



OECD SHORT-TERM ECONOMIC STATISTICS WORKING PARTY
(STESWP)

Task force on administrative data, overview of work done to date
2.2 General overview of the proposed administrative data handbook

Paper prepared by David Brackfield
Statistics Directorate, OECD

Submitted to the Working Party under item 2 of the draft agenda

Meeting:
26 – 28 June 2006

Franqueville Room
OECD Headquarters, Paris
Starting at 9:30 a.m. on the first day

2006 STESWP Meeting

Agenda item 2 – Task force on administrative data, overview of work done to date.

2.2 General overview of the proposed administrative data handbook

PAPER

1. Executive summary

1. The task force on administrative data has been mandated with two specific outcomes:
 - a) to produce a dynamic web-based framework of previously published administrative data papers; and,
 - b) from this first task, to put together a small handbook on administrative data good practices relevant to short-term economic statistics.

With the first task now near completion, this paper explores options on how the second task could be undertaken. To date over 80 previously published administrative data papers have been received from STESWP members for consideration for inclusion in the web-based framework. A paper was presented at the 2005 STESWP meeting outlining various directions and timelines that the proposed handbook might take, however this was before the wealth of administrative data papers were received and as such, the ideas outlined in that paper and those previously expressed at the STESWP meetings needs to be revisited.

2. Included among the administrative data papers provided by STESWP members were the following:

- a) Use of administrative sources for business statistics purposes, Handbook of Good Practices¹; Eurostat;
- b) Labour statistics based on administrative records: Guidelines on compilation and presentation; ILO;
- c) Use of registers and administrative data sources for statistical purposes, Best practices of Statistics Finland;
- d) Business register recommendations manual – Chapter 20, The use of administrative sources; EuroStat;
- e) We must use administrative data for official sources – but how should we use them; ILO;
- f) Issues in the use of administrative records for statistical purposes; Statistics Canada.

3. The first major issue that the 2006 STESWP members may wish to consider is:

- Is there an actual need for an additional OECD administrative data handbook given all the quality papers that are already available in the field and have partial relevance to short-term economic statistics as seen in paragraph 2?
- Should more focus be given to enhancing the web-based framework that has been established so it could become an effective tool for NSIs using administrative data, thus making the proposed handbook obsolete?

If STESWP members believe that there is a need for such an administrative data handbook, then the purpose and content of such a handbook needs to be clearly specified prior to its compilation. As such,

¹ This handbook contains a detailed description of the administrative data used in the NSIs of 17 EU member countries.

STESWP members are requested to provide comment on the various options presented below in Sections 3 and 4 of this paper.

2. Introduction

4. The administrative data task force was created from discussion during the ‘Future Topics’ agenda item of the 2004 STESWP meeting. The 2005 STESWP meeting saw the focus of the project defined both from discussion during the meeting and from the large number of written response received from STESWP members before the meeting. The task force was asked to focus on two outcomes:

- Creation of a dynamic web-based framework that will bring together administrative data issues as faced by NSIs producing short-term economic statistics using or about to use administrative data. As part of this output, STESWP members requested that this web-based framework should try and maintain consistency with the STES Timeliness Framework² created by a previous STESWP task force.
- Using the web-based framework as a foundation, the task force was then asked to create a very small guide (or handbook) summarising best practice for NSIs when using administrative data. In particular the handbook should only focus on good practices when using administrative data in the compilation of short-term economic statistics.

5. This paper explores the options available to the task force to meet the second output – the production of an administrative data handbook. During late 2005 and early 2006 a dynamic web-based framework of administrative data papers was created on the OECD website and is now available at: <http://www.oecd.org/std/research/administrativedata>
It was originally thought by the task force and STESWP that this web-based framework resource would provide the basis for the handbook, however, this was before the number and quality of administrative data papers received and reviewed was known.

6. With over 71 administrative data papers available on the web-based framework covering a wide range of topics, including good practices, any administrative data handbook to be produced by STESWP must be well defined.

3. Handbook Types

7. Assuming that there is general agreement that there is need for an administrative data handbook of good practices in short-term economic statistics and assuming that the handbook will not just be a summary of the previously published generalised handbook papers; then some general themes will need to be explored and commented on by STESWP members. What follows are some general outlines of approaches that could be taken for the proposed handbook.

3.1 Possible handbook options

8. One option would be to produce a handbook designed for NSIs looking at the option of using administrative data in the compilation of short-term economic statistics for the first time. While it would be rare for an OECD Member NSI never to have used administrative data before, there may be some NSIs

² www.oecd.org/std/research/timeliness

who have only used administrative data in registers or for structural statistics and who are now looking to start exploiting administrative data for short-term economic statistics. In this case the proposed handbook could provide these NSIs with a list of good practices of how to undertake this move into short-term economic statistics, mostly through the use of flow charts and action item lists.

9. Another option could be a handbook that is focused for NSIs that are presently only using administrative data in a limited fashion in short-term economic statistics, for example, just in sample frames, but now want to further this use. A number of the papers received for the framework dealt specifically with the issue of how to use administrative data in the actual production of short-term economic statistics. In particular, a number of NSIs undertook redevelopment of their Retail Trade survey with the express outcome of using more administrative data in the compilation.

10. The proposed handbook could be more focused on NSIs that are already using administrative data and who want a check-list of sorts just to confirm that they are utilising their administrative data fully and efficiently in the compilation of short-term economic statistics. This would be aimed at those NSIs who are looking to undertake reviews and/or redevelopment work on their administrative data use in short-term economic statistics and want a short handbook of good practices that provides a check-list of sorts. These NSIs would already be experienced administrative data users who could benefit from seeing what other NSIs like themselves have done to improve their usage of administrative data.

11. A fourth option could be for the proposed handbook to focus on a particular issue involving the use of administrative data in short-term economic statistics. To some degree this approach was taken by the Eurostat handbook that has a considerable section of the paper devoted to the issue of Statistical Laws (the legal framework) and their interaction with administrative bodies. In this vein the Eurostat paper also goes into detail about the relevant EU Council Regulations that are applicable to the issue and how countries can use them. This approach could possibly be taken in a new handbook focusing on such issues as:

- Cooperation between NSIs and administrative agencies, i.e. administrative protocols, service level agreements, and memorandums of understanding.
- Specific short-term economic statistics such as Retail Trade or Manufacturing. There are a number of papers from NSIs that go into detail about their experiences in using administrative data for these specific short-term economic statistics.
- Data validation tools – due to the sheer size of the administrative datasets and the requirement for timely publication in the world of short-term economic statistics, NSIs have been forced to develop data validation tools when using administrative data with short-term economic statistics. Such a topic could also include the graphical interfacing, modelling, forecasting and imputation methods used with administrative data in short-term economic statistics.

4. Handbook – General; Comments Required and Future Work

12. Another question that STESWP members may wish to comment on is the general layout of the possible proposed handbook. With the number of papers received and the quality of their information on administrative data and its uses, especially in short-term economic statistics, there is a good case for making the proposed handbook include web-links to these papers. This would give the user access to more in-depth information that could be followed-up, at a time suitable to them, if they desire. This would indicate that the proposed handbook would be more of an electronic creation rather than strictly paper-based. It has also been the intention from the beginning of the project that the proposed handbook should

be small and the STESWP members should keep this in mind when making comments or giving recommendations.

13. As discussed earlier, while there are no administrative data handbooks focusing exclusively on short-term economic statistics, there are a number of administrative data handbooks or very good general papers. Annex 1 presents some examples of the type of information these papers contain, these examples may give the STESWP members some help when formulating their comments to the administrative data handbook project. In particular, if it was thought that there was no added value in producing a handbook then STESWP members might want to comment on how the web-pages of the administrative data framework could be transformed to try and meet some of the needs first considered for the handbook.

EXTRACTS FROM PREVIOUSLY PUBLISHED ADMINISTRATIVE DATA HANDBOOKS

Eurostat – Business Register Recommendations Manual – Chapter 20 The Use of Administrative Sources – Status: Sixth Draft – December 2002

20F - Using Administrative Sources in Practice

20.25 The following paragraphs consider the processes that are often necessary to transform data from administrative sources into the variables required for statistical business registers. The main processes are listed and briefly explained below.

• *First steps and the transition from administrative to statistical data*

20.26 The first steps generally concern processing the administrative data to check their quality and coverage. Preliminary analyses may be useful, or even necessary, to check several points including:

- The main characteristics of the administrative data e.g. date of reference, the total number of units, procedures used to update the source, etc.
- The coverage of the main variables e.g. identity number, address, economic activity codes, number of employees, sales space, products sold etc.
- The error profile of these variables, especially the variables needed to determine continuity of statistical units (e.g. identity numbers, economic activity codes, location, dates, etc.)

ILO – Compilation and presentation of labour statistics based on administrative records; by R. J. Pember

Page 3.

“Review of strengths and weaknesses of administrative records as a source of statistics

The compilation of statistics from administrative records has several advantages as follows:

- (a) Low data collection costs for the statistics-producing agency: Since the data has already been collected as part of an administrative function, there are few costs in accessing the data for statistical compilation.
- (b) No additional response burden for the respondents: Similarly, the units of enquiry (persons, businesses, other organizations) are spared the inconvenience and cost of a separate statistical enquiry. The data which they have already provided as part of an administrative process (registration, application, inspection, notification) may be used for statistical compilation without them having to be involved in a separate statistical enquiry on the same or closely related topics.
- (c) It represents a full count of the 'clients' of the administrative system: A complete count is possible since all the records of the system are available for use in the statistical compilation. This means that statistics can be produced for small groups, such as small areas (districts, towns and provinces), without having to be concerned with problems of sampling precision.

However, statistical compilations from administrative systems are not without disadvantages. Possible disadvantages are linked to:

(a) The types of units described in the records: The units used in some administrative systems may not be the most appropriate to satisfy statistical user needs (for example, jobs versus persons, establishments versus enterprises) or may not use a definition of the unit which is compatible with other statistical sources. Some systems register persons or organizations, while others register events (which may occur several times for each individual in a given period). Different users may prefer one or the other, but may not be able to extract this information from the system.

(b) The scope of the registrations/applications: Some administrative systems may have too narrow a scope in that certain categories of units are excluded by design (legal exemptions) or otherwise (illegal non-registration, avoidance), while others may have too broad a scope in that it may include groups which are not of direct interest to a user. For example, registered job-seekers at employment exchanges exclude those unemployed who have not registered and may include those employed who are seeking a change of jobs or additional jobs.

(c) The content of the data collected: The forms used in an administrative system may not include all the information of interest for statistical users or for statistical processing. The data content may be constrained by legislation, limited resources, or other reasons. Coding of important data may be constrained because the distinctions needed for administrative use are fewer than what users of the statistical descriptions of the units/events will need. Thus the descriptions used as a basis for coding may be incomplete and the system may not use national standard coding classifications. Furthermore, data may be correct at first registration, but not be subsequently updated (see also next point).

(d) The procedures for handling data: The procedures used in the administrative system are designed to serve the administrative objectives, rules and regulations, and not to provide a basis for valid, reliable and timely statistics. Administrative procedures may not require the removal from a data base of expired records, or the updating of job details after a person is first registered. The operators of the system are more likely to edit and correct those data which affect their decision making, administrative action or output, and not to give much attention to other data which do not affect their work but which are important for statistical analysis. Administrative procedures and the flow of forms through a system may also lead to delays in updating a data base.

(e) High processing costs: Since administrative procedures are intended to achieve administrative (rather than statistical) output, considerably more attention may need to be given to detecting and correcting errors, and to coding of information which was not needed for the administrative system but is needed for statistical analysis. This processing may require expensive follow up and file amendment.