United Nations
Statistics Division

ESA/STAT/AC.88/18
7 April 2003

English only

Expert Group Meeting on
Setting the Scope of Social Statistics
United Nations Statistics Division
in collaboration with the Siena Group on Social Statistics
New York, 6-9 May 2003

Integration and Coordination of Social Statistics: Perspectives on Strategies and Issues*

by

Eivind Hoffmann **

* This document is being issued without formal editing.

** Senior Statistician, ILO. The views expressed in this report are those of the author and do not imply the expression of any opinion on the part of the United Nations Secretariat.

Contents

Integration and coordination of social statistics: Perspectives on strategies and issues[1]

Introduction

1. Based on the respective classifications for statistical programmes of the (former) *ACC Sub-Committee on Statistical Activities* and the *Integrated Presentation of UN/ECE, OECD and Eurostat (IP)*[2] there would seem to be broad agreement in the international statistical community about the scope of 'social' statistics, and that 'labour statistics' would be included. However, the scope of 'labour statistics as defined in *ILO Convention No. 160 Concerning Labour Statistics*, adopted by the *International Labour Conference (ILC)* in 1985 includes the Consumer Price Index (CPI) which IP considers to be 'economic statistics'. Convention 160 also covers statistics on households' incomes and expenditures and on occupational injuries. These areas are seen as part of 'social' statistics by the IP and ACC classifications, but there they belong to categories for 'distribution of income and consumption' and 'health statistics' respectively.

2. Among the most important and basic tools for any area of statistics are the classifications that are used when describing (measuring) the different topics of interest. The IP classification has 'classifications' as a separate category while the ACC list does not mention them, and one must assume that they are included under the most appropriate subject matter area. ILO's Convention 160 leaves the issue of classifications to be mentioned only briefly in the accompanying *Recommendation No. 170 concerning Labour Statistics*, which was also adopted by ILC in 1985.[3]

3. Thus differences in how international agencies in practice see the scope of 'social statistics' do emerge as soon as one examines the details below similar labels, and this may highlight one aspect of the need for 'more systematic and integrated development of social statistics', the main focus of this meeting of experts.

(Insufficient) mechanisms for coordinating and integrating official statistics

4. The advantages of having an 'integrated and coordinated approach' to the production of official statistics have been recognized for a long time. One result has been that most national statistical systems are based on a model where there is a strong central statistical office (CSO) responsible for producing statistics for a wide range of areas and users, and for maintaining the

---

1. Note prepared for the 6-9 May 2003 United Nations Expert Group Meeting *Setting the scope for social statistics* organized by the United Nations Statistics Division. The note was prepared by Eivind Hoffmann. A first draft was significantly improved by comments from Ralf Hussmanns and A. Sylvester Young, but the views expressed are those of the author and do not necessarily reflect those of the organizers, the ILO or its Bureau of Statistics. The author apologizes for all errors and omissions, and would welcome comments and suggestions for improvements and correction. Address: CH-1211 GENEVE 22, Switzerland; e-mail: hoffmann@ilo.org

2 See http://unstats.un.org/unsd/methods/statact/acc-class.htm and http://www.unece.org/stats/archive/docs.ip.e.htm respectively.

3 When ratified by the relevant national authorities an ILO Convention becomes part of a country's international legal obligations. An ILO Recommendation is not subject to ratifications and will normally be more detailed and technical than a corresponding Convention.

required operational infrastructure, i.e. business registers and the capacity for planning and implementing establishment and household surveys. Frequently specialized units in line ministries and operational agencies will supplement the CSO, with the task to produce statistics on the basis of administrative records that are created for operational reasons by the same agencies, and also to analyze the resulting statistics together with statistics produced by the CSO, as a basis for formulating, implementing and evaluating policies for which the agencies are responsible. The precise division of labour and areas of responsibility between the CSO and other agencies will differ between countries, as will the precise organization of work within the CSO. The latter will, however, often mirror to some extent the former as the policy formulating and executing agencies are important users of the statistics produced by the CSO: thus we get units for statistics on health, education, employment and unemployment, wages and income, social security, population and families, as well as on prices, wholesale and retail trade, foreign trade, transport, manufacturing, construction and public finance, to mention just some of the possible 'areas of statistics' that may be identified in the organizational chart of a CSO.

5. Those responsible for official statistics have also recognized for a long time that the 'real world' these statistics should aim to describe cannot be divided neatly along lines that correspond to those which are convenient for organizing the production of official statistics: because in the final count "everything depends on everything else". Inspired by the power and success of the *System of National Accounts (SNA)* (see e.g. ISWGNA, 1993) as a coordinating tool for different areas of 'economic' statistics there therefore have been several efforts to create similar tools for both 'environmental' statistics (see e.g. UN, 1984) and 'social and demographic' statistics (see e.g. UN, 1974). The latter work had as one of its objectives to serve a similar role for integrating and coordinating social and demographic statistics as SNA did for economic statistics. That aim was not achieved, however, (see e.g. UN, 1979) and it may be useful to reflect on possible reasons for that before discussing alternative approaches.

6. The success of SNA as a coordinating tool seems to depend on a combination of three factors:

(i) The first, and most obvious, factor is related to the fact that one of the characteristics of 'transactions', the main units of observation in economic statistics, is the size; and that the size of transactions, as measured by a continuous variable 'value', can be added/aggregated for transactions that are defined as 'similar' according to other variables observed (e.g. the parties involved or the kind of goods and services involved, their purpose, how they are financed). Through such aggregations one arrives at measuring concepts that are meaningful for a wide range of descriptive and analytical purposes. With a few exceptions that I'll return to later one can safely say that no corresponding measure exists for any of the areas that are normally regarded as part of 'social and demographic statistics'.

(ii) The second factor is that there exist reasonably clear descriptive and analytical purposes and models that have made it fairly easy to agree on criteria that would define important and interesting similarities and differences between transactions, and the aggregates that as a consequence could be measured were more interesting as units of analysis and descriptions than the individual transactions. Such agreement on 'interesting

aggregates' does not seem to exist in and among the areas included in 'social and demographic statistics', again with an exception to which I shall return.

(iii) The third factor is that the development of SNA happened through the combined efforts of users of the results and producers of the underlying relevant statistics, with the users being the main formulators of the system's specifications. This meant that there was a 'user guarantee' in the system from the start. The 'user' involvement in the 1970s efforts to create a system for 'social and demographic statistics' that could have similar coordinating effects was much more marginal. Those efforts were almost entirely undertaken by producers of statistics, probably because it was not possible to find any descriptive and analytical purposes or models in the relevant disciplines that as clearly would benefit from the expected outcome.

7. There are two possible exceptions to the above statement that there is no important unit of observation in social and demographic statistics for which one can observe a variable that can be aggregated in a meaningful way for 'similar' units. The first is related to the fact that (almost) all human 'activities' and many 'events' take time: for measuring 'duration' there is an obvious and easily defined continuous variable. However, it is not clear that for 'activities' and 'events' there exist enough analytical and descriptive purposes for which the usefulness of measuring 'duration' is as obvious and clearly defined as those we have for the value of transactions. But this may change, given that 'time' is a limited resource for us all, whether we are looking at a short reference period or at a lifetime.

8. A second exception is that there are clearly defined analytical purposes for which the human person is the relevant unit of measurement and aggregation. The most obvious of these are those concerning the study of the dynamics of human populations, i.e. demography. For such studies one has developed a clearly defined set of concepts, definitions and methods that can be used to coordinate and integrate the relevant areas of statistics on birth, deaths, migration and the age distribution of populations, based on the number of persons with certain characteristics, e.g. age and sex for all individuals, and the number of births and marital status for women. In demography and population statistics one also finds the exception mentioned under the second factor above, and it can be seen as an area where the third factor also have been present. This suggests that the 'secret' to the success of both SNA and the tools for demographic analysis is that both have been designed for and serve clearly defined and reasonably agreed areas of macro-analysis.

9. Possible conclusions from this brief review are (i) it cannot be expected that organizational arrangements that may be convenient for the production of official statistics will provide a clear and stable basis for arriving at coordinated and integrated 'social' statistics; (ii) the 'secret' to the success of both SNA and tools for demographic analysis is that both have clearly defined and reasonably agreed areas of macro-analysis for which their respective concepts, classifications and tools have been designed; and (iii) it is not possible to find macro-issues and variables that are common to areas conventionally regarded as included in 'social' statistics, and that can serve as basis for a coherent system of definitions and classifications to achieve widespread coherence and coordination between these areas.

Back to basics

10. Having rejected both organizational arrangements and conceptual frameworks as promising tools for achieving integration and coordination of 'social' statistics, it seems reasonable to start from the observation that most of the relevant analytical and descriptive reasons for using 'social' statistics represent micro- perspectives. This means that '**social' statistics can be defined as statistics that describe the behaviour and situations of natural persons as individuals, changes to such situations and behaviour, as well as events and institutions that directly influence both the situations and the behaviour**[4].

11. With this 'definition' of the scope of 'social' statistics, the relevant issue becomes how to observe the relevant situations, behaviours, events and institutions to ensure that the statistics we can produce on the basis of these observations are (i) valid and relevant for the issues that we want to use the statistics to give us information about; and (ii) coordinated and consistent to enhance the possibility that statistics produced for different purposes and by different procedures can be used together. With this formulation it becomes evident that we need to examine the basic elements and procedures that are involved when collecting and processing data for the production of 'social' statistics, namely the reference periods and periodicity, populations covered, the units of observation and the relationships between them, the variables associated with these units and their value sets (classifications), the measurement of quantities and the possibilities for 'accounting relationships'. The limits imposed by different sources with respect to these elements must be understood and be related to the different types of use that one would like for the statistics that can be produced. The main sources are of course: (i) Direct use of administrative records; (ii) indirect use of administrative records through request for information addressed to institutions and establishments, as part of administrative report procedures or statistical censuses and surveys; and (ii) replies to questions addressed to individuals in surveys and censuses of persons and households. The main areas of use are (a) description and analysis of relevant structures and behaviour in the respective areas; (b) monitoring actual developments; and (c) formulating, implementing and evaluating policies.[5]

12. In each subject matter area of 'social' statistics, e.g. labour, crime, health, time use, household consumption, there will be one or more types of primary units that are the main objects of interest, and for each of these there will be a set of descriptive (or distributive) variables that serve to describe them. Through defined relationships between these primary units it will also be possible to describe units of a particular type in terms of characteristics that serve to describe other units. For the area of labour statistics this is illustrated in Chart 1 and Table 1 (adapted from Hoffmann, 2000) where the former indicates (with double-headed arrows) units that are linked and the latter provides a list of relevant distributive variables for each of the primary units. Thus it is possible to describe employed and unemployed persons by the variable 'industry' by defining the relevant jobs and how a job is related to an establishment (employer), where it is the latter unit that is the primary unit for the variable 'industry'. For an employed

4 The situation and behaviour of families and households can either be added to the reference to 'natural persons', or be regarded as included in the reference to 'institutions'

5 There are many possible references for elaborations of these points. For this author it is easiest to refer to Hoffmann & Lawrence (1996), Hoffmann (1995), Hoffmann (1996) and Hoffmann (2000).

person it usually is the current (main) job that provides the basis for identifying the correct 'industry' group; for unemployed persons it will normally be a past job, e.g. the most recent.

Table 1. Examples of descriptive variables in labour statistics by primary units

1. Employers (establishments)
   - Ownership/legal organization*
   - Industry*
   - Size*
   - Location

2. Posts
   - Status in employment*
   - Occupation*
   - Contractual working hours
   - Shift system
   - Pay system
   - Collective agreement

3. Jobs
   - Income from employment*
   - Amount of compensation of employees*
   - Amount' of net operating surplus for self-employed persons (net mixed income of unincorporated enterprises)*
   - Normal or usual hours worked*
   - Hours paid for
   - Working time arrangements

4. Persons
   - Sex
   - Age*
   - Nationality
   - Ethnic group
   - Union membership
   - Education attained*
   - Actual hours worked*
   - Past occupation (and other life history variables)
   - Work desires
   - Injuries*
   - Activities (other than those defining labour force status)

5. Households
   - Domicile (location)
   - Type of household*
   - household income

* Indicates that there exists some type of international recommendation concerning the definition and classification of this variable. They are not necessarily coordinated with each other or with the SNA recommendations.

13. Chart 2 presents a similar framework that may be useful as basis for identifying units and variables of interest for statistics on vocational training, and for defining relationships between the units. Here the objective is to illustrate how this may depend on whether or not the issues to be described involve more than one time period, e.g. whether those who complete a training activity have a career that is different from those that never entered the training activity or those who dropped out.

Coordinating variables and classifications

14. Coordinating variables and classifications will be the main instrument for ensuring that different areas of 'social' statistics will be coordinated and coherent, and that these areas when needed can be coordinated with 'economic' statistics (e.g. for statistics on government and households expenditures and on the wages of workers and income of households) and environmental statistics (e.g. on the consumption of households and the amount and type of garbage generated, and on employment in activities that are considered to have significant impact on the environment). However, such coordination can only succeed if it, for each variable and classification, is based on a clear understanding of the descriptive and analytical needs of the different subject matter areas. It is particularly important to understand when a particular set of definitions and classifications developed and agreed upon in the context of a particular subject matter area may be relevant for use for other purposes, and when some other concept may have to be developed.

15. One example of this issue is how to treat 'second hand' goods when describing households' consumption: In the SNA it is recommended that expenditure on such goods should be recorded 'net', i.e. the receipts from sales of such goods should be deducted. This makes sense when looking at households as an aggregate collective and measuring the consumption of such goods for households in the aggregate. However, it seems much less clear that the 'net' treatment is the best when looking at individual households, their receipts and expenditures and how these are distributed over households, i.e. a micro perspective. Then it may be more relevant to see the sale of 'second hand' goods as a source of funds for the household, and the purchase of such goods as an issue of quality and price for the type of good purchased, in which case it is relevant to record receipts and expenditures separately for such goods, e.g. in a household income and expenditure survey (HIES), even if the sale of second hand goods must be seen as a reduction in assets and not as 'income'[6].

16. Another example is the measurement of 'labour force status', i.e. as 'employed', 'unemployed' or 'not in the labour force': For the purpose of monitoring (current) developments in the labour market it is accepted that the preferred reference period should be short as well as recent. One important reason for this is that this can be done more precisely than the measurement of similar concepts for longer reference periods[7], another is that variations in total 'employment' and 'unemployment' for different groups between short reference periods in itself

6 This implies that it will be an advantage to include in a HIES a (complete) registration of savings and assets of different forms, or at least those which can easily be sold.

7 See e.g. Hussmanns et al (1990) and United Nations & ILO (2002).

is important to observe. However, used as a background variable for the description and analysis of other issues, e.g. consumption or time use behaviour, criminal activities, fertility or need for social services it may be more relevant to look at the labour market experience over a longer reference period. One way of doing so will be to use the concept 'usual activity' with respect to a longer reference period, see e.g. ILO (1983), but for certain purposes it may be more relevant to use a variable that tries to reflect the total labour market experience, i.e. a combination of the 'employment', 'unemployment' and 'not active' episodes that a person may have experienced during the reference period. One possible variable of this type was suggested in Hoffmann (1996).

17. Thus one of the challenges for the custodians of international statistical recommendations will be to identify when existing recommendations can be applied outside the area of application for which they were originally developed, and for which; and how to work together to develop other solutions elsewhere: preferably solutions that will build on the recommendations that already exist and still be valid for the intended areas of application.

Location in time and space

18. Supplementing the above discussion is the observation that common reference periods as well as common geographic references may provide some degree of coordination between different areas of statistics, because groups that are in the same time and geographic location will be subject to similar influences and events (at least potentially) and their behaviour may impact on the same phenomena.[8].

Concluding remarks on practical implications

19. This note has tried to argue that the work to create better coordination and integration of 'social' statistics cannot expect to receive much help from existing systems for classifying the fields that fall under this heading, because they have been created to serve organizational purposes rather than to support analytical and descriptive needs. Nor can much help be found in conceptual frameworks of the type represented by e.g. SNA-93 for economic statistics, because the analytical and descriptive needs to be served by 'social' statistics are dominated by different types of micro oriented issues (with the exception of demography as a discipline for the study of population dynamics).

20. The main starting points for coordination and coherence in 'social' statistics therefore must be: (1) that the use of such statistics is predominantly to describe and analyse micro issues, i.e. the situation and behaviour of individuals and the events and institutions that influence them; and (2) to carefully examine the way the basic elements for such statistics are constructed, i.e. the definition of population and units, reference periods, variables and classifications (value sets), as well as clear definitions of the relationships between different units of observations as basis for assigning variable values to them. Coherence in time and geographic references will be important for certain areas and issues, as will be the fact that activities and many events are using

8 This must not be confused with 'ecological fallacy': i.e. the idea that e.g. a behaviour or experience which is 'typical' or common for a group can be expected to hold for any particular member of the group.

time, all individuals' ultimate limited resource. 'Complete' coherence and coordination cannot be achieved, and should not be aimed for: relevance and validity will have to be given priority in many situations.

21. The ILO work to provide a basis for producing labour statistics from different sources that are integrated and coherent, presented e.g. in Hoffmann (2000) and in ILO (2000), has to a large extent been based on the ideas presented above. ILO is also very aware of the fact that labour statistics straddle any dividing line that it may be convenient to draw between 'economic' statistics and 'social' statistics, not least because 'labour' and the 'human capital' embodied in its population will represent the most important productive resource for all nations, and income arising from their employment will be the most important source of livelihood and welfare for most of the world's inhabitants. In its work to improve and develop statistical standards of importance for labour statistics ILO therefore has endeavoured to make certain that these standards will be relevant for the whole range of statistics for which they may be used, not only for those areas which are seen as their primary applications. Similarly ILO has tried to ensure that the standards for which it serves as the custodian are closely coordinated with the standards developed for other areas. Some examples may be illustrative:

(a) The scope of the definition of 'employment' is directly linked to the production boundary as defined in SNA (see e.g. Hussmanns et al, 1990).
(b) In the *International Standard Classification of Occupations (ISCO-88)* the definitions used refer directly to ISCED and ISIC wherever relevant. (see e.g. ILO, 1990).
(c) ILO's work on defining the 'informal sector' for the purposes of measuring employment, and its work to revise the *International Classification of Status in Employment (ICSE-93)* contributed directly to corresponding parts of SNA-93 (see e.g. ISWGNA, 1993)
(d) ILO has worked with UNSD to develop guidelines on how to code 'industry' reliably and effectively in household based data collection (see e.g. UN & ILO, 2002)
(e) ILO has worked with UNSD to ensure that ISIC, rev. 3.1 and the future ISIC, rev. 4 will provide a better basis describing the structure of 'informal sector' activities and households' productive activities than was possible with earlier versions of ISIC.
(f) In its work to gather statistics on employment in the public sector ILO used the definitions provided by SNA-93 to indicate their desired scope. When examining the quality and limitations of these statistics some issues concerning the delineation of contractual arrangements for employment have been identified that may have serious implications for some key concepts in the SNA (see e.g. Hammouya, 1999 and Hoffmann, 2001).
(g) In its work to develop recommendations for statistics on occupational injuries ILO has made certain that they are consistent with relevant standards developed by the *World Health Organization (WHO)* (see e.g. ILO, 1998).
(h) ILO and WHO has started a dialogue on how to best delineate "human resources in health (HRH)" on the basis of the current versions of the international standard classifications for economic activities (ISIC), occupations (ISCO) and education (ISCED) respectively. This work may lead to suggestions for modifications to these classifications as well as to the delineation of HRH that is presented Annex A.1 in OECD (2000)[9].

22. It seems reasonable to expect that better integration and coordination of the various areas of 'social' statistics, and of 'social' statistics with relevant parts of 'economic' and

9 ILO's Bureau of Statistics was not consulted during the work presented in this publication.

‘environmental’ statistics will be achieved mostly through careful considerations of such issues when planning statistical data collections and developing the necessary tools. The responsibility of the statistical units in international and regional organizations will be to cooperate when developing statistical standards and tools such as model questionnaires, as well as data collection programmes, and to strongly advocate the same cooperation and coordination at the national level when delivering technical cooperation and advisory services.

References

Hammouya, M. (1999): *Statistics on Public Sector Employment: Methodology , Structures and Trends*. Sectoral Activities Programme and Bureau of Statistics. Working Paper SAP2.85/WP.144. International Labour Office, Geneva, 1999.

Hoffmann, E. (1995): "We must use administrative records for official statistics – but how should we use them?" *Statistical Journal of the United Nations ECE,* **12**, pp. 41-48.

Hoffmann E. (1996): "Requirements and possible sources for statistics on the dynamics of employment: a producer perspective" *Statistical Journal of the United Nations ECE,* **13**, pp. 335-347.

Hoffmann E. (2000): "Developing labour account estimates: issues and approaches" in United Nations Statistics Division: *Household Accounting: Experience in Concepts and Compilation. Volume 2: Household Satellite Extensions*. Handbook of National Accounting. Studies in Methods. Series F, No. 75 (Vol. 2) United Nations, New York. pp. 21-40.

Hoffmann, E. (2001): *Statistics on public sector employment: a review of quality issues*. Note presented to the 11th Statistical Days: A New Millennium – New Phenomena: Have statisticians been able to understand and measure them. Radenci, Slovenia, 26-28 November 2001.

Hoffmann, E. & Lawrence, S., (1996): *Statistics on International Labour Migration: A review of sources and methodological issues*. Interdepartmental Project on Migrant Workers, 1994-95. International Labour Office, Geneva, 1996.

Hussmanns, R., F. Mehran and V. Verma (1990): *The economically active population, employment, unemployment and underemployment: An ILO-manual on concepts and methods*. International Labour Office, Geneva.

ILO (1983): *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment*, adopted by the 13th International Conference of Labour Statisticians. See ILO (2000): *Current international recommendations in labour statistics* (2000 edition). International Labour Office, Geneva.

ILO (1990): *International Standard Classification of Occupations (ISCO-88)*. International Labour Office. Geneva.

ILO (1998): *Resolution concerning statistics of occupational injuries (resulting from occupational accidents*, adopted by the 16th International Conference of Labour Statisticians. See ILO (2000): *Current international recommendations in labour statistics* (2000 edition). International Labour Office, Geneva.

ILO (2000): *Developing a labour accounting framework for tourism: issues and approaches*. Paper annexed to *Note by the Secretary-General to Item 3 (b) of the provisional agenda for the 32nd session of the UN Statistical Commission*, 6-9 March 2001. E/CN.3/2001/3. United Nations. New York

ISWGNA (1993): *System of National Accounts 1993*. United Nations et al. New York

OECD (2000): *A System of Health Accounts*. Organization for Economic Cooperation and Development. Paris.

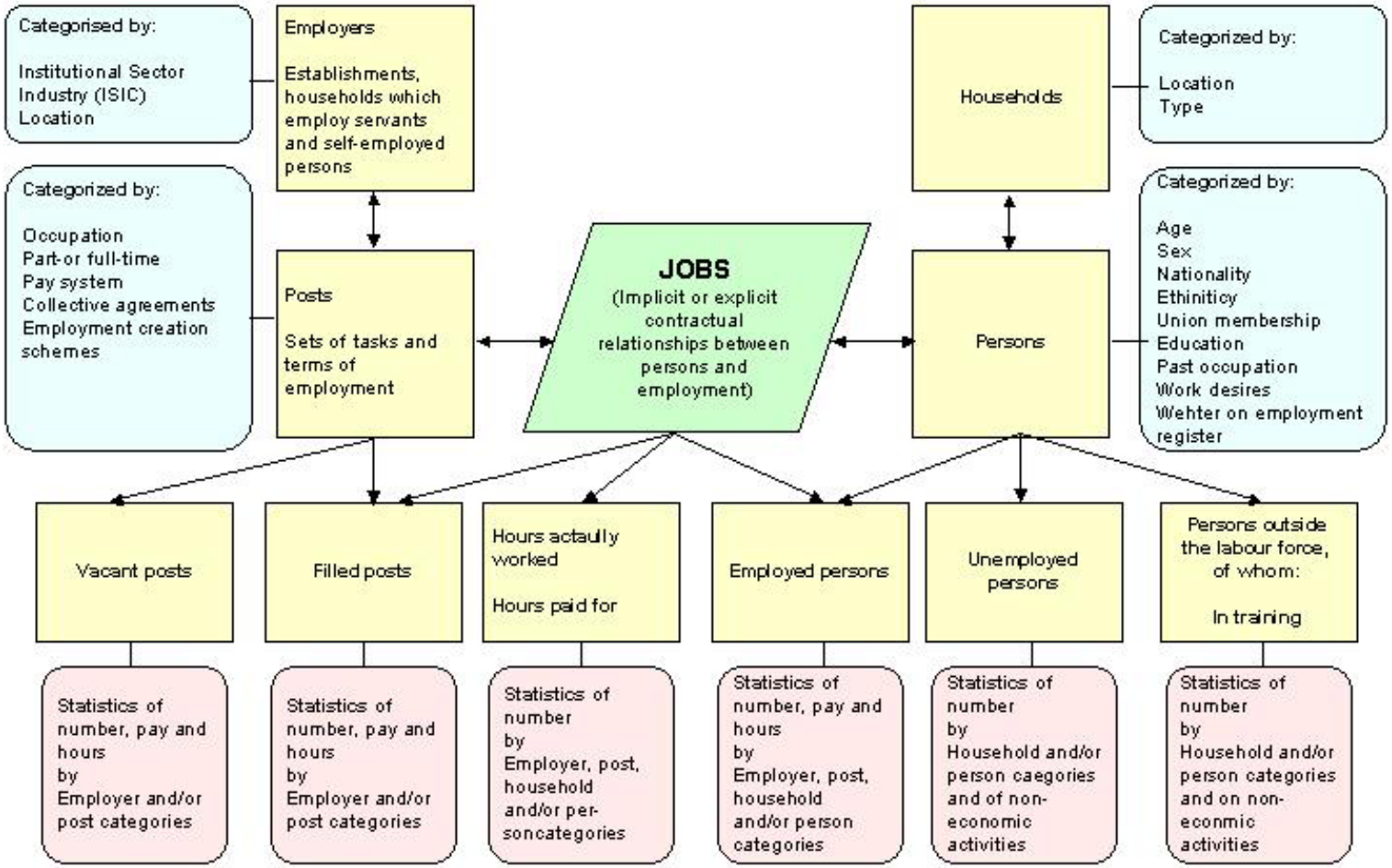
United Nations (1974): *Towards a system of social and demographic statistics*. Sales No. E.74. XVII.8. New York

United Nations (1979): *The development of integrated data bases for social, economic and demographic statistics*. Studies in Methods, Series F No. 27. United Nations, New York

United Nations (1984): *The development of environment statistics.* Statistical Papers; Series M, No. 78. United Nations, New York.

United Nations & ILO (2002): *Collection of Economic Characteristics in Population Censuses*. Technical Report, ST/ESA/STAT/119. United Nations, New York.

Chart 1. Conceptual Framework for a Labour Accounting System



Note: « Numbers » means stocks at a point in time, average over a time period or total over a period (for flows, e.g., hours actually worked or paid for). The statistics on total compensation by institutional sector and by industry provide links to the system of national accounts.

Chart 2. Framework for identifying units and variables for voca