK			
COMPANY/BUSINESS ACTIVITY **(P-19b)**	OCCUPATION **(P-19c)**	HOURS WORKED **(P-19d)**	PLACE OF WORK **(P-19e) (P-19f)**
If YES to P-18 **What does the business do (main economic activity)?** Write the MAIN INDUSTRY, economic activity, product or service of (the person's) employer or company. For example, gold mining, road construction, supermarket, police service, healthcare, hairdressing, banking. OR Write the activity of the person if self-employed. For example, subsistence farming. If doing PAID domestic work in a private household, write DOMESTIC SERVICE. Use CAPITAL LETTERS only.	If YES to P-18 **What is the main occupation of (the person) in this workplace?** Occupation refers to the type of work (the person) performed in the seven days before 10 October. Use two or more words. For example, street trader, cattle farmer, primary school teacher, domestic worker, fruit vendor, truck driver, warehouse manager, filing clerk, etc. Use CAPITAL LETTERS only.	If YES to P-18 **How many hours did (the person) work in the seven days before 10 October?** If (the person) was absent from work those seven days, but usually works, write the number of hours s/he usually works.	If YES to P-18 **Does (the person) work in the same sub-place in which s/he usually lives?** Y = Yes N = No Dot the appropriate box. If NO, **where is this place of work?** If NOT the same place, write PROVINCE *P R*, MAIN PLACE (city, town, tribal area, administrative area) and SUB-PLACE (suburb, ward, village, farm, informal settlement). If another country, write the name of the country in the boxes below.

B.   Questionnaires using the "type 1" portrait format

The type 1 portrait format,[3] which lists individuals along the columns and displays the questions in the rows, is a major improvement over the landscape format with respect to space, for both formulating questions and recording answers. Two examples of the block of economic questions in this format are presented, as follows:

Papua New Guinea, 2000. Six questions on economic characteristics were included in the census, referring to the last seven days. The questions include one on economic activity, two on occupation, one on industry, one on unemployment (whether sought work) and one on the main activity if not economically active.

Sri Lanka, 2001. The five questions in the census questionnaire deal with status in employment, occupation, industry, institutional sector of employment and principal non-economic activity, during the 12-month reference period.

[3] Several individuals are listed on one or more pages.

Papua New Guinea 2000

CONTINUE ONLY FOR PERSONS AGED 10 YEARS OR MORE

PERSON NUMBER ⟶	Person 01	Person 02	Person 03
15 **In the last seven days, has the person done any of the following?** *(If more than one answer is given record main activity.)* If code 7 "None" **GO TO Q.18**	1. Gardening/fishing for money 2. Gardening/fishing for own use only 3. Business (any type) with paid help 4. Business (any type) without paid help 5. Helping in family business without pay 6. Wage job (include temporary absence) 7. None **(GO TO Q.18)** ▢	1. Gardening/fishing for money 2. Gardening/fishing for own use only 3. Business (any type) with paid help 4. Business (any type) without paid help 5. Helping in family business without pay 6. Wage job (include temporary absence) 7. None **(GO TO Q.18)** ▢	1. Gardening/fishing for money 2. Gardening/fishing for own use only 3. Business (any type) with paid help 4. Business (any type) without paid help 5. Helping in family business without pay 6. Wage job (include temporary absence) 7. None **(GO TO Q.18)** ▢
16 (a) **What kind of work did the person do in the last seven days?** *(Describe person's work or occupation, e.g. plumber, streetseller, farmer, car mechanic.)*			
(b) **What are the main task/duties done in that work?** *(Describe kind of tasks or duties done in the work stated in Q.16(a). e.g. mending tap, selling betelnut, weeding garden, fixing engine.)*	▢▢▢▢	▢▢▢▢	▢▢▢▢
17 **What kind of business is done at the place where the person works?** *(Describe business by activity, e.g. house building, air transport, selling betelnut.)* **GO TO Q.20**	▢▢▢▢	▢▢▢▢	▢▢▢▢
18 **Has the person actively sought any kind of work in the last four weeks?**	**1. Yes 2. No** ▢	**1. Yes 2. No** ▢	**1. Yes 2. No** ▢
19 **What was the person doing mostly in the last seven days?** *(If other describe in the space provided.)*	1. Housework 2. Studying 3. Waiting to start a new job 4. Looking for work **Not working because** 5. Too young 6. Too old/retired 7. Sick 8. Permanently Disabled 9. Other, describe ▢	1. Housework 2. Studying 3. Waiting to start a new job 4. Looking for work **Not working because** 5. Too young 6. Too old/retired 7. Sick 8. Permanently Disabled 9. Other, describe ▢	1. Housework 2. Studying 3. Waiting to start a new job 4. Looking for work **Not working because** 5. Too young 6. Too old/retired 7. Sick 8. Permanently Disabled 9. Other, describe ▢

Sri Lanka 2001

	P 18	**ACTIVITIES IN THE LAST 12 MONTHS AND THEIR DURATIONS (IN WEEKS):** A. Paid employment B. Employer C. Own account worker D. Contributing to family enterprises (unpaid family worker) E. Available/ Seeking work F. Non economic activities (Student, house keeping, income recepient/pensioner, unable/too old to work, child not attending school, other non economic activity) If not engaged in any activity write '00' for that activity. The sum of all durations should be 52 weeks.	**Activity Weeks** A. Paid ☐☐ B. Emp. ☐☐ C. O.A.W. ☐☐ D. Family ☐☐ E. Available ☐☐ F. N.E.A. ☐☐	**Activity Weeks** A. Paid ☐☐ B. Emp. ☐☐ C. O.A.W. ☐☐ D. Family ☐☐ E. Available ☐☐ F. N.E.A. ☐☐	**Activity Weeks** A. Paid ☐☐ B. Emp. ☐☐ C. O.A.W. ☐☐ D. Family ☐☐ E. Available ☐☐ F. N.E.A. ☐☐
10 years & over **(For persons born on or before June 1991)**	P 19	**PRINCIPAL OCCUPATION OR KIND OF WORK DONE IN THE LAST 12 MONTHS :** **If P18A to P18D '00' go to P 22** State clearly exact occupation or kind of work done during the last 12 months. eg. Paddy cultivator, Tea plucker, Accounts clerk, Mathematics teacher, Vegetable seller (retail), Domestic helper etc. If the person had more than one occupation/work, give the principal occupation/work on which most time was spent.	If P18A to P18D '00' go to P 22 .... ☐☐☐☐	If P18A to P18D '00' go to P 22 .... ☐☐☐☐	If P18A to P18D '00' go to P 22 .... ☐☐☐☐
	P 20	**INDUSTRY, BUSINESS OR SERVICE :** Describe the kind of products made or services rendered at the work place of the principal occupation/ work Write the name of the work place. eg. Paddy cultivation, Tea cultivation, Education department, Textile shop (retail), Garment factory etc.	.... ☐☐☐☐	.... ☐☐☐☐	.... ☐☐☐☐
	P 21	**EMPLOYMENT STATUS AND SECTOR :** 1 Government employee 4. Employer 2. Semi government employee 5. Own account worker 3. Private sector employee 6. Unpaid family worker	1. Govt. 4. Emplo. 2. Semi 5. Own govt. acc.wor. 3. Pvt.sec. 6. Unpaid employ. fam.wor	1. Govt. 4. Emplo. 2. Semi 5. Own govt. acc.wor. 3. Pvt.sec. 6. Unpaid employ. fam.wor	1. Govt. 4. Emplo. 2. Semi 5. Own govt. acc.wor. 3. Pvt.sec. 6. Unpaid employ. fam.wor
	P 22	**PRINCIPAL NON ECONOMIC ACTIVITY IN THE LAST 12 MONTHS :** **If P18F is '00' go to P 23** 1 Student 4. Unable/Too old to work 2. Household work 5. Child not attending school 3. Income recipient/ Pensioner 6. Other non economic activity	If P18F is '00' go to P 23 1. Student 4. Unable/ 2. H.h. Too old work 5. Child 3. Inc.rec./ not att. Pensi. 6. Other	If P18F is '00' go to P 23 1. Student 4. Unable/ 2. H.h. Too old work 5. Child 3. Inc.rec./ not att. Pensi. 6. Other	If P18F is '00' go to P 23 1. Student 4. Unable/ 2. H.h. Too old work 5. Child 3. Inc.rec./ not att. Pensi. 6. Other

C.   Questionnaires using the "type 2" portrait format

The type 2 portrait format[4] provides the most ample space for both the questions and responses to them. The format is such that the set of questions for each household member is contained in two or more pages for each individual. Not many developing countries adopted this format in the 2000 round of censuses owing to cost considerations, but the number is increasing. Section C presents seven examples of the block of economic activity questions in the type 2 format, as follows:

Argentina, 2001. There are 13 questions dealing with economic activity. The first four are on activity status, employment, and unemployment. The next three cover industry and occupation (two questions). Status in employment and institutional sector are covered in three questions, and of the remaining questions, two deal with retirement contributions and one with the size of the establishment (with respect to the number of employees that it has). The set of questions is shown in Spanish and in English translation.[5]
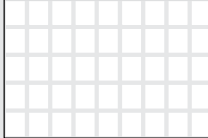
Bahamas, 2000. In the economic activity section of the Bahamas census questionnaire, there are 13 questions, and in a separate section on income, there is one question on the sources of income and amount of each. The first two questions in the economic activity section deal with activity status and unemployment. They are followed by one question on multiple jobs and one on hours worked. The reference period for those four questions is the past week. The next four questions are on activity status and number of weeks worked in a 12-month reference period. The final five questions are on industry (two questions), occupation, status in employment and informal trade.

Belize, 2000. The section on economic activity contains 14 questions. The first six questions deal with activity status, including one on hours worked and another on reasons for inactivity. Three questions deal with occupation, industry and status in employment (present for those currently employed and previous for those currently unemployed). In addition, one question concerns the name and address of the place of work and two questions relate to pay and income. Finally, there are two questions on remittances from abroad.

Jamaica, 2001. The block of questions presented is from the census questionnaire that was administered to a 10 per cent sample of the census population. The first four questions in the set of 20 questions cover activity status and the fifth covers hours worked. Status in employment, occupation and industry take up the next three, with the following two covering place of work. The size, name and address of the establishment take up another two questions, which are followed by a question on income from employment. For those not currently employed, there is one question on the time of last employment. Two questions cover usual activity status (12-month reference period), and two more deal with permanent layoffs. The two last questions are on benefits and pension.

Mexico, 2000. The block of questions presented is from the long form census questionnaire. There are 10 economic activity questions (although some actually consist of a pair of questions), broken down as follows: two on activity status, one each on occupation, status in employment, employment benefits, hours worked, labour income, industry, place of work and other sources of income (including the amount received from each source). The set of questions is shown in Spanish and in English translation.[6]

Turkey, 2000. The questionnaire has nine questions on economic activity: four relating to activity status (employment, unemployment, reasons for inactivity), two to occupation, two to industry and one to status in employment.

Zambia, 2000. There are five questions: one on current activity status (one-week reference period) and another on usual activity status (12-month reference period). The three remaining questions cover status in employment, occupation and industry in the 12-month reference period.

Argentina 2001

A PARTIR DE AQUI CONTESTAN LAS MUJERES Y LOS VARONES DE 14 AÑOS O MAS

24 ¿Durante la semana pasada, trabajó aunque sea por pocas horas?
- Sí
- No
- Ignorado

25 ¿Hizo algo en su casa para afuera o ayudó a alguien en un negocio, chacra o trabajo?
- Sí
- No
- Ignorado

26 Tenía trabajo pero, ¿estuvo de licencia por enfermedad, vacaciones, etc.?
- Sí
- No
- Ignorado

27 Durante las últimas cuatro semanas, ¿buscó trabajo?
- Sí
- No
- Ignorado

EN EL TRABAJO PRINCIPAL, EN EL QUE TRABAJA MAS HORAS...

28 El lugar o establecimiento donde trabaja, ¿a qué actividad se dedica o qué servicio presta?

29 ¿Cuál es el nombre de la ocupación?

30 ¿Qué tareas hace en ese trabajo?

31 En el trabajo principal, ¿es...
- obrero o empleado?
- patrón?
- trabajador por cuenta propia?
- trabajador familiar?

32 ¿Trabaja en ...
- un empleo estatal (nacional, provincial o municipal)?
- un empleo del sector privado?
- Ignorado

33 Por ese trabajo, ¿le descuentan para la jubilación?
- Sí → **36**
- No
- Ignorado

34 ¿En ese trabajo aporta por sí mismo para la jubilación?
- Sí
- No
- Ignorado

35 ¿Recibe sueldo?
- Sí → **33**
- No
- Ignorado

36 ¿Cuántas personas hay en total en el establecimiento o lugar donde trabaja?
- 1 a 5
- 6 a 39
- 40 o más
- Ignorado

Si es Mujer: Siga en Preg. 37
Si es Varón: Fin de la Persona

Argentina 2001 (*English translation*)

THE FOLLOWING SECTION IS FOR WOMEN AND MEN 14 YEARS OF AGE OR OLDER

24 Did the person work last week, even for a few hours?

Yes
No
Unknown

25 Did the person do any work from home, or help anyone in a business, farm or job?

Yes
No
Unknown

26 Did the person have a job, but was absent from work due to illness, vacations, etc.?

Yes
No
Unknown

27 Has the person looked for work over the last four weeks?

Yes
No
Unknown

IN THE PERSON'S MAIN JOB, IN WHICH HE/SHE WORKS MOST HOURS…

28 The place or establishment in which the person works, what activity does the business perform or what service does it provide?

29 What is the name of the occupation?

30 What tasks does the person perform in that job?

31 In his/her main job, the person is …

Worker or employee

Owner

Own-account worker

Family worker

32 Does the person work for …

The governement (national, provincial or municipal)

The private sector

Unknown

33 In that job, does the person's employer make deductions for retirement?

Yes → **36**
No
Unknown

34 In that job does the person make his/her own retirement contributions?

Yes
No
Unknown

35 Does the person receive a wage?

Yes → **33**
No
Unknown

36 How many people altogether are there in the establishment where the person works?

1 to 5
6 to 39
40 or more
Unknown

If the person is a woman, proceed to question 37. If the person is a man, stop here.

Bahamas 2000

SECTION 8: ECONOMIC ACTIVITY (PERSONS 15 YEARS AND OVER)

P33. DID YOU DO ANY WORK AT ALL, FOR ANY LENGTH OF TIME DURING THE WEEK

OF APRIL 24TH THRU 30TH

(This would include helping in a family business/farm, street vending etc.)

1 ☐ YES **(Skip to P35)**   2 ☐ NO

P34. WHAT WAS YOUR MAIN ACTIVITY DURING THAT WEEK?

01 ☐ HAD A JOB BUT DID   05 ☐ VOLUNTARY WORK
NOT WORK   WITHOUT PAY

02 ☐ LOOKED FOR WORK   06 ☐ HOME DUTIES

03 ☐ LOOKED FOR WORK DURING   07 ☐ STUDENT
THE PAST 4 WEEKS

04 ☐ DID NOT LOOK BUT WANTED   08 ☐ RETIRED
TO WORK AND WAS AVAILABLE
DURING PAST 4 WEEKS   09 ☐ DISABLED

10 ☐ OTHER _____ (Specify)

(ALL PERSONS ANSWERING P34 SKIP TO P37)

P35. DURING THAT WEEK, HOW MANY PAID JOBS DID YOU WORK AT?

1 ☐ ONE   3 ☐ THREE OR MORE

2 ☐ TWO   4 ☐ NONE

P36. HOW MANY HOURS DID YOU WORK ON YOUR MAIN JOB DURING THAT WEEK?

1 ☐ 1 – 8   4 ☐ 33 – 44

2 ☐ 9 – 15   5 ☐ 45 & OVER

3 ☐ 16 – 32   6 ☐ NONE

(ALL PERSONS ANSWERING P36 TICK 1 AT P37 AND THEN SKIP TO P40)

**P37. DID YOU WORK AT ALL, FOR ANY LENGTH OF TIME DURING THE PAST
TWELVE MONTHS? (This would include selling newspapers, peanuts and other
items, helping in a family business or farm, summer employment etc.)**

1 ☐ YES **(Skip to P40)**   2 ☐ NO

P38. WHAT WAS YOUR MAIN ACTIVITY DURING THE PAST 12 MONTHS?

01 ☐ HAD A JOB BUT DID NOT WORK **(Skip to P40)**

02 ☐ LOOKED FOR WORK

03 ☐ DID NOT LOOK BUT WANTED WORK AND WAS AVAILABLE

04 ☐ VOLUNTARY WORK WITHOUT PAY

05 ☐ HOME DUTIES

06 ☐ STUDENT   } **(Skip to P46)**

07 ☐ RETIRED (Skip to P46)

08 ☐ DISABLED

09 ☐ OTHER _____ **(Specify)**

P39. HAVE YOU EVER WORKED OR HAD A JOB FOR AT LEAST TWO (2) WEEKS?

1 ☐ YES   2 ☐ NO

(ALL PERSONS ANSWERING P39 SKIP TO P46)

P40. HOW MANY WEEKS DID YOU WORK IN THE PAST TWELVE MONTHS?

1 ☐ 1 – 4   5 ☐ 40 – 48

2 ☐ 5 – 13   6 ☐ 49 – 52

3 ☐ 14 – 26   7 ☐ NONE

4 ☐ 27 – 39

**P41. WHAT IS THE NAME OF THE COMPANY/BUSINESS WHERE YOU WORK OR FOR
WHICH YOU LAST WORKED?
(This and following questions refer to main job)**

_____

**P42. WHAT KIND OF BUSINESS OR ACTIVITY TAKES PLACE THERE?
(Describe the kind of business e.g. retail store, primary school, law firm, brewery,
etc.)**

_____

☐ ☐ ☐ ☐

P43. WHAT TYPE OF WORK DO YOU/DID YOU DO?

(Describe your job as accurately as possible e.g. sales clerk, typist, doctor, auto
mechanic, civil engineer, taxi driver, housemaid, etc.)

(Do Not Say engineer, mechanic, teacher, supervisor, clerk, etc. Be more specific e.g.
sales clerk, primary school teacher, auto mechanic, etc.)

_____

☐ ☐ ☐ ☐

**P44. ARE/WERE YOU SELF-EMPLOYED OR WORKING FOR SOMEONE ELSE IN YOUR
MAIN JOB?**

1 ☐ SELF-EMPLOYED (NO PAID HELPER)

2 ☐ SELF-EMPLOYED (1-4 PAID HELPERS)

3 ☐ SELF-EMPLOYED (5 OR MORE PAID HELPERS) ⎤

4 ☐ EMPLOYEE (GOVT/GOVT CORP)

5 ☐ EMPLOYEE (PRIVATE – 1-4 WORKERS)   ⎬ **(Skip to P46)**

6 ☐ EMPLOYEE (PRIVATE – 5 OR MORE WORKERS) ⎦

7 ☐ UNPAID FAMILY WORKER

**P45. DO YOU MOVE ALL YOUR GOODS DAILY; E.G. FRUITS, PEANUTS, NEWSPAPERS,
CLOTHING, EQUIPMENT?**

1 ☐ YES (INFORMAL TRADER)   2 ☐ NO

SECTION 9: INCOME (PERSONS 15 YEARS OF AGE AND OVER)

**P46   DURING THE PAST TWELVE MONTHS DID YOU RECEIVE INCOME FROM ANY OF THESE SOURCES? IF SO, STATE AMOUNT IN THE SPACE PROVIDED.
(B$ to the nearest whole number e. g. 12565.80 = 12566)**

(PRIMARY JOB)

1   WAGES, SALARY, COMMISSION, TIPS, ETC.   4   RETIREMENT PENSION   7   GIFTS AND DONATIONS

☐☐☐☐☐☐   ☐☐☐☐☐☐   ☐☐☐☐☐☐

(SECONDARY JOB)

2   WAGES, SALARY, COMMISSION, TIPS, ETC.   5   OLD AGE PENSION   8   INVESTMENTS

☐☐☐☐☐☐   ☐☐☐☐☐☐   ☐☐☐☐☐☐

3   OWN BUSINESS   6   GOVERNMENT ALLOWANCES   9   OTHER SOURCES

☐☐☐☐☐☐   ☐☐☐☐☐☐   ☐☐☐☐☐☐

10   TOTAL INCOME DURING LAST TWELVE MONTHS

☐☐☐☐☐☐☐

Belize 2000

SECTION 11 ECONOMIC ACTIVITY   FOR PERSONS 14 AND OVER   PERSON 1

11.1 Do you/Does. . . . own any land for farming?
① Yes   ② No   ⑨ Don't know/Not stated

11.2 Did you/. . . . do any work for pay, profit, or family gain for at least one hour during the week ending May 12th? This includes helping in a family business or farm, street vending, or work at home.
① Yes   (SKIP TO Q 11.5)   ⑨ Don't know/Not stated
② No

11.3 Did you/. . . . engage in any economic activity on the following list, for pay, profit or family gain, for at least one hour, during the week ending May 12th?
① Yes   (SKIP TO Q 11.5)   ⑨ Don't know/Not stated
② No

INTERVIEWER: READ THE FOLLOWING LIST

Babysitting   Nurse's Aid for pay
Laundry, ironing for pay   Subsistence farming
Cleaning yard/cutting yard

11.3 list continued.
Selling food or snacks   Sewing for pay
   at market/bus-stop/school   Cleaning of offices
Bicycle/cart deliveries   Car washing
Selling food from home   Taxi-driver
Selling sweets from   Lottery vendors
   home (fudge, etc.)   Any other similar activity

11.4 Did you/. . . . have a job during that week, from which you were temporarily absent?
① Yes
② No   (SKIP TO Q 11.6)
⑨ Don't know/Not stated (SKIP TO Q 11.6)

11.5 How many hours did you/. . . . . work during that week?
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
99 Don't know/Not stated
(SKIP TO Q 11.7a)

Belize 2000 (*continued*)

SECTION 11 ECONOMIC ACTIVITY FOR PERSONS 14 AND OVER PERSON 1

11.6 If you/.... had been offered a job that week what would have prevented you/.... from taking up that job?
① Nothing ④ Retirement
② School ⑤ Not interested in working
③ Home duties ⑥ Other (Specify _____)
⑨ Don't know/Not stated

11.7 Have you/Has . . . . . ever worked
① Yes (SKIP TO Q 11.7b)
② No (SKIP TO Q 11.13)
⑨ Don't know/Not stated (SKIP TO Q 11.13)

11.7a What sort of work do you/does . . . . . do in your/his/her (main) occupation? Please specify in detail.

```
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
(SKIP TO Q 11.8a)
```

11.7b What sort of work did you/ . . . . . do in your/his/her previous occupation? Please specify in detail.

```
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
(SKIP TO Q 11.8b)
```

11.8a What type of business is carried on at your/ . . . . . .'s (main) workplace? Please specify in detail.

```
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
(SKIP TO Q 11.9)
```

11.8b What type of business was carried on at your/ . . . . .'s previous workplace? Please specify in detail.

```
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
(SKIP TO Q 11.10b)
```

11.9 What is the name and address of your/ . . . . .'s (main) workplace?

```
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
0 1 2 3 4 5 6 7 8 9
```

11.10a Do you/Does . . . . . carry on your/his/her own business, work for a wage or salary or as an unpaid worker?
① Paid employee – Government (SKIP TO Q 11.11)
② Paid employee – Private (SKIP TO Q 11.11)
③ Unpaid worker (SKIP TO Q 11.13)
④ Own business with paid help (Employer) (SKIP TO Q 11.12)
⑤ Own business without paid help (Own Account) (SKIP TO Q 11.12)
⑨ Don't know/Not stated (SKIP TO Q 11.13)

11.10b Did you/ . . . . . carry on your/his/her own business, work for a wage or salary or as an unpaid worker?
① Paid employee – Government
② Paid employee – Private
③ Unpaid worker
④ Own business with paid help (Employer)
⑤ Own business without paid help (Own Account)
⑨ Don't know/Not stated
SKIP TO Q 11.13

11.11 How often do you/ . . . . . get paid?
① Daily ⑥ Annually
② Weekly ⑦ Other
③ Fortnightly (Specify _____)
④ Monthly ⑨ Not stated
⑤ Quarterly

11.12 What was . . . . .'s gross pay/income during the last pay period, that is before deductions?

(PRESENT FLASH CARD)

INTERVIEWER: FOR SELF-EMPLOYED PERSONS OBTAIN "NET INCOME," I.E., RECEIPTS LESS BUSINESS EXPENSES.

```
0 0
1 1
2 2
3 3
4 4
5 5
6 6
7 7
8 8
9 9
```

11.13 Do you/does he/she receive any money from family and/or friends abroad?
① Yes
② No (END PRESENT INTERVIEW)
39 Don't know/Not stated

11.14 Approximately how much money did you/he/she receive last year (. . . .) from family and/or friends abroad?

(PRESENT FLASH CARD)

```
0 0
1 1
2 2
3 3
4 4
5 5
6 6
7 7
8 8
9 9
```

Jamaica 2001

PC01C   *Confidential*

S E C T I O N 4

ECONOMIC ACTIVITY

(For persons 14 years and over)

4.1 Did you/did .... work for at least one hour during the first week of September 2001 ?

○ Yes *(Go to Q4.5)*   ○ No   ○ Not Stated

4.2 Did you/did .... do anything like farming, buying and selling during the first week of September 2001 ?

○ Yes *(Go to Q4.5)*   ○ No   ○ Not Stated

4.3 Did you/did .... do any type of odd job or hustling during the first week of September 2001 ?

○ Yes *(Go to Q4.5)*   ○ No   ○ Not Stated

4.4 What were you/was .... doing for most of the time during the first week of September 2001 ? (READ CATEGORIES)

○ Working in Agriculture or any other business without pay

○ With job not working *(Go to Q4.6)*

○ Seeking first job *(Go to Q4.15)*

○ Seeking a job which was not the first *(Go to Q4.7)*

○ Did not seek work but wanted work and was available *(Go to Q4.7)*

○ Student *(Go to Q4.14)*

○ Did Home Duties *(Go to Q4.14)*

○ Retired did not work *(Go to Q4.14)*

○ Disabled unable to work *(Go to Q4.14)*

○ Not interested in work *(Go to Q4.14)*

○ Other *(Go to Q4.14)*

○ Not Stated

4.5 How many hours did you/did .... work during the first week of September 2001 ?

[|]   ○ Not Stated

4.6 Which of the following categories best describes your/ ....'s main employment? (READ CATEGORIES)

○ Paid Government Employee

○ Paid Employee in Private Enterprise

○ Paid Employee in Private Home

○ Unpaid Employee in Agriculture or in any other type of business

○ Self Employed with Employees

○ Self Employed without Employees

○ Other

○ Not Stated

4.7 What kind of work do you do/does .... do/did you last do/did .... last do ?

_____

[| | |]

○ Never Worked *(Go to Q4.19)*   ○ Not Stated

4.8 What type of business is/was carried on at the work place ?

_____

[| | |]   ○ Not Stated

> **GO TO Q 4.14 IF RESPONDENT IS NOT CURRENTLY EMPLOYED OTHERWISE CONTINUE**

4.9 Where is your/is ....'s place of work located ? (READ CATEGORIES)

○ In own home or yard   ○ On a Farm

○ In another home or yard   ○ Not Stated

○ Not in a private home

4.10 In which parish do you/does .... work ?

○ Kingston   ○ Hanover

○ St. Andrew   ○ Westmoreland

○ St. Thomas   ○ St. Elizabeth

○ Portland   ○ Manchester

○ St. Mary   ○ Clarendon

○ St. Ann   ○ St. Catherine

○ Trelawny   ○ More than one parish

○ St. James   ○ Not Stated

Jamaica 2001 (*continued*)

4.11 How many persons including yourself / including .... are working in the business or at the work place?

- ○ 1 person
- ○ 10 - 19 persons
- ○ 2 - 4 persons
- ○ 20 + persons
- ○ 5 - 9 persons
- ○ Not Stated

4.12 What is the Name and Address of the Business Establishment where you/where .... work(s) ?

_____

_____

- ○ Don't Know
- ○ Not Stated

4.13 What is your/is .... 's weekly, monthly or annual income from all employment? ($JA)

○ Not Stated

Weekly
- ○ Less than 1,000
- ○ 10,000 - 19,999
- ○ 1,000 - 1,499
- ○ 20,000 - 29,999
- ○ 1,500 - 5,999
- ○ 30,000 - 59,999
- ○ 6,000 - 9,999
- ○ 60,000 and over

Monthly
- ○ Less than 3,500
- ○ 40,000 - 79,999
- ○ 3,500 - 5,999
- ○ 80,000 - 129,999
- ○ 6,000 - 24,999
- ○ 130,000 - 249,999
- ○ 25,000 - 39,999
- ○ 250,000 and over

Annually
- ○ Less than 40,000
- ○ 500,000 - 999,999
- ○ 40,000 - 79,999
- ○ 1,000,000 - 1,499,999
- ○ 80,000 - 299,999
- ○ 1.5 million - 2,999,999
- ○ 300,000 - 499,999
- ○ 3 million and over

▶ **(GO TO Q. 4.15)**

4.14 When was the last time that you /that .... worked ?

Year				Month	

- ○ Never Worked **(Go to Q4.19)**
- ○ Not Stated

4.15 What did you/did .... do most during the past twelve months ?

- ○ Worked or had a job
- ○ Looked for first job **(Go to Section 5)**
- ○ Looked for work which was not the first **(Go to Q4.17)**
- ○ Student **(Go to Q4.17)**
- ○ Home Duties **(Go to Q4.17)**
- ○ Retired did not work **(Go to Q4.17)**
- ○ Disabled unable to work **(Go to Q4.17)**
- ○ Not Interested in work **(Go to Q4.17)**
- ○ Other **(Go to Q4.17)**
- ○ Not Stated

4.16 How many months did you/did .... work?

○ Not Stated

4.17 Have you/has .... ever been laid off permanently or made redundant during the past 5 years ?

- ○ Yes
- ○ No **(Go to Q4.19)**
- ○ Not Stated

4.18 In what Industry were you/was .... working at the time of lay-off or redundancy ?

_____

○ Not Stated

Q4.19 & Q4.20 FOR PERSONS 60 YEARS AND OVER ONLY

IF UNDER 60 YEARS ▶ **GO TO SECTION 5**

4.19 Do you/does .... currently receive any Social Welfare benefits or pension ?

- ○ Yes
- ○ No **(End Interview)**
- ○ Not Stated

4.20 What benefits or pension ? (Tick all applicable)

- ○ Employment related pension
- ○ Other
- ○ National Insurance
- ○ Not Stated
- ○ Food Stamps
- ○ Other Public Assistance/Poor Relief

Mexico 2000

PERSONA 1   PARA PERSONAS DE 12 AÑOS CUMPLIDOS O MÁS

---

22. CONDICIÓN DE ACTIVIDAD

¿La semana pasada (NOMBRE):

LEA LAS OPCIONES HASTA OBTENER UNA RESPUESTA AFIRMATIVA Y CIRCULE UN SOLO CÓDIGO

trabajó? 1   PASE A 24

tenía trabajo, pero
no trabajó? 2   PASE A 24

buscó trabajo? 3

¿Es estudiante? 4

¿Se dedica a los
quehaceres de su hogar? 5

¿Es jubilado(a) o
pensionado(a)? 6

¿Está incapacitado(a)
permanentemente
para trabajar? 7   PASE A 31

¿No trabaja? 8

---

23. VERIFICACIÓN DE ACTIVIDAD

Además de (RESPUESTA DE 22),
la semana pasada (NOMBRE):

LEA LAS OPCIONES HASTA OBTENER UNA RESPUESTA AFIRMATIVA Y CIRCULE UN SOLO CÓDIGO

ayudó en un negocio
familiar? 1

vendió algún producto? 2

hizo algún producto
para vender? 3

ayudó trabajando en el campo
o en la cría de animales? .. 4

a cambio de un pago realizó
otro tipo de actividad?
Por ejemplo: lavó o planchó ajeno,
cuidó coches 5

¿No trabaja? 6   PASE A 31

---

24. OCUPACIÓN U OFICIO

¿Qué hizo (NOMBRE) **en su trabajo de la semana pasada?**

_____

_____

ANOTE LAS ACTIVIDADES O TAREAS

¿Cuál es el nombre de su ocupación, oficio o puesto?
Por ejemplo: campesino(a), maestro(a) de primaria, vendedor(a) ambulante.

_____

_____

ANOTE LA OCUPACIÓN, OFICIO O PUESTO

---

25. SITUACIÓN EN EL TRABAJO

¿(NOMBRE) en su trabajo de la semana pasada fue:

*LEA LAS OPCIONES HASTA OBTENER UNA RESPUESTA AFIRMATIVA
Y CIRCULE UN SOLO CÓDIGO*

empleado(a) u obrero(a)? 1

jornalero(a) o peón? 2

patrón(a)? (contrata trabajadores) 3

trabajador(a) por su cuenta? 4   PASE A 27

trabajador(a) sin pago en el negocio o predio familiar? ... 5

---

26. PRESTACIONES LABORALES

¿(NOMBRE) recibe por su trabajo:

LEA TODAS LAS OPCIONES Y CIRCULE EL CÓDIGO SEGÚN LA RESPUESTA

	Sí	No
vacaciones pagadas?	1	2
aguinaldo?	3	4
reparto de utilidades?	5	6
servicio médico (IMSS, ISSSTE u otro)?	7	8
ahorro para el retiro (SAR o AFORE)?	1	2

---

27. HORAS TRABAJADAS

En total, ¿cuántas
horas trabajó
(NOMBRE) **la
semana pasada?**

|__|__|__|__|
ANOTE CON NÚMERO

---

28. INGRESOS POR TRABAJO

En total, ¿cuánto gana o recibe (NOMBRE)
por su trabajo?

*ANOTE EN PESOS LA CANTIDAD RECIBIDA Y EL PERIODO
NO RECIBE INGRESOS, ANOTE "0" EN PERIODO*

Periodo: A la semana 1
A la quincena ... 2
Al mes 3
Al año 4

$ |__|__|__|__| , |__|__|__|__| |__|__|
       ANOTE CON NÚMERO   *PERIODO*

Continúe con la pregunta 29 ➡

Mexico 2000 (*continued*)

PERSONA 1

29. ACTIVIDAD ECONÓMICA

¿En dónde trabajó (NOMBRE) **la semana pasada?**
Por ejemplo: en el campo, en una fábrica, en un taller mecánico.

_____

_____

ANOTE EN DÓNDE TRABAJÓ

El negocio, empresa o lugar donde trabajó, ¿a qué se dedica?
Por ejemplo: a cultivar maíz, a hacer muebles, a vender ropa.

_____

_____

ANOTE A QUÉ SE DEDICA

30. LUGAR DE TRABAJO

¿En qué municipio (delegación) está el negocio, empresa o lugar donde trabajó (NOMBRE) **la semana pasada?**

Aquí, en este municipio o delegación .... 1 ⇨ *Pase a 31*

En otro municipio o delegación

_____

ANOTE EL MUNICIPIO O DELEGACIÓN

¿En qué estado (o país)?

Aquí, en este estado 3

En otro estado o país

_____

ANOTE EL ESTADO O PAÍS

31. OTROS INGRESOS

¿(NOMBRE) recibe dinero por:

Periodo: A la semana 1
A la quincena 2
Al mes 3
Al año 4

LEA TODAS LAS OPCIONES Y CIRCULE EL CÓDIGO SEGÚN LA RESPUESTA.
EN LAS AFIRMATIVAS, PREGUNTE "¿CUÁNTO RECIBE?" Y ANOTE CON NÚMERO LA CANTIDAD RECIBIDA Y EL PERIODO

	Sí	No		
jubilación o pensión?	1	2	¿Cuánto recibe?	$___/___/___/ , /___/___/___/ *ANOTE CON NÚMERO* /___/ *PERIODO*
ayuda de familiares desde otro país?	3	4	¿Cuánto recibe?	$___/___/___/ , /___/___/___/ *ANOTE CON NÚMERO* /___/ *PERIODO*
ayuda de familiares dentro del país?	5	6	¿Cuánto recibe?	$___/___/___/ , /___/___/___/ *ANOTE CON NÚMERO* /___/ *PERIODO*
Procampo o Progresa?	7	8	¿Cuánto recibe?	$___/___/___/ , /___/___/___/ *ANOTE CON NÚMERO* /___/ *PERIODO*
¿Otro tipo como becas, renta, intereses bancarios?	1	2	¿Cuánto recibe?	$___/___/___/ , /___/___/___/ *ANOTE CON NÚMERO* /___/ *PERIODO*

Mexico 2000 (*English translation*)

PERSON 1	**ASK PERSONS AGED 12 OR OVER**

22. ACTIVITY STATUS

Last week (NAME):

READ THE ALTERNATIVES UNTIL A CORRECT ONE IS FOUND. SELECT ONE ONLY

Worked .... 1 ⇒ Go to 24
Had a job but was absent from work .... 2 ⇒ Go to 24
Looked for work 3
Is a student 4
Undertook household chores 5
Is retired or living from a pension 6
Is permanently disabled for work .... 7 ⇒ Go to 31
Does not work 8

23. VERIFICATION OF ACTIVITY

In addition to (REPLY TO QUESTION 22)**,** **last week** (NAME):

READ THE ALTERNATIVES UNTIL A CORRECT ONE IS FOUND. SELECT ONE ONLY

Helped in family business 1
Sold a product 2
Made a product for sale 3
Helped in crop growing or livestock breeding activities 4
Undertook another type of activity for pay (e.g. external washing or ironing, looked after cars) 5
Does not work .... 6 ⇒ Go to 31

24. OCCUPATION OR TRADE

What did (NAME) **do in his/her work last week?**

_____

_____

INDICATE ACTIVITIES OR TASKS

What is the name of his/her occupation, trade or position?
(e.g. peasant farmer, primary school teacher, travelling salesperson)

_____

_____

INDICATE OCCUPATION, TRADE OR POSITION

25. STATUS AT WORK

At work last week, (NAME) **was:**

READ THE ALTERNATIVES UNTIL THE CORRECT ONE IS FOUND. SELECT ONE ONLY

Employee or manual worker 1

Day labourer 2

Boss (hires workers) 3

Self-employed 4

Unpaid worker in family business or farm 5

} Go to 27

26. EMPLOYMENT BENEFITS

Does (NAME) **receive for work:**

READ THE ALTERNATIVES UNTIL THE CORRECT ONE IS FOUND. SELECT ONE ONLY

	Yes	No
Paid vacations?	1	2
Bonus?	3	4
Profit share?	5	6
Medical service (IMSS, ISSSTE or other)**?**	7	8
Saving for retirement (SAR or AFORE)**?**	1	2

27. HOURS WORKED

How many hours did (NAME) **work last week altogether?**

|__|__||__|__|
AMOUNT

28. LABOUR INCOME

In total, how much did (NAME) **earn or receive for his/her work?**

NOTE IN PESOS THE AMOUNT RECEIVED AND THE PERIOD. IF NO INCOME RECEIVED, PUT "0" IN THE PERIOD

Period: Per week 1
Per fortnight . 2
Per month 3
Per year 4

$ |__|__|__| |__|__|__| |__|
AMOUNT   *PERIOD*

Mexico 2000 (*English translation*) (*continued*)

29. ECONOMIC ACTIVITY

Where did (NAME) **work last week?**
(e.g. in the fields, in a factory, in a mechanical workshop)

_____

_____

INDICATE PLACE OF WORK

What does the business, firm or place where (NAME) **worked do?** (e.g. cultivates maize, makes furniture, sells clothing)

_____

_____

INDICATE ACTIVITY

30. PLACE OF WORK

In what municipality (delegation) is the business, firm or place where (NAME) **worked last week located?**

Here in this municipality or delegation .... 1 ⇒ Go to 31

In another municipality or delegation

_____

_____

INDICATE MUNICIPALITY OR DELEGATION

In what state or country?
Here in this state 3
In another state or country

_____

_____

INDICATE STATE OR COUNTRY

31. OTHER INCOME

(NAME) **receives money in respect of:**

READ ALL OPTIONS AND CIRCLE CORRECT REPLIES. FOR EACH AFFIRMATIVE REPLY, ASK **"HOW MUCH DO YOU RECEIVE"** *AND ENTER THE AMOUNT RECEIVED AND PERIOD*

Period: Per week 1
Per fortnight ... 2
Per month 3
Per year 4

	Yes	No			
Retirement or other pension	1	2	**How much received:**	$ |__||__||__||__||__| *AMOUNT*	|__| *PERIOD*
Family assistance from abroad	3	4	**How much received:**	$ |__||__||__||__||__| *AMOUNT*	|__| *PERIOD*
Family assistance within country	5	6	**How much received:**	$ |__||__||__||__||__| *AMOUNT*	|__| *PERIOD*
Procampo or Progresa	7	8	**How much received:**	$ |__||__||__||__||__| *AMOUNT*	|__| *PERIOD*
Income from other source, e.g. scholarships, rent, bank interest	1	2	**How much received:**	$ |__||__||__||__||__| *AMOUNT*	|__| *PERIOD*

Turkey 2000

ASK PERSONS AGED 12 AND OVER

35 **Did you work at any job during the last week?**

(Persons who worked to earn cash or income in kind even 1 hour during the last week should be marked as "worked". For the persons who did not work last week because of vacation, illness, travel, etc., the choice "Did not work but has a job attachment" will be marked.)

1   Worked

2   Did not work but has a job attachment

3   Did not work ——————————→ *Go to Question 40*

36 **What was your main task or duty at the job that you worked or had a job attachment in the last week?**

(Occupations such as; farmer, mechanical engineer, bank manager, typist, history teacher, greengrocer, shoe repairer, barber, carpet machine operator, master workman builder, etc. should be written.) **General names such as; tradesman, official, worker, self employed should not be written.**

37 **What is the name of the work place or organization where you worked or had a job attachment in the last week?**

(For example; field, municipality, security office, ministry of justice, courthouse, grocery, greengrocery shop, barber shop, construction, home, primary school, high school, university etc.)

38 **What is the main activity of the work place (firm/organization) where you worked or had a job attachment in the last week?**

(For example; public service, retail trade, banking, refrigerator manufacturing, construction works, field agriculture, animal feeding, poultry.)

39 **What is your employment status in the job where you worked or had a job attachment in the last week?**

1   Regular or casual employee

2   Employer

3   Self-employed

4   Unpaid family worker

→ *Go to Question 43*

40 **Are you seeking a job?**

1   Yes   2   No ——→ *Go to Question 42*

41 **When did you last take any steps to look for work?**

1   1 month ago

2   2-3 months ago

3   4-6 months ago

4   7 months or more

→ *Go to Question 43*

5   Have done nothing

42 **Why are you not seeking a job or not working?**

1   Found a job but waiting to start   4   Retired

2   Student   5   Income recipients

3   Housewife/Housegirl   6   Other

43 **What is your main occupation?**

(If you do not have mark "none")   None

Zambia 2000

FOR PERSONS 12 YEARS AND OLDER

P23 What were you mainly doing in the last 7 days?

Worked - Paid non-seasonal [1]
Worked - Unpaid non-seasonal [2]
Worked - Paid seasonal [3]
Worked - Unpaid seasonal [4]
On leave [5]
Unpaid work on household holding or business [6]
Unemployed and seeking work [7]
Not seeking work but available for work [8]
Full-time housewife/homemaker [9]
Full-time student [10]
Not available for work for other reasons [11]

P24 What have you mainly been doing for the past 12 months?

Worked - Paid non-seasonal [1]
Worked - Unpaid non-seasonal [2]
Worked - Paid seasonal [3]
Worked - Unpaid seasonal [4]
On leave [5]
Unpaid work on household holding or business [6]
Unemployed and seeking work [7])
Not seeking work but available for work [8]) **GO**
Full-time housewife/homemaker [9]) **TO**
Full-time student [10]) **P28**
Not available for work for other reasons [11])

P25 Employment Status

Since August 1999 what have you been mainly:

An employer ▭
An employee ▭
Self-employed ▭
An unpaid family worker ▭

P26 Occupation

What was your main occupation for the last 12 months?

Write name of occupation and enter code.

[0]	[0]	[0]
[1]	[1]	[1]
[2]	[2]	[2]
[3]	[3]	[3]
[4]	[4]	[4]
[5]	[5]	[5]
[6]	[6]	[6]
[7]	[7]	[7]
[8]	[8]	[8]
[9]	[9]	[9]

P27 Industry

What kind of main product or service is/was produced where you work?

Write name of industry and enter code.

[0]	[0]	[0]
[1]	[1]	[1]
[2]	[2]	[2]
[3]	[3]	[3]
[4]	[4]	[4]
[5]	[5]	[5]
[6]	[6]	[6]
[7]	[7]	[7]
[8]	[8]	[8]
[9]	[9]	[9]

D. Questionnaires from developed countries

All the examples are from the questionnaires in portrait format, three of the type 1 format and two of the type 2 format.

1. Type 1 format

Australia, 2006. There are 19 questions in the economic and related characteristics block, which starts with a question on gross income and ends with questions on activities relating to volunteer work and household services that are not economic activities as defined in this handbook, but of interest to many users. The second question is on activity status, and the next three questions relate to status in employment. There are two questions on occupation and a set of four questions relating to name, address, and industry of employer's business. There is also a question on hours worked, another on mode of travel to work and a pair on unemployment.

Canada, 2006. The questionnaire has 20 questions on household activities, labour market activities and income. The first question relates to unpaid household activities that are not economic activities, but of interest to many users. This is followed by six questions on employment and unemployment. There are two questions related to industry, two related to occupation and two on status in employment. There are also questions on the address of the place of work, mode of travel to work and languages used on the job. The last two questions in the labour market activities section cover number of weeks worked during the last calendar year (2005) and whether the work was part-time or full-time. The last question is on income in 2005.

Japan, 2005. The six questions on economic characteristics cover activity status and employment, hours worked, place of work, status in employment, industry (two sub-questions) and occupation.
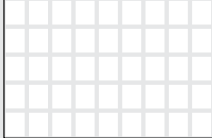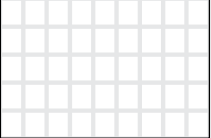
2. Type 2 format

France, 1999. There are 14 questions, with the first three covering activity status, followed by one on past work experience and occupation. The next question is on occupation, and the last two on professional categorization of the employment and principal function. There is one question on part- or full-time work and one on whether work is in family enterprise. There is a four-part question related to industry and place of work. The remaining questions deal with mode of travel to work, status in employment and size of business (with respect to number of workers), and the type of contract. The set of questions is shown in French and in English translation.[7]
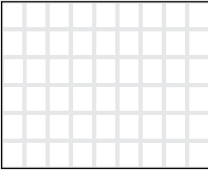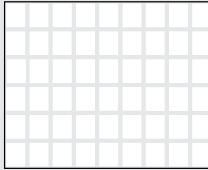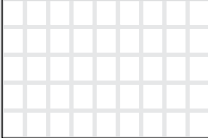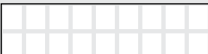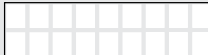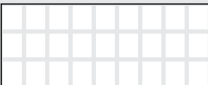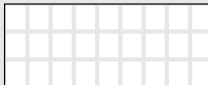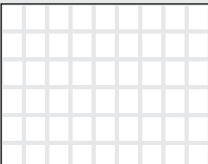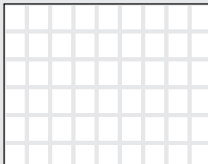
Italy, 2001 (20 per cent sample). The 24 questions are organized under the following headings: (*a*) professional and non-professional status, five questions; (*b*) work activity, 12 questions; and (*c*) place of work, seven questions. The first five questions deal with activity status, including unemployment and previous employment. In the next set of questions, the first four cover hours worked, reason for not working and full- or part-time work. The remaining eight questions cover status in employment, contractual status, occupation and industry. The last set of questions is on commuting status, address of place of work, mode of transport and travel time to the workplace.

7 Translated by the Statistics Division of the United Nations Secretariat.

Australia 2006

■ Please use CAPITAL letters only.	10	Person 1	Person 2

33 What is the *total* of all wages/salaries, government benefits, pensions, allowances and other income the person *usually* receives?

- MARK **ONE** BOX ONLY.
- **Do** *not* **deduct:** tax, superannuation contributions, health insurance, amounts salary sacrificed, or any other automatic deductions.
- Include the following:

Pensions/Allowances	Other income
family tax benefit	interest
parenting payment	dividends
unemployment benefits	rents (exclude expenses of operation)
Newstart allowance	business/farm income (exclude expenses of operation)
rent assistance	income from superannuation
pensions	any other income
student allowances	
maintenance (child support)	**Wages/salaries**
workers' compensation	regular overtime
any other pensions/allowances	commissions and bonuses

- Information from this question provides an indication of living standards in different areas.

Person 1:
- $2,000 or more per week ($104,000 or more per year)
- $1,600 - $1,999 per week ($83,200 - $103,999 per year)
- $1,300 - $1,599 per week ($67,600 - $83,199 per year)
- $1,000 - $1,299 per week ($52,000 - $67,599 per year)
- $800 - $999 per week ($41,600 - $51,999 per year)
- $600 - $799 per week ($31,200 - $41,599 per year)
- $400 - $599 per week ($20,800 - $31,199 per year)
- $250 - $399 per week ($13,000 - $20,799 per year)
- $150 - $249 per week ($7,800 - $12,999 per year)
- $1 - $149 per week ($1 - $7,799 per year)
- Nil income
- Negative income

Person 2:
- $2,000 or more per week ($104,000 or more per year)
- $1,600 - $1,999 per week ($83,200 - $103,999 per year)
- $1,300 - $1,599 per week ($67,600 - $83,199 per year)
- $1,000 - $1,299 per week ($52,000 - $67,599 per year)
- $800 - $999 per week ($41,600 - $51,999 per year)
- $600 - $799 per week ($31,200 - $41,599 per year)
- $400 - $599 per week ($20,800 - $31,199 per year)
- $250 - $399 per week ($13,000 - $20,799 per year)
- $150 - $249 per week ($7,800 - $12,999 per year)
- $1 - $149 per week ($1 - $7,799 per year)
- Nil income
- Negative income

34 *Last week,* did the person have a full-time or part-time job of any kind?

- Mark one box only.
- A 'job' means any type of work including casual, temporary, or part-time work, if it was for one hour or more.
- See page 12 of the Census Guide for more information.
- Remember to mark box like this: ▬

Person 1 / Person 2:
- Yes, worked for payment or profit
- Yes, but absent on holidays, on paid leave, on strike, or temporarily stood down
- Yes, unpaid work in a family business ▶ Go to 38
- Yes, other unpaid work ▶ Go to 46
- No, did not have a job ▶ Go to 46

35 In the main job held *last week,* was the person:

- Mark one box only.
- If the person had more than one job last week, then 'main job' refers to the job in which the person usually works the most hours.
- For all persons conducting their own business, including those with their own incorporated (e.g. Pty Ltd) company, as well as sole traders, partnerships and contractors, mark the second box.

Person 1 / Person 2:
- Working for an employer? ▶ Go to 38
- Working in own business? ▶ Go to 36

36 Was the person's business:

- Mark one box only.
- Incorporated means a limited liability company.

Person 1 / Person 2:
- Unincorporated?
- Incorporated (e.g. Pty Ltd)?

37 Does the person's business employ people?

- Mark one box only.

Person 1 / Person 2:
- No, no employees
- Yes, 1 - 19 employees
- Yes, 20 or more employees

38 In the main job held *last week,* what was the person's occupation?

- Give full title.
- For example: CHILDCARE AIDE, MATHS TEACHER, PASTRY COOK, TANNING MACHINE OPERATOR, APPRENTICE TOOLMAKER, SHEEP AND WHEAT FARMER.
- For public servants, provide official designation and occupation.
- For armed services personnel, provide rank and occupation.

Occupation (Person 1) [blank grid]

Occupation (Person 2) [blank grid]

Australia 2006 (*continued*)

		Tasks or duties	Tasks or duties
39	**What are the main tasks that the person usually performs in the occupation reported at Question 38?** • Give full details. • For example: LOOKING AFTER CHILDREN AT A DAY CARE CENTRE, TEACHING SECONDARY SCHOOL STUDENTS, MAKING CAKES AND PASTRIES, OPERATING LEATHER TANNING MACHINE, LEARNING TO MAKE AND REPAIR TOOLS AND DIES, RUNNING A SHEEP AND WHEAT FARM. • For managers, provide main activities managed.		
40	**For the main job held *last week,* what was the employer's business name?** • For self-employed persons, write name of business. • For teachers, write name of school.	Business name	Business name
41	**For the main job held *last week,* what was the person's workplace address?** • For persons who usually worked from home, write the home address. • For persons with no fixed place of work: - if the person usually travels to a depot to start work, write the depot address; - otherwise write 'NONE' in 'Suburb/Locality' box. • This information is used to calculate daytime populations and to plan transport activities.	Street number Street name Suburb/Locality State/Territory Postcode Building/Property name (if any)	Street number Street name Suburb/Locality State/Territory Postcode Building/Property name (if any)
42	**Which best describes the *industry* or *business* of the employer at the location where the person works?** • Mark one box only. • Examples of 'Other – please specify' are: REPAIRS AND MAINTENANCE, EDUCATION, AGRICULTURE, FINANCE, TELECOMMUNICATIONS SERVICE. • Remember to mark box like this: ▬	○ Manufacturing ○ Wholesaling ○ Retailing (incl. Take-aways) ○ Accommodation ○ Pubs, cafes and restaurants ○ Road freight transport ○ House construction ○ Health service ○ Community care service Other – please specify	○ Manufacturing ○ Wholesaling ○ Retailing (incl. Take-aways) ○ Accommodation ○ Pubs, cafes and restaurants ○ Road freight transport ○ House construction ○ Health service ○ Community care service Other – please specify
43	**What are the *main* goods produced or *main* services provided by the employer's *business*?** • Describe as fully as possible, using two words or more. • For example: WHEAT AND SHEEP, BUS CHARTER, HEALTH INSURANCE, PRIMARY SCHOOL EDUCATION, CIVIL ENGINEERING CONSULTANCY SERVICE, HOUSE BUILDING, STEEL PIPES.	Goods produced/services provided	Goods produced/services provided

Australia 2006 (*continued*)

44 *Last week,* how many hours did the person work in all jobs? • Add any overtime or extra time worked and subtract any time off. • Remember to mark box like this: ▬	☐☐ Hours worked ◌ None	☐☐ Hours worked ◌ None
45 How did the person get to work on Tuesday, 8 August 2006? • If the person used more than one method of travel to work, mark all methods used. • Remember to mark boxes like this: ▬	◌ Train ◌ Bus ◌ Ferry ◌ Tram (including Light Rail) ◌ Taxi ◌ Car – as driver ◌ Car – as passenger ◌ Truck ◌ Motorbike or motor scooter ◌ Bicycle ◌ Walked only ◌ Worked at home ◌ Other ◌ Did not go to work	◌ Train ◌ Bus ◌ Ferry ◌ Tram (including Light Rail) ◌ Taxi ◌ Car – as driver ◌ Car – as passenger ◌ Truck ◌ Motorbike or motor scooter ◌ Bicycle ◌ Walked only ◌ Worked at home ◌ Other ◌ Did not go to work

46 Did the person actively look for work at any time in the *last four weeks*?
• Examples of actively looking for work include: being registered with Centrelink as a job seeker; checking or registering with any other employment agency; writing, telephoning or applying in person to an employer for work; or advertising for work.
	◌ No, did not look for work
	   ▶ **Go to 48**
	◌ Yes, looked for full-time work
	◌ Yes, looked for part-time work
		◌ No, did not look for work
		   ▶ **Go to 48**
		◌ Yes, looked for full-time work
		◌ Yes, looked for part-time work

47 If the person had found a job, could the person have started work *last week*? • Remember to mark box like this: ▬	◌ Yes, could have started work last week ◌ No, already had a job to go to ◌ No, temporarily ill or injured ◌ No, other reason	◌ Yes, could have started work last week ◌ No, already had a job to go to ◌ No, temporarily ill or injured ◌ No, other reason
48 In the *last week* did the person spend time doing unpaid domestic work for their household? • Include all housework, food/drink preparation and cleanup, laundry, gardening, home maintenance and repairs, and household shopping and finance management. • See page 14 of the Census Guide for more information.	◌ No, did not do any unpaid domestic work in the last week ◌ Yes, less than 5 hours ◌ Yes, 5 to 14 hours ◌ Yes, 15 to 29 hours ◌ Yes, 30 hours or more	◌ No, did not do any unpaid domestic work in the last week ◌ Yes, less than 5 hours ◌ Yes, 5 to 14 hours ◌ Yes, 15 to 29 hours ◌ Yes, 30 hours or more
49 In the *last two weeks* did the person spend time providing unpaid care, help or assistance to family members or others because of a disability, a long term illness or problems related to old age? • Recipients of Carer Allowance or Carer Payment should state that they provided unpaid care. • Ad hoc help or assistance, such as shopping, should only be included if the person needs this sort of assistance because of his/her condition. • Do not include work done through a voluntary organisation or group.	◌ No, did not provide unpaid care, help or assistance ◌ Yes, provided unpaid care, help or assistance	◌ No, did not provide unpaid care, help or assistance ◌ Yes, provided unpaid care, help or assistance
50 In the *last two weeks* did the person spend time looking after a child, *without pay*? • Only include children who were less than 15 years of age. • Mark all applicable responses.	◌ No ◌ Yes, looked after my own child ◌ Yes, looked after a child other than my own	◌ No ◌ Yes, looked after my own child ◌ Yes, looked after a child other than my own
51 In the *last twelve months* did the person spend any time doing voluntary work through an organisation or group? • Exclude anything you do as part of your paid employment or to qualify for a Government benefit. • Exclude working in a family business.	◌ No, did not do voluntary work ◌ Yes, did voluntary work	◌ No, did not do voluntary work ◌ Yes, did voluntary work

Canada 2006

Remember, these questions are only for persons aged 15 and over.

HOUSEHOLD ACTIVITIES

Note:

Last week refers to Sunday, May 7, to Saturday, May 13, 2006.

In Question 33, report all time spent on each activity, even if two or more activities took place at the same time.
Visit www.census2006.ca or call 1 877 594-2006 for more information.

FOR INFORMATION ONLY

33 **Last week**, how many hours did this person spend doing the following activities:

(a) doing **unpaid** housework, yard work or home maintenance for members of this household, or others?

Some examples include: preparing meals, washing the car, doing laundry, cutting the grass, shopping, household planning, etc.

○ None	○ None
○ Less than 5 hours	○ Less than 5 hours
○ 5 to 14 hours	○ 5 to 14 hours
○ 15 to 29 hours	○ 15 to 29 hours
○ 30 to 59 hours	○ 30 to 59 hours
○ 60 hours or more	○ 60 hours or more

(b) looking after one or more of this person's own children, or the children of others, **without pay**?

Some examples include: bathing or playing with young children, driving children to sports activities or helping them with homework, talking with teens about their problems, etc.

○ None	○ None
○ Less than 5 hours	○ Less than 5 hours
○ 5 to 14 hours	○ 5 to 14 hours
○ 15 to 29 hours	○ 15 to 29 hours
○ 30 to 59 hours	○ 30 to 59 hours
○ 60 hours or more	○ 60 hours or more

(c) providing **unpaid** care or assistance to one or more seniors?

Some examples include: providing personal care to a senior family member, visiting seniors, talking with them on the telephone, helping them with shopping, banking or with taking medication, etc.

○ None	○ None
○ Less than 5 hours	○ Less than 5 hours
○ 5 to 9 hours	○ 5 to 9 hours
○ 10 to 19 hours	○ 10 to 19 hours
○ 20 hours or more	○ 20 hours or more

Canada 2006 (*continued*)

Remember, these questions are only for persons aged 15 and over.

LABOUR MARKET ACTIVITIES

34 **Last week**, how many hours did this person spend working for pay **or** in self-employment?

Please enter the total number of hours worked for pay or in self-employment at all jobs held last week.

Include:

- *working for wages, salary, tips or commission;*
- *working in his/her own business, farm or professional practice, alone or in partnership;*
- *working directly towards the operation of a family farm or business without formal pay arrangements (e.g., assisting in seeding, doing accounts).*

Number of hours (to the nearest hour)

☐☐☐ → **Go to Question 40**

OR

○ None → **Continue with the next question**

Number of hours (to the nearest hour)

☐☐☐ → **Go to Question 40**

OR

○ None → **Continue with the next question**

FOR INFORMATION ONLY

35 **Last week**, was this person on temporary lay-off or absent from his/her job or business?

Mark "Ⓧ" one circle only.

○ No
○ Yes, on temporary lay-off from a job to which this person expects to return
○ Yes, on vacation, ill, on strike or locked out, or absent for other reasons

○ No
○ Yes, on temporary lay-off from a job to which this person expects to return
○ Yes, on vacation, ill, on strike or locked out, or absent for other reasons

36 **Last week**, did this person have definite arrangements to start a new job within the next four weeks?

○ No
○ Yes

○ No
○ Yes

Canada 2006 (*continued*)

Remember, these questions are only for persons aged 15 and over.

37 Did this person look for paid work **during the past four weeks**?

For example, did this person contact an employment centre, check with employers, place or answer newspaper ads, etc.?

Mark "⊗" one circle only.

○ No → **Go to Question 39**

○ Yes, looked for full-time work

○ Yes, looked for part-time work (less than 30 hours per week)

○ No → **Go to Question 39**

○ Yes, looked for full-time work

○ Yes, looked for part-time work (less than 30 hours per week)

FOR INFORMATION ONLY

38 Could this person have started a job **last week** had one been available?

Mark "⊗" one circle only.

○ Yes, could have started a job

○ No, already had a job

○ No, because of temporary illness or disability

○ No, because of personal or family responsibilities

○ No, going to school

○ No, other reasons

○ Yes, could have started a job

○ No, already had a job

○ No, because of temporary illness or disability

○ No, because of personal or family responsibilities

○ No, going to school

○ No, other reasons

39 When did this person last work for pay **or** in self-employment, even for a few days?

Mark "⊗" one circle only.

○ In 2006 → **Continue with the next question**

○ In 2005 → **Continue with the next question**

○ Before 2005 → **Go to Question 51**

○ Never → **Go to Question 51**

○ In 2006 → **Continue with the next question**

○ In 2005 → **Continue with the next question**

○ Before 2005 → **Go to Question 51**

○ Never → **Go to Question 51**

Canada 2006 (*continued*)

Remember, these questions are only for persons aged 15 and over.

Note: Questions 40 to 48 refer to this person's job or business last week. If this person held no job last week, answer for the job of longest duration since January 1, 2005. If this person held more than one job last week, answer for the job at which he/she worked the **most hours**.

40 For whom did this person work?

For self-employed persons, enter the name of their business. If the business does not have a name, enter the person's name.

Print in capital letters as follows:

Name of firm, government agency, etc.

A B C   C O N C R E T

E   P R O D U C T S

L I M I T E D

Name of firm, government agency, etc.	Name of firm, government agency, etc.

FOR INFORMATION ONLY

Section, plant, department, etc. (if applicable)	Section, plant, department, etc. (if applicable)

41 What kind of business, industry or service was this?

Please give details. For example:
• new home construction
• primary school
• municipal police
• wheat farm
• shoe store
• food wholesale
• car parts factory

Kind of business, industry or service	Kind of business, industry or service

42 What was this person's work or occupation?

Please be specific. For example:
• legal secretary
• plumber
• fishing guide
• wood furniture assembler
• secondary school teacher

(If in the Armed Forces, give rank.)

Occupation	Occupation

43 In this work, what were this person's main activities?

Please give details. For example:
• prepared legal documents
• installed residential plumbing
• guided fishing parties
• made wood furniture products
• taught mathematics

Main activities	Main activities

Canada 2006 (*continued*)

Remember, these questions are only for persons aged 15 and over.

44 In this job or business, was this person mainly:

Mark "⊗" one circle only.

- ○ working for wages, salary, tips or commission?
  ↳ **Go to Question 46**
- ○ working without pay for his/her spouse or another relative in a family farm or business?
  ↳ **Go to Question 46**
- ○ self-employed without paid help (alone or in partnership)?
- ○ self-employed with paid help (alone or in partnership)?

- ○ working for wages, salary, tips or commission?
  ↳ **Go to Question 46**
- ○ working without pay for his/her spouse or another relative in a family farm or business?
  ↳ **Go to Question 46**
- ○ self-employed without paid help (alone or in partnership)?
- ○ self-employed with paid help (alone or in partnership)?

45 If self-employed, was this person's farm or business incorporated?

- ○ No
- ○ Yes

- ○ No
- ○ Yes

46 At what address did this person **usually** work **most of the time**?

Example: 365 Laurier Ave. West

Number
Name
Type
Direction

If direction (e.g., North, South, East or West) is a part of the street address, please include it.

If street address is unknown or if the address is a post office box, specify the building or nearest street intersection. Do not give a post office box number.

Please give the name of the city or town rather than the metropolitan area of which it is a part.

For example:

- Saanich rather than Victoria (metropolitan area);
- St. Albert rather than Edmonton (metropolitan area);
- Laval rather than Montréal (metropolitan area).

If the address of work is different than the address of the employer, please provide the address where this person actually works (e.g., school teachers should provide the address of their school, not the address of the school board).

- ○ Worked at home (including farms)
  ↳ **Go to Question 48**
- ○ Worked outside Canada
  ↳ **Go to Question 48**
- ○ No fixed workplace address
  ↳ **Go to Question 47**
- ○ Worked at the address specified below:
  Specify complete address

Street address (see example)

City, town, village, township, municipality or Indian reserve

Province/territory

Postal code

- ○ Worked at home (including farms)
  ↳ **Go to Question 48**
- ○ Worked outside Canada
  ↳ **Go to Question 48**
- ○ No fixed workplace address
  ↳ **Go to Question 47**
- ○ Worked at the address specified below:
  Specify complete address

Street address (see example)

City, town, village, township, municipality or Indian reserve

Province/territory

Postal code

FOR INFORMATION ONLY

Canada 2006 (*continued*)

Remember, these questions are only for persons aged 15 and over.

47 How did this person **usually** get to work?

*If this person used more than one method of transportation, **mark the one used for most of the travel distance**.*

○ Car, truck or van — as driver	○ Car, truck or van — as driver
○ Car, truck or van — as passenger	○ Car, truck or van — as passenger
○ Public transit *(e.g., bus, streetcar, subway, light-rail transit, commuter train, ferry)*	○ Public transit *(e.g., bus, streetcar, subway, light-rail transit, commuter train, ferry)*
○ Walked to work	○ Walked to work
○ Bicycle	○ Bicycle
○ Motorcycle	○ Motorcycle
○ Taxicab	○ Taxicab
○ Other method	○ Other method

48 (a) In this job, what language did this person use **most often**?

○ English	○ English
○ French	○ French
Other — *Specify*	Other — *Specify*

(b) Did this person use any other languages **on a regular basis** in this job?

○ No	○ No
○ Yes, English	○ Yes, English
○ Yes, French	○ Yes, French
Yes, Other — *Specify*	Yes, Other — *Specify*

FOR INFORMATION ONLY

Canada 2006 (*continued*)

Remember, these questions are only for persons aged 15 and over.

49 In how many **weeks** did this person work **in 2005**?

Please enter the total number of weeks worked for pay or in self-employment at all jobs held in 2005.

Include **those weeks** *in which this person:*

- was **on vacation** or **sick leave with pay**;

- worked **full time** or **part time**;

- worked for **wages**, **salary**, **tips** or **commission**;

- was **self-employed**;

- worked directly towards the operation of a **family farm or business** without formal pay arrangements.

○ None → **Go to Question 51**

OR

[ ][ ] ← Number of weeks

○ None → **Go to Question 51**

OR

[ ][ ] ← Number of weeks

FOR INFORMATION ONLY

50 During **most** of those **weeks**, did this person work **full time** or **part time**?

Mark "⊗" one circle only.

○ Full time (30 hours or more per week)

○ Part time (less than 30 hours per week)

○ Full time (30 hours or more per week)

○ Part time (less than 30 hours per week)

INCOME IN 2005

51 To save time, each person can give Statistics Canada permission to use the income information already available in his/her income tax files instead of answering **Question 52**.

- *This option is only available for persons who filed a tax return for the year ending December 31, 2005.*
- *Please note that your income tax information will be used for statistical purposes only.*

Does this person give Statistics Canada permission to use the income information already available in his/her income tax files for the year ending December 31, 2005?

○ Yes → Person 1 agrees. **Go to Question 53**

○ No → **Continue with Question 52**

○ Yes → Person 2 agrees. **Go to Question 53**

○ No → **Continue with Question 52**

Canada 2006 (*continued*)

Remember, these questions are only for persons aged 15 and over.

52 During the year ending December 31, 2005, did this person receive any income from the sources listed below?

*Answer "Yes" or "No" for all sources. If "Yes", also **enter the amount**; in case of a loss, also mark "Loss".*

PAID EMPLOYMENT:

(a) Total **wages** and **salaries**, *including commissions, bonuses, tips, taxable benefits, research grants, royalties, etc., before any deductions*

○ Yes ⤷ $ [ ][ ][ ][ ][ ] .00 ○ No

○ Yes ⤷ $ [ ][ ][ ][ ][ ] .00 ○ No

SELF-EMPLOYMENT:

(b) **Net farm income** *(gross receipts minus expenses), including grants and subsidies under farm-support programs, marketing board payments, gross insurance proceeds*

○ Yes ⤷ $ [ ][ ][ ][ ][ ] .00 ○ No ○ Loss

○ Yes ⤷ $ [ ][ ][ ][ ][ ] .00 ○ No ○ Loss

(c) **Net non-farm income** from **unincorporated business, professional practice**, etc. *(gross receipts minus expenses)*

○ Yes ⤷ $ [ ][ ][ ][ ][ ] .00 ○ No ○ Loss

○ Yes ⤷ $ [ ][ ][ ][ ][ ] .00 ○ No ○ Loss

INCOME FROM GOVERNMENT:

(d) **Child** benefits, *such as child tax benefits, family allowances (federal, provincial and territorial)*

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

(e) **Old Age Security Pension, Guaranteed Income Supplement, Allowance** and **Allowance for the Survivor** *from federal government only (provincial income supplements should be reported in (h))*

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

(f) Benefits from **Canada** or **Quebec Pension Plan**

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

(g) Benefits from **Employment Insurance** *(total benefits before tax deductions)*

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

○ Yes ⤷ $ [ ][ ][ ] .00 ○ No

Canada 2006 (*continued*)

(h) **Other income** from **government sources**, *such as provincial income supplements and grants, the GST/QST/HST credit, provincial tax credits, workers' compensation, veterans' pensions, welfare payments*	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No
OTHER INCOME: (i) **Dividends, interest** on bonds, deposits and savings certificates, and **other investment income**, *such as net rents from real estate, interest from mortgages. Do not include capital gains/losses.*	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No ○ Loss	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No ○ Loss
(j) **Retirement pensions, superannuation** and **annuities,** *including those from RRSPs and RRIFs. Do not include withdrawals from a pension plan or RRSP.*	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No
(k) **Other money income**, *such as alimony, child support, scholarships*	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No
TOTAL INCOME in 2005 from all sources	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No ○ Loss	○ Yes ➜ $ ⬚⬚⬚ ⬚⬚⬚ .00 ○ No ○ Loss
INCOME TAX PAID on 2005 income (federal, provincial and territorial)	➤ $ ⬚⬚⬚ ⬚⬚⬚ .00	➤ $ ⬚⬚⬚ ⬚⬚⬚ .00

The following question is for all persons who usually live here including those less than 15 years old. If you are answering on behalf of other people, please consult each person.

53 The *Statistics Act* guarantees the confidentiality of your census information. Only if you mark "YES" to this question will your personal information be made public, 92 years after the 2006 Census. If you mark "NO" or leave the answer blank, your personal information will never be made publicly available.

Does this person agree to make his/her 2006 Census information available for public release in 2098 (92 years after the census)?	○ Yes ○ No	○ Yes ○ No

Japan 2005

This is the English translation of the Japanese questionnaire form.

1

For pevery member of your household	**6 Did the person work at any time during the week from 24th to 30th of September?** • "Work" means any work for pay or profit. It also includes work performed in a family business (farming or retail business), as well as home handicraft and part time work. • "School" includes a non-regular school such as a Japanese language school, preparatory school, professional school.

If worked to get income. ➤

Mostly worked | Worked in addition to regular housework | Worked while attending school

Fill out Q.7 to 11

If not worked at all and did not get income. ➤

Had a job but was temporarily absent from work — Looked for a job — Did housework — Attended shool — Other (young children, elderly persons, etc.)

Fill out Q.8 to 11 | End of the questionnaire | Fill out Q.8 | End of the questionnaire

---

For persons who worked	**7 Hours actually worked during the week** • Total hours actually worked during the week from 24th to 30th of September. "Work" includes side work and home handicraft job.

☐☐☐ Hours

[Round up fractions greater than 30 min.
Round off fractions smaller than 30 min.]

For persons who worked or attended school	**8 Place of work or location of school** • For a person who worked and also attended school during the week, give the place of work. • If the place is other than one's own home, write the address of the place of work or the location of the school (i.e., the name of the *to*, *do*, *fu* or *ken* (prefecture), and the *shi* (city), *ku* (ward), *machi/cho* (town), or *mura/son* (village)). Do not omit the name of the *ku*, if the address is in one of the 15 major cities. • The major 15 cities include *ku* area of Tokyo, Sapporo, Sendai, Saitama, Chiba, Yokohama, Kawasaki, Shizuoka, Nagoya, Kyoto, Osaka, Kobe, Hiroshima, Kitakyushu and Fukuoka.

Home (Including living-in employee) | Same *ku*, *shi*, *machi/cho*, *mura/son* | Other *ku*, *shi*, *machi/cho*, *mura/son*

(Write the address)

to, *do*, *fu*, *ken*
shi, *gun*
ku, *machi/cho*, *mura/son*

For persons who worked	**9 Employment status** • "Temporary employee" means the employee on a daily basis or the one with employment period limited to less than a year. • "Self-employed" includes a proprietor of unincorporated business (including a farmer) or a person working on his/her own account.

Employee: Regular employee — Temporary employee | Director of a firm or corporation

Self-employed: Employing others — Not employing others | Family worker | Doing piece rate job at home (Side job)

For persons who worked	**10 Name of establishment and the kind of business** • Write the name (including the name of the department if the establishment is a government office) of the establishment (head office, branch, sales office, factory, shop, etc.) where the person worked during the week, and describe the kind of work done there. • Describe in detail the major business conducted at the place of work.	Name of establishment	Please describe in detail by referring to the example in the "Instructions".
		Kind of business	

---

11 Kind of work

• Describe in detail the kind of work the person did during the week.

France 1999

---

(11) QUELLE EST VOTRE SITUATION?

• VOUS TRAVAILLEZ
Cochez la case et passez au verso (questions 15 à 24)
y compris si vous êtes en **congé de maladie** *ou de* **maternité**, *si vous* **aidez** *un membre de votre famille dans son travail ou si vous êtes* **apprenti sous contrat, stagiaire rémunéré** ☐ 1

• VOUS NE TRAVAILLEZ PAS (OU PLUS)
Cochez la case et répondez aux questions 12 à 14 ☐ 2

(12) ÊTES-VOUS?

• **Étudiant** (facultés, IUT, grandes écoles, etc.) . . . . ☐ 1

• **Élève** (collège, lycée) . . . . ☐ 2

• **Stagiaire non rémunéré** . . . . ☐ 3

• **Chômeur** (inscrit ou non à l'ANPE) . . . . ☐ 4

• **Préretraité** . . . . ☐ 5

• **Retraité :** ▸ ancien **salarié** . . . . ☐ 6

          ▸ ancien **indépendant** (agriculteur, artisan, commerçant, etc.) . ☐ 7

• **Autre** (femme ou homme au foyer, personne ne percevant qu'une pension de réversion ou d'invalidité, etc.) . . . . ☐ 8

---

(13) CHERCHEZ-VOUS UN EMPLOI?

• Vous ne cherchez pas d'emploi. . . . . ☐ 1

• Vous cherchez un emploi depuis :

    moins d'un an ☐ 2   plus d'un an ☐ 3

---

(14) AVEZ-VOUS DÉJÀ TRAVAILLÉ?

    OUI ☐ 1 ➡ Quelle était votre **profession principale?**

    NON ☐ 2

Signez le bulletin au bas de la page 2

France 1999 (*continued*)

VOUS TRAVAILLEZ

(15) PROFESSION EXERCÉE ACTUELLEMENT

Soyez précis. (Ex. : électricien **d'entretien de robot**, comptable **d'assurances**, technicien **chimiste**, etc.)

Si vous êtes agent de la Fonction Publique de l'État ou des collectivités (y compris HLM, hôpitaux publics), **précisez votre grade** (corps, catégorie…)

(16) TRAVAILLEZ-VOUS ?

• À temps complet .... ☐ 1

• À temps partiel :

    plus d'un mi-temps ☐ 2   à mi-temps ou moins ☐ 3

Le temps partiel est déterminé par rapport au temps de travail normal dans votre entreprise.

(17) VOTRE ACTIVITÉ PRINCIPALE consiste-t-elle à aider un membre de votre famille dans son travail ? *(que vous perceviez ou non un salaire)*

(Exploitation agricole ou artisanale, commerce, profession libérale, etc.)

    OUI ☐ 1   NON ☐ 2

(18) OÙ TRAVAILLEZ-VOUS ?

(a) **Adresse de votre lieu de travail :** (Ex. : 18, boulevard Pasteur)

Si travail à domicile, indiquez « à domicile »
Si travail chez un particulier, indiquez « particulier »
Si lieu de travail variable, indiquez « variable »

  ▸ Est-ce dans la **commune** où vous résidez ?
  (ou dans l'arrondissement s'il s'agit de Paris, Lyon, Marseille)

    OUI ☐ 1   NON ☐ 2

Si non, indiquez la commune où vous travaillez :
(précisez l'arrondissement)

Commune :

Département : ☐☐

(pays pour l'étranger)

(b) **Nom (ou raison sociale) de l'établissement qui vous emploie ou que vous dirigez :**

(c) **Adresse de cet établissement**, si elle est différente de celle déclarée à la question 18 a.

(d) **Activité de cet établissement :**
Soyez très précis. (Ex. : commerce **en gros** de fruits et légumes, fabrication d'outillage **mécanique**, etc.)

(19) QUEL MODE DE TRANSPORT UTILISEZ-VOUS LE PLUS SOUVENT POUR ALLER TRAVAILLER ?

• **Pas de transport** (travail à domicile).......................... ☐ 1

• **Marche à pied uniquement**............................. ☐ 2

• **Un seul mode de transport**

  Deux-roues ☐ 3   Voiture ☐ 4   Transports ☐ 5
              particulière   en commun

• **Plusieurs modes de transport** .... ☐ 6

(20) ÊTES-VOUS ?

• **Indépendant ou à votre compte,**
  y compris aide familial non salarié............................ ☐ 1

• **Chef d'entreprise salarié,**
  PDG, gérant minoritaire de SARL, co-gérant .... ☐ 2

• **Salarié,** y compris aide familial salarié......................... ☐ 3

(21) SI VOUS ÊTES À VOTRE COMPTE OU CHEF D'ENTREPRISE Combien de salariés employez-vous ?

Aucun ☐ 1   1 ou 2 ☐ 2   3 à 9 ☐ 3   10 ou plus ☐ 4

Ne comptez ni les apprentis ni les gens de maison. Dans l'agriculture, comptez seulement les salariés permanents.

Les questions 22 à 24 ne s'adressent qu'aux salariés et aux stagiaires rémunérés

(22) INDIQUEZ VOTRE TYPE DE CONTRAT OU D'EMPLOI

• Apprenti sous contrat... ☐ 1

• Placé par une agence d'intérim .... ☐ 2

• CES ou autre emploi aidé (CIE, qualification, emploi-jeune, etc.) .... ☐ 3

• Stagiaire rémunéré (SIFE, etc.)............................... ☐ 4

• Contrat à durée déterminée
  (y compris contrat court, saisonnier, etc.)......................... ☐ 5

• Titulaire de la Fonction Publique (État, locale, hôpitaux)....... ☐ 6

• Contrat (ou emploi) à durée indéterminée .... ☐ 7

(23) INDIQUEZ LA CATÉGORIE PROFESSIONNELLE DE VOTRE EMPLOI

• Manœuvre, ouvrier spécialisé (OS1, OS2, OS3, etc.) .... ☐ 1

• Ouvrier qualifié ou très qualifié (P1 à P3, TA, OQ, OHQ, etc.) . ☐ 2

• Agent de service, aide soignant(e), employé de maison .... ☐ 3

• Employé de commerce, employé de bureau, personnel
  administratif de catégorie C ou D de la Fonction Publique ..... ☐ 4

• Agent de maîtrise dirigeant des ouvriers, maîtrise administrative,
  commerciale, informatique...................................... ☐ 5

• Agent de maîtrise dirigeant des techniciens ou d'autres agents
  de maîtrise .... ☐ 6

• Technicien, dessinateur, VRP.................................... ☐ 7

• Instituteur, infirmier(e), travailleur social, technicien médical,
  personnel administratif de catégorie B de la Fonction Publique. ☐ 8

• Ingénieur, cadre d'entreprise (les techniciens et agents de maîtrise ne
  devront pas se classer ici, même s'ils cotisent à une caisse de retraite
  des cadres)... ☐ 9

• Personnel de catégorie A de la Fonction Publique et assimilés . ☐ 0

(24) INDIQUEZ LA FONCTION PRINCIPALE DE VOTRE EMPLOI

• Production, fabrication, chantier, exploitation .... ☐ 1

• Installation, réglage, réparation, maintenance .... ☐ 2

• Gardiennage, nettoyage, entretien ménager .... ☐ 3

• Manutention, magasinage, transports, logistique............. ☐ 4

• Secrétariat, guichet, saisie, standard, accueil .... ☐ 5

• Gestion, comptabilité, fonction administrative, organisation... ☐ 6

• Directeur général ou adjoint direct, état-major .... ☐ 7

• Commerce, vente, technico-commercial .... ☐ 8

• Recherche, études, méthodes, informatique.................... ☐ 9

• Enseignement, formation, santé, travail social, information,
  publicité, arts, spectacles, sports............................ ☐ 0

Nous vous remercions de votre participation

France 1999 (*English translation*)

---

(11) *WHAT IS YOUR SITUATION?*

• YOU ARE WORKING
Check the box and move to the next page (questions 15 to 24) including if you are on **sick leave** or **maternity leave**, *if you are* **helping** *a member of your family in his work or if you are an* **apprentice under contract** *or a* **remunerated trainee** ☐ 1

• YOU ARE NOT WORKING OR ARE NO LONGER WORKING ☐ 2
Check the box and answer questions 12 to 14

(12) **ARE YOU? ...**

• **A student in higher education** . . . . ☐ 1

• **A student at secondary school** . . . . ☐ 2

• **An unpaid trainee** . . . . ☐ 3

• **Unemployed** (*whether registered with the ANPE or not*) . . . . ☐ 4

• **Taking early retirement** . . . . ☐ 5

• **Retired:** ▶ former **dependent worker** . . . . ☐ 6

    ▶ former **independent worker** (*farmer, craftsman, shopkeeper, etc.*) ☐ 7

• **Other** (*housewife or house-husband, person receiving only a reversion or invalidity pension, etc.*) . . . . ☐ 8

(13) **ARE YOU LOOKING FOR WORK?**

• You are not looking for work . . . . ☐ 1

• You have been looking for work for

    less than one year ☐ 2   more than one year ☐ 3

(14) **HAVE YOU WORKED IN THE PAST?**

    YES ☐ 1 ➡ What was your **principal occupation?**

    NO ☐ 2

Sign the bulletin at the bottom of page 2

France 1999 (*English translation*) (*continued*)

YOU ARE WORKING

(15) CURRENT OCCUPATION

Be precise (*e.g., "electrician **working on robot maintenance**", "accountant **in an insurance company**", not "technician" but "**chemist** technician", etc.*)

If you are employed by the State civil service or local authorities (*including social housing organisations, public hospitals*), **state your grade** (*corps, category*)

(16) ARE YOU WORKING?

- Full-time . . . . □ 1
- Part-time:
  more than half-time □ 2   half-time or less □ 3
Part-time is determined in relation to the normal working time in your firm.

(17) DOES YOUR PRINCIPAL ACTIVITY consist of helping a member of your family in his work? (*whether receiving a wage or not*)
(Farm or workshop, commerce, professional office, etc.)

YES □ 1   NO □ 2

(18) WHERE DO YOU WORK?

(a) Address of workplace (*for example, 18, boulevard Pasteur*)
If you work at home, state "at home"
If you work for a private person, indicate "private person"
If your workplace is variable, indicate "variable"

▶ Is it in the **commune** where you live?
(arrondissement in the case of Paris, Lyon, Marseille)

YES □ 1   NO □ 2

If no, indicate the commune where you work
(specify arrondissement)

Commune:

Département: □ □
(country, if abroad)

(b) Name of the establishment employing you or which you manage:

(c) Address of this establishment, if different from that reported in question 18 a.

(d) Activity of this establishment:
Be very precise (*for example, **wholesale** of fruit and vegetables, manufacturer of **mechanical** tools, etc*).

(19) WHAT FORM OF TRANSPORT DO YOU USE MOST OFTEN TO GO TO WORK?

- **None** (working at home) . . . . □ 1
- **Entirely on foot** . . . . □ 2
- **One single form of transport**
  Two-wheel □ 3   Private car □ 4   Public transportation □ 5
- **Several modes of transportation** . . . . □ 6

(20) ARE YOU ... ?

- **Independent or freelance,** including unpaid family help . . . . . . □ 1
- **Salaried head of firm,** chairman, minority manager of a limited company, co-manager . . . . □ 2
- **Dependent worker,** including wage-earning family help . . . . . □ 3

(21) IF YOU ARE FREELANCE OR HEAD OR ENTERPRISE, how many paid workers do you employ?

None □ 1   1-2 □ 2   3-9 □ 3   10 or more □ 4

Do not include apprentices or domestic staff. In agriculture, count only the permanent wage earners.

> *Questions 22 to 24 apply only to dependent workers and paid trainees*

(22) INDICATE YOUR TYPE OF CONTRACT OR EMPLOYMENT

- Apprentice under contract . . . . □ 1
- Placed by a temporary employment agency . . . . □ 2
- Government-aided employment scheme . . . . □ 3
- Paid trainee . . . . □ 4
- Fixed-term contract (including short-term, seasonal, etc.) . . . . □ 5
- Accredited civil servant (government, local, hospitals) . . □ 6
- Indefinite contract of employment . . . . □ 7

(23) INDICATE THE PROFESSIONAL CATEGORY OF YOUR EMPLOYMENT

- Unskilled worker . . . . □ 1
- Skilled or highly skilled worker . . . . □ 2
- Service agent, hospital assistant, domestic staff . . . . □ 3
- Shop worker, office worker, administrative personnel in categories C or B of the civil service . . . . . . □ 4
- Foreman supervising workers, administrative supervisor, commercial supervisor, IT supervisor . . . . □ 5
- Foreman supervising technicians or other foremen . . . . □ 6
- Technician, draughtsman, sales representative . . . . □ 7
- Teacher, nurse, social worker, medical technician, administrative personnel in category B of the civil service □ 8
- Engineer, executive (*technicians and foremen should not be classified here even if they are members of an executive retirement fund*) . . . . □ 9
- Staff in category A of the civil service and similar . . . . . . □ 0

(24) INDICATE THE PRINCIPAL FUNCTION OF YOUR EMPLOYMENT

- Production, manufacture, building site, extraction or cultivation □ 1
- Installation, adjustment, repair, maintenance . . . . □ 2
- Caretaking, cleaning, house cleaning . . . . □ 3
- Packaging, storage, transport, logistics . . . . □ 4
- Secretariat, contact with the public, data entry, switchboard, reception . . . . □ 5
- Management, accounts, administrative, organization . . . □ 6
- General manager or immediate deputy, senior staff . . . . □ 7
- Distribution, sales, technical/commercial . . . . □ 8
- Research, O&M, IT . . . . □ 9
- Education, training, health, social work, information, advertising, arts, entertainment, sports . . . . □ 0

Thank you for your participation

Italy 2001

Section II – Information on persons who usually reside in the dwelling
Individual form for persons

- Persons over 15 years of age must answer questions starting from point 6
- Persons under 15 years of age must answer questions starting from point 8

6. Professional and non-professional status

6.1 Indicate whether, during the week preceding the date of the Census (from 14 to 20 October, 2001), the person was

Employed	01	→ go to point 7
Looking for first employment	02	
Unemployed (looking for a new job)	03	
Waiting to begin a job already obtained	04	
Student	05	
Looking after home/household	06	
Retired	07	
On national military service or substitute civil service	08	} go to point 8
Disabled for work	09	
Other conditions	10	

6.2 Indicate whether, during the week from 14 to 20 October, 2001 the person worked <u>one or more hours of paid work</u>, or as contributing family worker

Yes .... 1 → go to point 7
No 2

If No

6.3 Indicate whether, during the last 4 weeks, the person has actively tried to find a job or attempted to open own activity

Yes............... 1
No 2

6.4 Indicate whether, should the opportunity arise, the person would be willing to start work within 2 weeks

Yes............... 1
No 2

6.5 Indicate whether the person has ever held a paid job or worked as a contributing family worker during their lifetime

Yes ....1 → | **6.6 Indicate the year when terminated** | → go to point 8

No ....2 → go to point 8

7. Work activity

To answer questions from 7.1 to 7.12, refer to the main work activities (activities where the greater number of hours were worked)

7.1 Indicate the number of work hours during the week of 14 to 20 October

None ....1 → | **7.2 Specify why**

Vacation .. 1	Maternity leave 3	Income support 5
Illness 2	Leave of absence 4	Lack of orders .... 6 Other ... 7

One or more ...... 2 → | **7.3 Indicate the number** hours worked

7.4 Indicate whether the person has a job

Full time 1   Part time 2

7.5 Indicate whether the person works as

Employee or other subordinate position .. 1	→ go to question 7.7
Entrepreneur............... 2	} **7.6 Indicate if there are paid workers**
Professional 3	Yes 1
Own-account worker.... 4	No .... 2 } go to question 7.9
Member of producers and/or services cooperative 5	
Contributing family worker 6	

7.7 Indicate if the person has a job for

Indefinite .... 1 → go to question 7.9

Definite .... 2 ↓

7.8 Indicate the type of work contract

Work and training contract 1
Apprentice contract 2
Interim contract 3
Other 4

Italy 2001 (*continued*)

Section II – Information on persons who usually reside in the dwelling
Individual form for persons

7.9 Indicate what your work activity consists in
[For each answer proposed, in parenthesis there are examples of professions where these activities are carried out]

Work as laborer or non-specialized service
(Farmhand, Janitor, Building laborer, Domestic servant, Mailman, Concierge, Porter, Travelling salesman) 9

Work on fixed manufacturing systems, machinery, assembly lines or drive vehicles
(Fork lift driver, Assembly of electronic equipment, Truck driver, Taxi driver) 8

Work as specialized laborer
(Mason, Mechanic, Air Conditioning technician, Shoemaker, Taylor, Carpenter) 7

Grow plants and/or raise animals
(Farmer, Fruit grower, Cattle breeder) 6

Sale to the public or provide services to people
(Sales clerk, Traffic policeman, Hairdresser, Cook, Waiter, Customs officer) 5

(continues in next column)...

... (question 7.9 continues)
Work as non-technical office worker
(Clerk, Typist, Bank teller, Switchboard operator) 4

Work in a technical, administrative, sporting or artistic field requiring an average level of qualification
(Physiotherapist, Accountant, Electronic technician, Computer technician, Athlete, Elementary school teacher, Sales representative) 3

Work in an organizational, technical, intellectual, scientific or artistic field requiring a high level of qualification
(Cardiologist, University professor, Grade school or Secondary school professor, Engineer, Chemist, Procedural analyst, Researcher) 2

Direct business or manage the work of complex organizational structures
(Businessman, public or private Manager, Head of clinical, Schoolmaster, Merchant) 1

Work as officer, non-commissioned officer, cadet or volunteer in the Armed Forces - Army, Navy, Air Force, Carabinieri
(Lieutenant General, Colonel Medical Corp, Sargent, Carabinieri cadet)... 0

7.10 Indicate the work, profession or job, in as much detail as possible, (for example, Professional Accountant, Senior High School Professor, Truck driver) **avoiding the use of general terms like clerical worker or laborer.**

7.11 Indicate the sector to which pertain the exclusive or main activities of the plant, office, organization, farm, shop, professional office, etc. where the person works or which he owns.
[Examples of economic activities are provided in parenthesis for some of the answers].

Agriculture and Fishing
Agriculture, hunting and forestry... 01

Fishing, fish-breeding and related services 02
Industry
Extraction of coal, crude petroleum, minerals, etc................... 03

Food, drink and tobacco industry 04

Textile, clothing, fur and leather industry 05
Wood, wood products (except furniture), paper, press and publishing industry.. 06
Cookery, refinery, chemical and pharmaceutical industry, rubber and plastics industry 07
Processing of non-metalliferous minerals (cement, glass, ceramics) 08
Iron and steel industry, engineering industry, electronics industry and manufacture of transportation vehicles 09
Furniture manufacture and other manufacturing industries including recuperation and preparation for recycling 10

Production and distribution of electricity, water and gas........... 11
Construction and System Installation
Building construction, public works and installation of services in buildings 12
Commerce, repairs, public practices, transportation and communications
Commerce, maintenance and repair of vehicles and motorcycles, sale of traction fuel.. 13
Wholesale trade, and sales representatives, except for vehicles and motorcycles 14

(continued in the next column)...

... (question 7.11 continues)
Retail sales, except vehicles and motorcycles, repair of personal goods and goods for the house 15

Hotels, camping areas, bars, restaurants, etc 16
Transportation (public and private), warehousing, post and telecommunications 17

Banking, insurance and other production and/or consumer services
Bank, insurance, monetary and financial brokerage........................... 18

Computer and related activities, research and development............. 19
Professional consulting activities, real estate and rentals (legal, planning, market, accounting offices, surveillance, cleaning) .. 20
Social and people services
Central and local public administration (e.g. Ministries, ANAS, regional, provincial, municipal administrations, mandatory social security) 21
Public and private education and training (including schools, colleges and military academies) 22
Health and public & private social assistance (pubic and military hospitals, doctor studios, out patients clinics, rest homes) 23
Political and union member organizations (Italian Manufacturer's Association, Chamber of Commerce) 24

Recreational, cultural and sports activities (cinema, museums) 25
Other service oriented activities (dry cleaners, beauty salons, parking, garbage collection) 26
Live-in domestic help (in this sector belong, for example, domestic helpers, baby-sitters, drivers, custodians, gardeners) 27
International organizations and organizations of other countries
Extra-territorial organizations and representations (UN, FAO, Embassies in Italy) 28

7.12 Indicate, in as much detail as possible, the sector to which pertain the exclusive or main activities of the plant, office, organization, farm, shop, professional office, etc. where the person works or which he owns.
(e.g. auto repairs, cereal farming, cattle breeding).

Italy 2001 (*continued*)

Section II – Information on persons who usually reside in the dwelling
Individual form for persons

8. Place of study or work

8.1 Indicate whether the person goes everyday out of his/her accommodation to reach the place of study or work

Yes, he/she reaches the place of study
(including kindergarten, nursery school
and professional training course............... 1 } go to question 8.2

Yes, he/she reaches the place of work... 2

No, because he/she studies at home 3

No, because he/she works at home .... 4 } the questionnaire terminates here

No, because he/she does not have a
steady work place (travelling salesman,
representatives, etc) 5

No, because he/she does not work,
study   nor   attend   professional

8.2 Indicate from which accommodation the person goes out to reach the place of study or work

From the accommodation 1   the questionnaire terminates here

From accommodation other than this....... 2 ➡ the questionnaire terminates here

If from the accommodation

8.3 Indicate whether the person re-enters from his/her usual place of study or work to the accommodation every day

Yes 1   the questionnaire terminates here

No .... 2 ➡ the questionnaire terminates here

8.4 Write out the full name and address of the usual place of study or work
[Working students must indicate the name and address of their place of work. Anyone working on a means of transportation (drivers, rail and tram workers, pilots, maritime workers, etc) must indicate the name and address of the location where he/she enters on duty (parking area, station, deposit, airport, port, etc.)]

Name of school, company, office, organization, etc: ....

Address of school, company, office, organization, etc: ....

....

In this Municipality.. 1

In another Italian Municipality .... 2 ➡ indicate which Municipality

Abroad .... 3 ➡ indicate which foreign country ✴ specify the abbreviation for the province

Persons who went to their usual place of study or work on <u>Wednesday of last week</u> have to answer questions 8.5, 8.6 and 8.7

8.5 Indicate at what time the person left the house for his/her usual place of study or work.

:
hours   minutes   (e.g. 07:30)

8.6 Indicate the time it takes (one way) to go to the usual place of study or work
[If the person accompanied children to school before going to work or place of study, consider the overall time take]

Up to 15 minutes 1

From 16 to 30 minutes 2

From 31 to 45 minutes 3

From 46 to 60 minutes 4

Over 60 minutes 5

| Reserved for use by the Municipal Census Office |
| ☐ |☐☐☐☐☐☐☐☐ |

8.7 Indicate the means of transportation used to cover the longest stretch (in terms of distance and not time) of your trip from the accommodation to the usual place of study or work

Train 01

Tram 02

Subway 03

City bus, trolley bus 04

Bus, suburban bus line 05

Company or school bus 06

Private Automobile (as driver) 07

Private Automobile (as passenger) 08

Motorcycle, moped, scooter 09

Bicycle 10

Other means (boat, cable cat, etc.) 11

By foot 12

Annex II

List of tabulations (recommended and additional) relating to economic characteristics

All the tabulations listed below are also produced at the lowest level of geographical detail and with urban and rural classification. Outlines for the basic tabulations are shown in *Principles and Recommendations for Population and Housing Censuses, Revision 2.*[1]

> Owing to the volume of some of the tabulations, it is recommended that such tabulations be available in electronic form only, since the number of printed pages would be immense. They may be disseminated through websites or compact disks or in other electronic forms.

> Printed output is appropriate for some of the tabulations, especially for tabulations relating to higher levels of geographical aggregation.

Recommended tabulations[2]

- P7.1-R: Population ...[3] years of age and over, by current (or usual) activity status, educational attainment, age and sex.*
- P7.2-R: Currently (or usually) active population, by activity status, main occupation, age and sex.*
- P7.3-R: Currently (or usually) active population, by activity status, main industry, age and sex.*
- P7.4-R: Currently (or usually) active population, by activity status, main status in employment, age and sex.*
- P7.5-R: Currently (or usually) active population, by activity status, main status in employment, main industry and sex.
- P7.6-R: Currently (or usually) active population by activity status, main status in employment, main occupation and sex.
- P7.7-R: Currently (or usually) active population by activity status, main industry, main occupation and sex.*
- P7.8-R: Population not currently (or usually) active, by functional category, age and sex.*

Additional tabulations

- P7.1-A: Currently (or usually) active population, by activity status, main status in employment, place of work, main occupation and sex.
- P7.2-A: Currently (or usually) active population, by activity status, institutional sector of employment, main industry and sex.
- P7.3-A: Currently (or usually) active population, by activity status, main occupation, educational attainment, age and sex.
- P7.4-A: Currently (or usually) active population, by activity status, main industry, educational attainment, age and sex.

---

1 United Nations publication, Sales No. E.07.XVII.8, annexes II and IV.

2 An asterisk (*) indicates a basic/essential tabulation.

3 The minimum age adopted by the country for enumerating the economically active population.

- P7.5-A: Usually active population, by activity status, sex, main status in employment and number of weeks worked in all occupations during the last year.
- P7.6-A: Currently employed population, by main status in employment, sex and number of hours worked in all occupations during the last week.
- P7.7-A: Currently (or usually) active population, by activity status, main occupation, marital status, age and sex.
- P7.8-A: Currently (or usually) active population, by activity status, main status in employment, marital status, age and sex.
- P7.9-A: Currently (or usually) active population in the informal sector, by activity status, main status in employment, place of work, main occupation and sex.
- P7.10-A: Usually active population, by monthly or annual income, main occupation and sex.
- P7.11-A: Households and population in households, by annual income and size of household.
- P7.12-A: Population not currently active (in other words, not in the labour force), by primary reason for inactivity, age and sex.
- P7.13-A: Heads or other reference members of households ...[3] years of age and over, by economic activity status, age and sex.
- P7.14-A: Households and population in households, by size of household and number of currently (or usually) employed members.
- P7.15-A: Households, by size, number of currently (or usually) unemployed members and dependent children under 15 years of age in household.
- P7.16-A: Currently (or usually) active heads or other reference members of households ...[3] years of age and over, by activity status, main status in employment, main industry and sex.

Other tabulations of interest

- P2.4-R (Recommended): Economically active foreign-born population ...[3] years of age and over, by year or period of arrival, main occupation and sex.
- P8.3-R (Recommended): Population ...[3] years of age and over, by disability status, current (or usual) activity status, age and sex.
- P2.2-A (Additional): Foreign-born population ...[3] years of age and over, by current (or usual) activity status, age and sex.

References

Australian Bureau of Statistics (1991). *Australian Standard Classification of Occupations (ASCO)–Expert Coding System: Occupation Level.* Floppy disk, version 5.0 for DOS. ABS Catalogue No. 1226.0. Canberra.

_____ (1992). *Australian Standard Classification of Occupations—ASCO Coding System: Unit Group Level.* ABS Catalogue No. 1225.0. Canberra.

_____ (1993). *Australian Standard Classification of Occupations (ASCO)—Manual Coding System: Occupation Level.* ABS Catalogue No. 1227.0. Canberra.

_____ (1994). *1996 Census of Population and Housing: Nature and Content of the Census.* ABS Catalogue No. 2008.0. Canberra.

_____ (1996). *1996 Census of Population and Housing: How Australia Takes a Census.* ABS Catalogue No. 2903.0. Canberra.

_____ (1997). *Australian Standard Classification of Occupations (ASCO), Second Edition.* ABS Catalogue No. 1220.0; and ASCO coder for Windows on CD-ROM. ABS Catalogue No. 1220.0.30.001. Canberra.

Bain, D. (1996). *A Guide to Estimating the Value of Non-Market Production in Pacific Island Developing Countries.* Nouméa: South Pacific Commission.

Campanelli, P., and others (1997). The quality of occupational coding in the United Kingdom. In *Survey Measurement and Process Quality*, L. Lyberg and others, eds. Wiley Series in Probability and Statistics. New York: John Wiley & Sons, Inc.

Carew, J., R. Woods and B. Brady (1999). 1996 census: labour force status. Census Working Paper 99/2. Canberra: Australian Bureau of Statistics.

Cho, Lee-Jay, and Robert L. Hearn, eds. (1984). *Censuses of Asia and the Pacific: 1980 Round.* Honolulu: East-West Center, Population Institute.

Commission of the European Communities, and others (1993). *System of National Accounts, 1993.* Sales No. E.94.XVII.4.

Deming, W. E. (1964). *Statistical Adjustment of Data*, chap. VII. New York: Dover Publications.

Deming, W. E., and F. F. Stephan (1940). On a least square adjustment of a sample frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics*, vol. 11.

Elbers, C., J. O. Lanjouw and P. Lanjouw (2002). Micro-level estimation of welfare. Policy Research Working Paper No. 2911. Washington, D.C.: World Bank.

Elias, P. (1996). Automatic coding of occupational information for the 2001 census of population: a feasibility study. Coventry, United Kingdom: University of Warwick, Institute for Employment Research.

Elias, P., K. Halstead and K. Prandy (1993). CASOC: *Computer-Assisted Occupational Coding.* London: Her Majesty's Stationery Office.

Embury, B. L. (1991). The ASCO expert coding system. Mimeo. Canberra: Australian Bureau of Statistics.

_____ (1997). Constructing a map of the world of work: how to develop the structure and contents of a national standard classification of occupations. STAT Working Papers, No. 95-2. Geneva: International Labour Office.

Ghosh, M., and J. N. K. Rao (1994). Small area estimation: an appraisal. *Statistical Science*, vol. 9, No. l.

Gilbert, R. J. (1982). *1980 National Census Project*. Port Moresby: Government Printer.

_____ (1986). The first complete enumeration of Papua New Guinea: the 1980 population census. *Journal of Official Statistics*, vol. 2, No. 4.

_____ (2001). Asking questions on economic characteristics in a population census. STAT Working Papers, No. 2001-1. Geneva: International Labour Office.

Hoffmann, E. (1994). Mapping the world of work: an international review of the use and gathering of occupational information. In *Labour Statistics for a Market Economy: Challenges and Solutions in the Transition Countries of Central and Eastern Europe and the Former Soviet Union*, I. Chernyshev, ed. Budapest, London and New York: Central European University Press.

_____ (2000). Developing labour account estimates: issues and approaches. In *Household Accounting: Experience in Concepts and Compilation,* vol. 2: *Household Satellite Extensions*. Studies in Methods, Series F, No. 75, vol. 2. Handbook of National Accounting. Sales No. E.00.XVII.16.

_____ (2001). Coding occupation and industry in a population census. STAT Working Papers, No. 2001-2. Geneva: International Labour Office.

Hoffmann, E., and others (1995). What kind of work do you do? Data collection and processing strategies when measuring "occupation" for statistical surveys and administrative records. STAT Working Papers, No. 95-1. Geneva: International Labour Office.

Hussmanns, R. F. (2007). Measurement of employment, unemployment and underemployment—current international standards and issues in their application. *Bulletin of Labour Statistics*, No. 2007-1. Available from http://www.ilo.org/public/english/bureau/stat/papers/listart.htm.

Hussmanns, R., F. Mehran and V. Verma (1990). *Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods*. Geneva: International Labour Office.

International Labour Office (1955). *Eighth International Conference of Labour Statisticians, 1954: Report of the Conference*. Geneva.

_____ (1969). *International Standard Classification of Occupations, Revised Edition, 1968*. Geneva.

_____ (1983). Resolution concerning statistics of the economically active population, employment, unemployment and underemployment. *Thirteenth International Conference of Labour Statisticians, October 1982: Report of the Conference*, appendix I, resolution 1. Geneva.

_____ (1988). *Fourteenth International Conference of Labour Statisticians, October-November 1987: Report of the Conference*. Geneva.

_____ (1990a). *International Standard Classification of Occupations (ISCO-88)*. Geneva.

_____ (1990b). *Statistical Sources and Methods,* vol. 5: *Total and Economically Active Population, Employment and Unemployment (Population Censuses)*. Geneva.

_____ (1993a). Resolution concerning the International Classification of Status in Employment (ICSE). *Fifteenth International Conference of Labour Statisticians, Geneva, 19-28 January 1993: Report of the Conference*, appendix I, resolution 3. Geneva.

_____ (1993b). Resolution concerning statistics of employment in the informal sector. *Fifteenth International Conference of Labour Statisticians, Geneva, 19-28 January 1993: Report of the Conference*, appendix I, resolution 2. Geneva.

_____ (1996a). *Sources and Methods: Labour Statistics,* vol. 5: *Total and Economically Active Population, Employment and Unemployment (Population Censuses).* Geneva.

_____ (1996b). Recommendations of the joint ILO/Czech Statistical Office meeting on the statistical treatment of persons on extended types of leave in respect to the international definitions of employment and unemployment (Prague, 15-17 November 1995). *Bulletin of Labour Statistics*, No. 1996-1.

_____ (2000). *Current International Recommendations on Labour Statistics, Second Edition*. Geneva.

_____ (2002). On measuring place of work. Working paper. Geneva.

_____ (2005). *Sources and Methods: Labour Statistics*, vol. 3, *Third Edition: Economically Active Population, Employment, Unemployment and Hours of Work*. Geneva.

Jabine, T., and B. Tepping (1973). Controlling the quality of industry and occupation data. *Bulletin of the International Statistical Institute, Forty-fifth Session*, vol. 3.

Kish, Leslie (1979). Rotating samples instead of censuses. *Asian and Pacific Census Forum*, vol. 6, No. 1.

_____ (1980). Design and estimation for domains. *The Statistician*, vol. 29, No. 4.

_____ (1987). Discussion. In R. Platek and others, eds. *Small Area Statistics: An International Symposium*. New York: John Wiley & Sons, Inc.

_____ (1995). *Survey Sampling*. New York: John Wiley & Sons, Inc.

Kish, L., R. M. Groves and K. Krotki (1976). Sampling errors for fertility surveys. World Fertility Survey Occasional Papers, No. 17. Voorburg, Netherlands: International Statistical Institute.

Kish, L., and V. Verma (1986). Complete censuses and samples. *Journal of Official Statistics*, vol. 2, No. 4.

Lyberg, L. (1982). Coding of occupation and industry: some experiences from Statistics Sweden. *Bulletin of Labour Statistics*, No. 1982-3.

Lyberg, L., and D. Kasprzyk (1997). Some aspects of post-survey processing. In *Survey Measurement and Process Quality*, L. Lyberg and others, eds. Wiley Series in Probability and Statistics. New York: John Wiley & Sons, Inc.

Lyberg, L., and others, eds. (1997). *Survey Measurement and Process Quality*. Wiley Series in Probability and Statistics. New York: John Wiley & Sons, Inc.

Murthy, M. N. (1980). Use of sampling in population censuses with reference to the experience in some countries of Asia and the Pacific. In *Report of the Training Seminar on the Use of Sampling in Population Censuses*. New Delhi: Office of the Registrar General of India.

Murthy, M. N., and A. S. Roy (1970). A problem of integration of surveys: a case study. *Journal of the American Statistical Association*, vol. 65, No. 329.

_____ (1975). Development of the sample design of the Indian national sample survey during its first twenty-five rounds. *Sankhya: The Indian Journal of Statistics*, vol. 37, series C, part one.

Pember, R. J. (1998). Compilation and presentation of labour statistics based on administrative records. *Bulletin of Labour Statistics*, No. 1998-1. Available from http://www.ilo.org/public/english/bureau/stat/papers/listart.htm.

Platek, R., and others, eds. (1987). *Small Area Statistics: An International Symposium*. New York: John Wiley & Sons, Inc.

Purcell, N. J., and L. Kish (1980). Postcensal estimates for local areas (or domains). *International Statistics Review*, vol. 48, No. 1.

Rao, J. N. K. (1999). Some recent advances in model-based small area estimation. *Survey Methodology*, vol. 25, No. 2.

_____ (2003). *Small Area Estimation*. Wiley Series in Survey Methodology. New York: Wiley and Sons, Inc.

United Nations (1967). *Principles and Recommendations for the 1970 Population Censuses*. Sales No. E.67.XVII.3.

_____ (1968). *International Standard Industrial Classification of All Economic Activities*, Statistical Papers, Series M, No. 4, Rev.2. Sales No. E.68.XVII.8.

_____ (1971). *Handbook of Population and Housing Census Methods—Part VI: Sampling in Connection with Population and Housing Censuses*. Studies in Methods, Series F, No. 16. ST/STAT/SER.F/16 (Part VI).

_____ (1977). *Provisional Guidelines on Statistics of the Distribution of Income, Consumption and Accumulation of Households*. Statistical Papers, Series M, No. 61. Sales No. E.77.XVII.11.

_____ (1978). *Recommendations for the 1980 Censuses of Population and Housing in the ECE Region: Regional Variant of Parts 2 and 3 of the World Recommendations for the 1980 Censuses of Population and Housing*. Statistical Standards and Studies/Conference of European Statisticians, No. 31. Sales No. E.78.II.E.6.

_____ (1980). *Principles and Recommendations for Population and Housing Censuses*. Statistical Papers, Series M, No. 67. Sales No. E.80.XVII.8.

_____ (1984). *Handbook of Household Surveys*. Studies in Methods, Series F, No. 31. Sales No. E.83.XVII.13.

_____ (1988). *Improving Statistics and Indicators on Women Using Household Surveys*. Studies in Methods, Series F, No. 48. Sales No. E.88.XVII.11.

_____ (1990a). *International Standard Industrial Classification of All Economic Activities, Third Revision*. Statistical Papers, Series M, No. 4, Rev.3. Sales No. E.90.XVII.11.

_____ (1990b). *Supplementary Principles and Recommendations for Population and Housing Censuses*. Statistical Papers, Series M, No. 67, Add.1. Sales No. E.90.XVII.9.

_____ (1992). *Handbook of Population and Housing Censuses—Part I: Planning, Organization and Administration of Population and Housing Censuses*. Studies in Methods, Series F, No. 54 (Part I). Sales No. E.92.XVII.8.

_____ (1993). *Methods of Measuring Women's Economic Activity*. Technical report. Studies in Methods, Series F, No. 59. Sales No. E.93.XVII.6.

_____ (1996). *Handbook of Population and Housing Censuses—Part IV: Economic Activity Status*. Studies in Methods, Series F, No. 54 (Part IV). Sales No. E.96.XVII.13.

_____ (1998). *Principles and Recommendations for Population and Housing Censuses, Revision 1*. Statistical Papers, Series M, No. 67, Rev.1. Sales No. E.98.XVII.8.

_____ (2000). *Household Accounting: Experience in Concepts and Compilation*, vol. 2: *Household Satellite Extensions*. Studies in Methods, Series F, No. 75, vol. 2: Handbook of National Accounting. Sales No. E.00.XVII.16.

_____ (2001). *Handbook on Census Management for Population and Housing Censuses*. Studies in Methods, Series F, No. 83. Sales No. E.00.XVII.15 Rev.1.

_____ (2002). *International Standard Industrial Classification of All Economic Activities (ISIC), Revision 3.1*. Statistical Papers, Series M, No. 4, Rev.3.1. Sales No. E.03.XVII.4. Available from http://unstats.un.org/unsd/publication/SeriesM/seriesm_4rev3_1e.pdf.

_____ (2003). *Handbook on Non-Profit Institutions in the System of National Accounts*. Studies in Methods, Series F, No. 91. Sales No. 03.XVII.9.

_____ (2005). *Guide to Producing Statistics on Time Use: Measuring Paid and Unpaid Work*. Studies in Methods, Series F, No. 93. Sales No. E.04.XVII.7.

_____ (2007). Principles and guidelines for managing statistical confidentiality and microdata access. Background document. Statistical Commission, thirty-eighth session, 27 February–2 March 2007. Available from http://unstats.un.org/unsd/statcom/doc07/BG-Microdata-E.pdf.

_____ (2008a). *International Standard Industrial Classification of All Economic Activities (ISIC), Revision 4*. Statistical Papers, Series M, No. 4, Rev.4. Sales No. E.08.XVII.25.

_____ (2008b). *Principles and Recommendations for Population and Housing Censuses, Revision 2*. Statistical Papers, Series M, No. 67, Rev.2. Sales No. E.07.XVII.8.

United Nations, Economic Commission for Europe; and Statistical Office of the European Communities (1998). *Recommendations for the 2000 Censuses of Population and Housing in the ECE Region*. Statistical Standards and Studies, No. 49. Sales No. E.98.II.E.5.

United Nations Educational, Scientific and Cultural Organization (1976). *International Standard Classification of Education (ISCED)*. Paris.

_____ (1997). *International Standard Classification of Education (ISCED)*. Paris.

Verma, V. (1988). Improving the design and structure of major BPS surveys. Mimeo. Development Studies Project, Statistical Papers, No. 4. Jakarta: BPS—Statistics Indonesia.

_____ (1989). The use of sampling in conjunction with population censuses. *Asian and Pacific Population Forum,* vol. 3, Nos. 1-2.

_____ (2001). Sample designs for national surveys: surveying small-scale economic units. *Statistics in Transition*, vol. 5, No. 3.

Index

Reference numbers refer to paragraphs, not to pages. Subparagraph references take a suffix, "a", "b", etc.; and notes are designated by "n.". For figures, tables, annexes and sample questionnaires, page numbers rather than paragraph references are given.

A

Absence from work, temporary
    of employed persons, 178-180, 288, 925
    proper recording of, 180
Absent members of household, defined by
        a time limit, 64
Accounts, set of, 399
Accuracy
    complete, an unattainable goal, 489
    and quality of a census, 489
Activities, economic and non-economic
    in developing countries, 126
    identification of, by interviewers, 127
Administrative reporting systems and
        registers, 35
    statistics based on, 727-728, 944
        limitations of, 728
Administrative units, as primary sampling
        units, 833
Age-based bias, in census-taking, 83
Age cut-offs, for excluding segments of
            the population from questions on
            economic activity, 133, 148-154
Agricultural activities
    census questions about, 334
    commercial distinguished from
            subsistence, 436b
    employment in, 181d, 181f, 335
    excluded from the informal sector, 436b
    own-account, 181f, 335, 375a
Agricultural censuses, 712e
    delimiting of survey population, 775
    sampling frames for, 855
Agricultural establishments, small,
            surveys of, 786

Agricultural season, as reference period,
            140
Agricultural workers
    market-oriented, 334
    self-employed, 318b
    subsistence farmers, 334
Area frames, 780-790
    census data as source of, 823-835, 856
    common problems with, 795-801
    exclusions from, 774
    updating of, 808
Area units, errors and changes in
            boundaries, 797c
Argentina, sample questionnaire from,
            page 282
Australia, 108, 109, 364, 657
    sample questionnaire from, page 296
*Australian Standard Classification of
            Occupations* (ASCO), 657, 675
Automated data capture, repair and
            coding systems, 555, 557
    possible data quality problems from,
            490
Available for work, 215

B

Bahamas, sample questionnaire from,
            page 282
Batch processing against flat files, 559
Bayes prediction procedures, 947
Belize, sample questionnaire from,
            page 282
Border workers, 344c, 349
Botswana, sample questionnaire from,
            page 270
Brand names, in job descriptions, to be
            removed, 642n.
Bulletin of Labour Statistics, 122
Business partners, 315
Business registers. *See* Establishments,
            registers of