



**Economic and Social
Council**

Distr.
RESTRICTED

CES/1996/R.19
28 March 1996

ORIGINAL: ENGLISH

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

Forty-fourth plenary session
(Paris, 11-13 June 1996)

**REQUIREMENTS AND POSSIBLE SOURCES FOR
STATISTICS ON THE DYNAMICS OF EMPLOYMENT:
A PRODUCER PERSPECTIVE**

Submitted by the International Labour Office¹

A. Introduction

1. The objective of this paper is to present a framework for discussion of the requirements which the users of statistics on the dynamics of employment would like these statistics to satisfy, with some observations on the possibilities of data collection instruments for meeting these requirements.

2. The term 'dynamics of employment' (DE) is used here to refer to any type of change in the 'activity status'¹ of individuals, or to a change of jobs held by employed persons. Statistics on DE will either describe the number of persons who have experienced relevant types of changes in their 'activity status' or jobs, or they will describe the duration of completed spells in status or job situations. Statistics on consequent changes in descriptive variables for jobs (e.g. 'occupation', 'status in employment', 'wage rates', hours worked') and establishments (e.g. 'size', 'industry', 'sector') may also be considered as DE related.²

¹ Prepared by Eivind Hoffmann; Bureau of Statistics; International Labour Office; CH-1211 GENEVE 22; Switzerland. Fax: + 4122 799 6957. The views expressed are those of the author and they are not necessarily shared by the ILO or its Bureau of Statistics. Comments and corrections are welcomed.

3. Indicators of DE can be derived from registrations of events which results in statistics on e.g. the number of hirings and separations of persons or the number of 'births' and 'deaths' of enterprises during a reference period. However, the discussion in this paper focuses on data collection situations where it is possible to determine whether two events involve the same or different persons or establishments. Such observations can either take the form of 'process registrations', i.e. registration of the whole sequence of events and situations for persons or establishments, or take the form of a 'sequence of snapshots', i.e. by observation of the same units at different points in time.³ Both 'process registrations' and 'sequences of snapshots' can either be done retrospectively or by following individual units over time.

4. A discussion of what DE statistics users require could be organised according to the different types of users, e.g. the 'general public', 'politicians', 'planners' in public and private sector organisations, persons with responsibilities for budgets and administration, 'researchers'. However, it seems simpler to organise the discussion with respect to the following broad types of use of the statistics:

- a. to formulate, implement and evaluate policies designed to influence DE;
- b. to monitor actual developments with respect to DE;
- c. to describe and analyze DE structures and behaviour.

5. Sections B-D discuss the type of DE statistics needed for each of these types of use, with respect to the following quality dimensions

- population coverage;
- units of observation;
- timeliness and frequency;
- geographic resolution;
- consistency with other statistics and over time;
- main and descriptive variables, in terms of
 - validity of definitions;
 - resolution and validity of value sets;
 - reliability of measurements.

6. Section E contains some observations on the potentials of the different forms of sources, i.e. household and establishment surveys and direct use of administrative records, for providing DE related statistics, before making some concluding remarks and suggestions in section F.

B. Statistics for formulating, implementing and evaluating policies designed to influence DE

7. These policies are designed to influence the number of persons who move between the different types of activity states as well as the type and timing of such moves: e.g. to delay or hasten the flow of workers out of employment into retirement; to facilitate the transfer into employment of persons who are unemployed or outside the labour force; to ease the transfer to employment of persons leaving education or compulsory military service; to

facilitate the transfer between employers of persons already employed; to facilitate the establishment of new businesses by persons who are unemployed. The designers and administrators of such policies will want to assess in advance e.g. the number of persons who potentially will be affected by the policies or the number of persons who may become, or may want to become, clients of the schemes to be initiated. Such estimates will provide a basis for determining the funds and the administrative capacities needed to implement the policies under consideration or to be implemented.

8. The population to be covered by statistics needed for this type of assessment of scope and consequences of planned policies will normally depend on the specifics of the laws and regulations which provide the foundation for the implementation of the policies. These laws and regulations will often include particular groups of persons in the schemes and exclude others. The groups may be defined by their activity status or personal characteristics, as well as by characteristics of their place of residence (locality) or type of employer (past or prospective).

9. The preferred units of observation and units of measurement will depend on the objectives and operational basis for the schemes: Establishment oriented policies will normally prefer statistics based on information from establishments, providing only summary data on the number of 'workplaces' or 'work years' which may be affected. Concern with policies to assist persons would normally lead to requests for statistics (which have) to be based on information from individuals, and which can lead to estimates of the number of persons (expected to be) affected, characterized by descriptive variables such as 'age' and 'sex'.

10. The demand for consistency with other areas of statistics is normally modest, even if under certain circumstances there are requests for statistics which can be directly linked to information from public sector budgets and accounts. The need to have consistent time series is also limited, as the main concern is that the statistics should in some way reflect the 'current' situation and provide a good basis for assessing the budgetary and administrative capacity needs in the immediate future, as defined by the government's budgetary cycle and planning horizon. Time series may, however, be important if the scope of the scheme is such that the number of 'clients' may fluctuate rapidly, e.g. in accordance with seasonal or business cycle variations.

11. The explicit demand for precision in the estimates are normally fairly modest: 'order of magnitude' estimates are frequently said to be sufficient. However, estimates may be needed for groups which must be considered 'small' in the context of the possible data collection methods. Thus even an 'order of magnitude' estimate may imply a very precise estimate in this context.

12. That the timeliness and frequency demand of statistics to be used as a basis for policy formulation will tend to be modest is a consequence of the modest demand for precision. However, the implementation and evaluation of these policies may require observations which are both very timely and frequent, if e.g. the amount of funds dispensed are supposed to fluctuate

with the 'current needs' in the labour market and if these vary over the business cycle.

13. The value sets of the descriptive characteristics are normally used to define 'insiders' and 'outsiders' with respect to a particular set of policies, which is a "simple" binary value set. The quotes around 'simple' reflect the fact that the observation of the dividing line may be quite complicated both from a definitional and an implementation perspective, and may require some very precise measurements of e.g. 'industry', 'occupation' or 'geographic location' to ensure that the statistics provide distinctions corresponding to the intended groups.

14. The overall impression is that the demands on the statistics on the DE are relatively modest for policy formulation and implementation, with the exception of those which arise from the need to cover particular population groups. In general, users seem satisfied with statistics which provide 'order of magnitude' estimates and a broad structural picture. Policy evaluation will, however, frequently entail the same type of demand on the statistics as when they are used for analysis of the corresponding phenomenon.

C. Statistics for monitoring developments

15. Traditionally the statistics used to monitor the development in labour markets, whether in the aggregate or with respect to particular sub-markets, have been used as indicators of macro-economic developments and tensions. They serve as basis for judgement about the outcome of and need for adjustment in general activity regulating policies. The focus is on net changes in the overall unemployment rate, in the number of long term unemployed persons and in the number of persons employed, and none of these indicators reflects the dynamics of employment: they reflect only the outcome of these dynamics. However, it would seem reasonable to expect, however, that indicators which reflect DE will increasingly be demanded also for monitoring, if such indicators can be produced regularly and if the commentators and policy makers feel that they understand the relationship between the policy instruments and the movement of indicators such as: 'the number of persons who have changed jobs during the last x months'; 'the number of persons who have left the labour force' from employment and unemployment respectively; 'the number of persons who have entered the labour force' as employed and unemployed respectively; 'the number of establishments which have increased and which have decreased the number of persons employed, on own account or through sub-contracting'.

16. Monitoring activities will prefer to use statistics which can reflect the development of the total population of persons and establishments, not only those units which can be most directly affected by policies. This because the concern is with the overall developments with respect to employment, as a consequence both of general government policies and of influences beyond the control of governments. However, as compromises often will have to be made between having statistics which reflect all units and statistics which can satisfy other requirements, e.g. providing DE data which

are timely and consistent with data on changes in production, trade and investments, priority will tend to be given to statistics for those (parts of the) labour markets which are seen as particularly sensitive to business cycle variations, whether defined in terms of sectors, occupations or regions. The most relevant units of observation may be establishments in these cases and 'work hours' should supplement 'persons' as unit of measurement, e.g. where the relevant concern is with the analysis of factors influencing the development of productivity over the business cycle.

17. The main types of descriptive variables required for monitoring are already alluded to above, i.e. 'industry', 'occupation', '(type of) region'. These will be supplemented by the standard demographic variables 'age' and 'sex', and may be joined by variables related to 'type of training' and 'labour market experience', 'nationality' or 'ethnic group', depending on the type of current policy concerns. Variables related to 'working time arrangements' and 'status in employment' may also be considered important for the monitoring of certain forms of DE.

18. The concern with consistent time series is particularly strong among those monitoring current and longer term developments. Their concern is frequently with the changes which have taken place over the last month, quarter, year or five years, and with whether such changes are becoming more or less pronounced or are in a new direction. Consistency in time series therefore frequently tends to be given higher priority than improvement in the statistical series, even when such improvements are necessitated by changes in the underlying reality which over time have undermined the ability of statistics produced according to the previous methods to reliably reflect what is happening. The importance of estimates of changes will often imply very stringent demands for precision in the measurement of the main variables, in particular when these changes have to be estimated as differences between observed levels. The frequently expressed position that "the precise numbers do not really matter, as long as the order of magnitude is correct" therefore cannot be taken seriously, unless it is quite clear that this "order of magnitude" position is consistent with the way the statistics are being produced. Requests for information about detailed groups according to the descriptive variables will, of course, further complicate matters.

19. The choice of reference point or period also has to be made with regard to the rhythm of the DE under consideration: If there is a tendency for transfers into employment and out of unemployment to happen on certain week days and immediately after or before moveable holidays such as Easter, then the estimated changes may be dominated by such patterns and therefore be highly misleading as indications of current developments, e.g. if the last working day of the month is used as the reference day for monthly or quarterly statistics.

20. There is evidence to support the belief that in many respects the demand for statistics is a function of the supply. Drawing on the current experience of the demand for statistics for monitoring employment and unemployment, there is reason to expect that if and when DE descriptive

statistics can be produced on a regular basis, then users' requirements may prove to be operationally incompatible: high resolution of certain descriptive variables (e.g. 'region', 'industry', 'occupation') used simultaneously; precise estimates of changes; high degree of frequency and timeliness. It therefore seems likely that the short term monitoring of DE will be based on partial indicators which can be produced quickly, e.g. 'inflow to registered unemployment'; 'outflow of registered unemployment' (by destination if possible); 'advertised vacancies'. More complete and detailed statistics will have to be based on less frequent data collection programmes.

D. Statistics for description and analysis

21. It may seem fairly meaningless to try to outline in general terms the statistics required for description and analysis of DE. The relevant theoretical and methodological traditions and perspectives range from economics to social anthropology; from micro to macro perspectives; from national to local; from anecdotal quoting of numbers to formal statistical testing; and from simple cross tabulations to formal models with hundreds of relationships and thousands of parameters to be estimated. The following reflections stay away from the most extreme positions on these dimensions when trying to outline some patterns.

22. The population to be covered will depend upon the issue to be studied. However, the point of departure often is, or should be, the total population of individuals, households and establishments, even if the analysis is intended to focus on sub-groups defined by characteristics related to 'degree of involvement with the labour market', 'status in employment', 'industry', 'occupation', 'region', 'educational attainment', demographic characteristics or labour market experience. Many studies which focus on a particular sub-group will also need statistics for a reference population, which can be the total population.

23. The descriptive variables requested are almost as numerous and varied as the number of different issues to be analyzed, but certain variables will be important for a wide range of issues and analytical approaches: For issues related to the use of labour as a factor of production the central variables will be 'wage rates', 'industry' and 'occupation'. The latter variable, as well as variables such as 'educational attainment', 'age', 'sex' and 'labour market experience' will be central when concerned with possible improvements, or otherwise, in the quality of employed labour. For gender related issues and issues concerning the dynamics of moving between employment and outside the labour force relevant variables for persons in both groups will be related both to the respective labour markets, i.e. variables such as 'wage rates', 'working time arrangements', 'industry', 'region' and 'occupation', as well as to the personal and household situation, e.g. variables such as 'labour market experience', 'educational attainment', 'sex', 'age', 'marital status', 'number of children' and 'age of the youngest child', in addition to 'educational attainment', 'activity status', 'income' and 'working time arrangements' of the spouse and other household members. Based on theories of dual or fragmented labour markets

one may expect requests for descriptive variables which can reflect characteristics of the establishments in which the persons have been employed, e.g. 'industry', 'sector', 'size', and the career potentials of different jobs. 'Commuting regions' will be a relevant variable for work with issues related to regional labour markets and the situation and development within sub-national administrative regions. Variables related to the work environment and physical working conditions and work postures will be relevant for the analysis of DE phenomena which may lead to the exclusion of certain types of people from the labour market.

24. Concern with consistency between various statistical series will also differ considerably with the issues being investigated and with type of analytical approach. In economic analysis consistency with statistics on 'production', 'income' and 'wage rates' will normally be seen as particularly important, while other perspectives will see consistency with e.g. demographic, health and training statistics as more relevant. The consistency requirements will also be expressed with respect to time references and geographic delineations.

25. The demand for coherent time series arise in two different contexts: (i) Data referring to t different points in time are used to represent t different manifestations of the same underlying phenomenon or process. The longer the time series, i.e. the larger t is, the better will be the precision with which one can expect to describe this reality, provided that it has been stable over the observation period or has changed in ways which are known. (The estimation of econometric models from time series observations is an example.) (ii) Studies which are concerned with the development over time as such, e.g. as basis for projections into the future on the basis of observed trends and variations.

26. The needed precision of measurement of the various variables and their degree of resolution will be directly related to the issues being investigated. It would seem, however, that the need for precise and detailed measurements is more acute when the statistics are being used to estimate quantitative relations in complex models than when they are to be used for simpler cross tabulations.

27. Analysis which is expected to yield results which are to bear directly on policies will normally require timely statistics. However, overall timeliness seems generally to be of less concern in this context than most of the other quality considerations. The periodicity requested for time series studies will normally be related to the a priori assumptions about the speed of significant changes in the phenomenon under study.

28. It does not seem possible to present a uniform and coherent picture of the statistical needs of analysts, and the expressed demand does seem very much to be a function of supply. Some research activities are designed to support monitoring activities, and will therefore have corresponding demands on the statistics to be used, although the need for very recent statistics will normally be less. Otherwise the most dominant complaint concerning existing statistical programmes tends to be that they do not include the

large number of variables which the researchers would like to be able to use for their investigations.

E. Sources for DE statistics

29. In paragraph 2 above reference was made to two different forms of observation for DE, namely process registrations (PR) and sequence of snapshots (SS). In principle, either type can be carried out retrospectively or by following individual units over a period of time. The objective of this section is to briefly review the possibilities and main issues related to the use of these forms when working with the main types of data collection instruments: household surveys (HS), establishment surveys (ES) and direct use of administrative records (DUAR).

30. An HS can include a module for retrospective DE relevant PR by asking the respondent to recall all transfers between jobs and between activity status situations experienced over the last T months or years, where T corresponds to the reference period for which the statistics are supposed to describe the DE. Asking the respondent to provide information about their situation at particular points in time during this reference period will provide an SS observation set. With both approaches the main methodological problem involved is the design of the questionnaire in a way which reduces the amount of recall error below the maximum which is consistent with reliable statistics. The questions need to be formulated in a way which will overcome well known problems of accurate recall of all relevant transfers. Their timing and duration as well as characteristics of the situation before and after each transfer need to be recorded (e.g. their duration and their type: 'type of training activity' and 'industry', 'occupation' and 'status in employment'). There is no need to elaborate on those issues here.

31. It does not seem possible to implement PR by following individuals over time through an HS survey programme. However, using a panel sample for a sequence of e.g. Labour Force Surveys will provide the basis for a SS approach. This approach avoids, or at least reduces significantly, the problems of recall, but makes it particularly important that the processing of survey rounds are linked and that the questionnaires are designed to eliminate as far as possible the creation of false transfers between different situations. The length of the panel period and the number of 'snapshots' which will be made during that period will of course depend on the particular rotation scheme used.

32. When using an ES as the basis for DE statistics one has the choice of (i) asking the establishments to provide information about randomly selected posts and the corresponding staff members; or (ii) asking the establishments to provide summary information about their staff according to certain specified (cross-)tabulations. Both approaches require that the establishments keep good, easily accessible personnel records. The advantage of the former approach is that a larger range of variables may be included in the survey, that cross-tabulations are not pre-determined and that certain coding activities, e.g. of occupation and educational attainment, can be done by expert staff at the statistical office. The advantage of the latter

approach is that costly coding need not be done in the statistical office (but there may be more queries from and to responding establishments). Only the former approach will be able to provide PR based observations. Furthermore, the personnel records are not likely to provide a basis for much information about where employees came from when joining the establishment or where they went when departing.

33. Both retrospective and follow up approaches to the collection of DE data from establishments will have to cope with the effects of their complicated demographics: With a retrospective approach the point of departure will be the end-of-period population of establishments, and the population in place at the start of the reference period will have to be created through information about 'deaths' and 'births', mergers and splits. With a follow-up approach the point of departure will be the population at the start of the reference period, and rules must be formulated on how to deal with 'deaths', mergers and splits which happen during the period.

34. DUAR can be used to produce DE statistics if the administrative system has a personal identification number (PIN) for individuals and an establishment identification number (EIN), respectively, which will allow one to link registrations made at different points in time for the same units⁴. Only if the same (i.e. unique) identifier (U-PIN) is used by several relevant administrative registration systems, will it be possible to link together the appearance of these units in the different registers, to establish their situation when they are absent from a particular register. Without a U-PIN each system must limit itself to identify 'insiders' and 'outsiders', 'arrivals' and 'leavers'. While they can ask where their clients are coming from, it is normally very difficult, except through special tracer surveys, to establish the fate of 'leavers'.

35. Administrative registrations are either event based or carried out at regular intervals. The registration of persons who claim unemployment benefits is an example of the former type of process, and the registration of annual employer reports to tax or social security authorities about the earnings of persons who have been employed by them is an example of the latter. Both PR and SS are possible with either procedure, at least in principle. However, with the latter procedure PR will depend on accurate dating of events, retrospectively, by the reporting person or establishment. This is unlikely to be achieved in practice. Event based registrations can only provide accurate stock and flow estimates if care is taken to register both the date at which the event took place and the date at which it was registered. Experience has shown that with event based registrations the reports which lead to new entries into a stock for a particular reference date may arrive at a rhythm which is different from the reports of events which lead to departures from the same stock. As the production of statistics from this system will have to wait until one can be confident that a very high proportion (e.g. at least 95 percent) of all reports have arrived, this may influence significantly the timeliness of the statistics, in particular if the variation in the changes in total stock is an important concern, e.g. when monitoring business cycle developments.

F. Concluding remarks

36. Only in relation to a specified use of the statistics is it possible to set priorities among the different, possibly conflicting, quality requirements which can be set for DE related statistics. However, the following list seems to provide a fairly safe set of conclusions, as it is not very different from a corresponding list of conclusions about employment statistics in general:

- a. Statistics on the dynamics of employment must describe the type of transitions referred to in the **Introduction** to this paper, and/or the duration of the various situations.
- b. The statistics must be made available in formats which will make them easy to use by different types of users and which will facilitate the further manipulation of the data by those who need to do so.
- c. A programme of DE related statistics should consist of a few regularly produced statistical indicators, supplemented with estimates from more comprehensive investigations (based on surveys and/or the processing of administrative records) carried out at longer intervals.
- d. The main objective of the regularly produced indicators should be to provide a basis for the monitoring of current developments and for the type of research which will assist to strengthen the capacity to understand these developments.
- e. The main objective of the investigations carried out at longer intervals should be to describe and analyze structural aspects of DE, as well as changes in these aspects.
- f. The timeliness and precision requirements specified for DE related statistics must be realistic in the context of the available data collection instruments.
- g. It will be important to have consistency between the regularly produced statistics and the more comprehensive statistics produced at longer intervals. Such consistencies should be aimed at both for the main variables, i.e. those which are used as indicators of the transitions representing the DE, and for the descriptive variables.
- h. The potentials and limitations of the various data collection mechanisms must be carefully considered with respect to the possibility of obtaining complete process registrations or only approximations through sequences of snapshots. In principle the optimal choice will be different for different use of the resulting statistics. However, it will normally not be possible to tailor separate data collection and processing programmes to the different uses. The statistics producers and the users must therefore cooperate to establish the extent to which particular statistical products can satisfy the needs, to develop the best way of compensating for weaknesses, and to identify and explain inappropriate types of use of a particular statistical product.

37. The discussion above has mostly implied concern with the dynamics of changing between being 'employed' (E), 'unemployed' (U) and 'outside the labour force' (O), or with the change of employers and/or jobs for those who are 'employed', using the concepts of the ILO Labour Force framework. This will represent an extension of this framework beyond the short reference

period for which it was developed. This extension opens possibilities which are different from the concept of 'usual activity' that so far has been used to overcome limitations of this framework for those who need statistics which describe the situation with respect to a longer reference period. These possibilities are particularly interesting because experience seems to indicate that the concept of 'usual activity, and the categories specified for it, does not correspond well to the concerns of those would like to describe or refer to the labour market situation of persons with reference to a longer period. Their concern seems to be less with describing that situation in terms of just E, U and O as dominant situations than with expressing the totality of the experience over the reference period.

38. One strategy for achieving this may be to define a variable 'pattern of activities' which will summarize the labour market experience of persons over this longer reference period, not in terms of 'main' or 'usual' activity, but in terms of sequences and duration of activity states. This variable, which conceptually is closely related to the 'attachment variable' defined for the Swedish Labour Force Survey, would obviously need to have a larger value set than 'current activity', in order to reflect the patterns which would be of central interest to users. Using the activity at the start of the period as the main organising feature the following typology is presented as a possible point of departure for further discussions:

- a. Employed whole period: same job;
- b. Employed whole period: changed industry at least once, same occupation;
- c. Employed whole period: changed occupation at least once, same industry;
- d. Employed whole period: changed both industry and occupation at least once;
- e. Employment followed/interrupted by at least one unemployment spell;
- f. Employment followed by spell outside labour force, no unemployment spell;
- g. Unemployment followed/interrupted by at least one employment spell of t weeks or more;
- h. Unemployment followed/interrupted by at least one spell outside labour force of t weeks or more;
- i. Outside labour force followed/interrupted by at least one spell of employment lasting t weeks or more;
- j. Outside labour force followed/interrupted by at least one spell of unemployment lasting t weeks or more;
- k. Unemployed whole period;
- l. Outside labour force whole period: at least t weeks in training;
- m. Outside labour force whole period: other situations.

39. This typology has 13 categories, and is complicated in comparison with the basic distinction between E, U and O and possibly also in comparison with a typology parsimonious enough to serve as basis for newspaper headlines and cabinet briefs. Depending on the selected value for t it could perhaps be reduced to: (i) Stable employment (a); (ii) Mobile employment (b, c, and d); (iii) Unstable employment (e and f); (iv) Long term unemployment (g, h, j and

k); (v) Outside labour force (i, l and m). However, relative to the total number of possible situations and 'careers' over a reference period of e.g. one year, the typology represented by these 13 categories is quite simplistic, in particular when considering that the concept of 'underemployment' has not been used when delineating its categories.

40. It not evident that a single typology will suffice, i.e. will be both analytically meaningful and sufficiently parsimonious to be possible to implement with the limited precision allowed by sample surveys. One alternative to the above typology would be to start from the situation at the end of the reference period when defining the categories. Groups a-d and k-m above could then be the same, but the distinctions to be made among those who during the reference period experienced some turbulence in their employment and unemployment situation would have to be different. Experiments and the reactions of users will be needed before it is possible to establish standard typologies⁵. The typologies to be developed must also be 'realistic' with respect to the known limitations with respect to achievable precision and sources of error of the various data collection instruments.

Notes:

1. 'Activity status' can take the following values: 'employed (E)', 'unemployed (U) and 'outside the labour force (O)'.
2. Statistics on DE can be seen as a subset of statistics on 'labour market dynamics', which would also include statistics describing changes in e.g. wage rates and income from employment unrelated to job changes.
3. Note that a 'sequence of snapshots' can, in principle, only result in an approximation to the true dynamics described through a process registration. However, with the 'snapshots' close enough the resulting description may be close enough for most practical purposes, cf. the principles behind the production of motion pictures.
4. Strictly speaking one only needs clear identifiers, which need not necessarily be a PIN. However, in practice the use of PINs has many advantages.
5. Strictly speaking it is not necessary to develop 'standard' typologies as compromises of partly conflicting users' requirements, if basic data registrations are sufficiently detailed. However, the standard groups serve as useful references also for those wanting typologies tailored to their specific analytical requirements.