



OECD SHORT-TERM ECONOMIC STATISTICS EXPERT GROUP (STESSEG)

Task Force on Administrative Data:
Progress Report

Paper prepared by David Brackfield
Statistics Directorate, OECD

Submitted to the Expert Group under item 4 of the draft agenda

Meeting:
27 – 28 June 2005

Room 3
OECD Headquarters, Paris
Starting at 9:30 a.m. on the first day

Administrative Data Task Force – Progress Report

Introduction

1. During the 2004 STESEG meeting a questionnaire was distributed to delegates seeking their views on possible future projects for the Group. This was required as some of the current projects were finishing in either 2004 or 2005. On the final day of the meeting there was a session devoted to discussion on possible future topics. This followed presentations from Denmark and New Zealand on administrative data. The results from the questionnaire, and interest shown in the two presentations on administrative data, clearly indicated that most members would like to see future work on this topic. Further support for the task force was received from NSOs in response to a letter sent out in July 2004 concerning the future work of the Short-term Economic Statistics Experts Group.

2. In October 2004 a 'Draft Proposal for a STESEG Administrative Data Task Force' was disseminated to possible future task force members for comments. In total, comments were received from 5 countries; all who expressed interest and gave positive future directions for the task force based on their own experiences and areas of strength. As a result of this feedback, a Task Force comprising representatives from these countries was formed.

3. In late 2004 the Task Force agreed upon a Scope Paper that set the proposed future direction for the project and laid out an agreed timeframe. The main objective of the Task Force is to provide NSOs who are using or planning to use administrative data in the compilation of short-term economic statistics with a readily available source of help and this will be achieved through the production of a handbook. The initial undertaking was for task force members to check within their own NSOs for previously published papers or articles and to forward these to the chair. Since the inception of the Task Force over twenty such papers have been forwarded (a list of which is provided in Annex 1) from task force members with more promised. The initial outline of the draft handbook is planned to be finalised by 31 July 2005 after input from the STESEG 2005 meeting and a specific meeting of the task force at the same time to begin the handbook process.

4. After reviewing this short paper, we ask that STESEG representatives provide written comments on the following points to aid the discussion at STESEG and the Task Force meeting to be held on 27 June.

- Can the handbook meet the needs of both centralised and decentralised NSOs in different stages of development in regards the use of administrative data or should it focus on a specific group, e.g. NSOs that are starting to use administrative data for the first time?
- Should the handbook try and deal with specific complex statistical methods employed to deal with administrative data especially in regards to the sheer size and sampling techniques involved? Alternatively the handbook could just focus on the general models and underlying assumptions for dealing with administrative data and only provide references to detailed technical papers which explain complex statistical methods that have been applied in some countries.
- Should the proposed handbook look at expanding on the work done so far by Eurostat and Finland (see paragraph 9)? This could involve looking at forming a partnership with Eurostat or using the Eurostat work to date as a starting point?

- Should the proposed handbook place more emphasis on presenting case studies or focus more on the issues presented in the scope document such as; quality, timeliness, and coverage – in a general overall sense?
- Should the proposed handbook be an actual publication or be presented in a more dynamic format such as the timeliness framework (<http://www.oecd.org/std/research/timeliness>). This later approach would gather together papers and documents, logically structured and presented, and place these in a usable and easily accessible format together with the appropriate content information on the internet?
- Additional papers; STESEG members are encouraged to provide administrative data papers of interest to the task force to assist them with the production of the handbook.

Scope of proposed Handbook

5. The following box shows in more detail the scope for the handbook that was agreed upon by the task force members.

The use of administrative data for short-term economic statistics has three key challenges:

- timeliness – administrative data is usually available too slowly for direct use in the publication of short-term economic statistics;
- quality – depending on who is collecting the data and the reasons, there can be issues with the quality of administrative data, i.e. questions of definition, classification, lack of data validation, etc;
- coverage – administrative data in most cases isn't designed or collected as an input into short-term economic statistics, e.g. problems of statistical units, variable definitions, etc.

The use of administrative data is seen to provide three key advantages for short-term economic statistics:

- reduction of response burden – the use of administrative data can significantly reduce response burden, particularly for small and medium sized businesses;
- census coverage – administrative data can provide census coverage for key short-term economic statistic variables;
- cost – in some cases there can be substantial cost savings from using administrative data, for possibly both statistical agencies and respondents.

In recent years a number of NSOs have started to seriously investigate the use of administrative data in the compilation of short-term economic statistics. These investigations have lead NSOs to also ask how to ensure that their organisation is getting the full benefit for short-term economic statistics from the administrative data they currently use and from administrative data that is available but not currently used. These countries now have:

- Good investigative reports that can provide useful examples of how they overcame the challenges involving the use of administrative data outlined previously, and in particular provide information on how the use of administrative data was first implemented, in for example; sample design,

imputation, substitution of surveyed businesses, etc.

- Reports and papers on how their NSO furthered the usefulness (timeliness, coverage, accuracy etc.) of their administrative data through statistical techniques, statistical processes, and improvements in informational technology – both hardware and software.

The desired outcome of the task force will be an administrative data handbook for compilers of short-term economic statistics. The main focus of the handbook will be on how to maximise the benefits that the use of administrative data can provide. In particular, it will address this issue with respect to both short-term economic statistics (i.e. not structural statistics) and the key issues surrounding the use of administrative data – that is to overcome the problems of quality, timeliness and coverage of statistical variables.

It is envisaged that the handbook would have the following basic layout:

- A brief introduction/section aimed at those countries beginning to look at using administrative data for the first time. Countries that have recently incorporated administrative data into the compilation of their short-term economic statistics will be asked to provide relevant papers and reports that will be either included fully or referenced via links. There is also some literature in regards to this subject relating to structural statistics that could either be referenced directly or alternatively the basic thrust of these papers could be modified to be applicable to short-term economic statistics.
- The creation of a table showing the current uses of administrative data in OECD countries for short-term economic statistics. Where possible the table will provide a brief description of how the administrative data is used and reference general links to papers on the statistical processes involved in compiling data for short-term economic statistics.
- The creation of a frame-work that discusses the main issues of administrative data – coverage, timeliness, and accuracy. The frame-work would hopefully be completed by countries that have been using administrative data for a reasonable period of time and focus intensely on the statistical compilation of the statistics. Topics that should also be covered that are sure to be encountered by NSOs who want to maximise the use of their administrative data are:
 - Dealing with the sheer size of administrative datasets, the IT resources required and the computational methods. In particular, this covers capture/processing, validation tools for editing and checking, and the further integration of administrative data into an NSOs statistical compilation.
 - Statistical methods, actual statistical techniques that are used on administrative data including the theory and mathematics. This will most likely be focused on sampling methods, imputation and estimation methods, and data optimisation.
- Administrative issues when receiving administrative data from entities, be they government or private. While in some cases NSOs have the legal right to request administrative data, it is still necessary to form working relationships with the providers to ensure their full cooperation. It would be expected that in most cases a formal contract is entered into between the NSO and the provider. This part of the handbook would deal with all the issues involved in maintaining a good working relationship between the providers of the administrative data and the NSO.

Review of Papers Received to Date

6. There have been several points that have arisen from the papers received so far that the task force will be forced to deal with and make decisions on. The first point is the differences between decentralised national

statistical systems where each part of the system may have very close links and working relationships with administrative data providers or actually be the administrative data provider. This contrasts with centralised national statistical systems that have contracts or service agreements with the administrative data providers, and these differences lead to different attitudes to administrative data. Another theme that arises is the length of time that NSOs have been using administrative data. Some countries, such as the United States, have been using this type of data to produce economic statistics for almost 100 years. For example, the Statistics of Income Division of the IRS began to publish income statistics from their administrative data in 1913.

7. For countries where there is heavy reliance on administrative data and where this data has been used for a long time it has become the norm. These countries have dealt with the issues raised earlier such as timeliness, quality and coverage a long time ago, and to a certain degree the short-term economic statistics produced are completely modelled around what administrative data is available instead of the administrative data being used as an input into producing short-term economic statistics. What this can lead to is very sophisticated and complex statistical models to deal with the administrative data to ensure that good quality short-term economic statistics are produced, as the administrative data may be the only source for the statistics required.

8. There is a possibility that these complex statistical models could be a focus of the handbook, thus providing NSOs using or planning to use administrative data with detailed methodological information. A large part of these statistical methods focus on issues such as benchmarking and modelling and involve high level statistical theories and methods. In recent years a lot of work has been done in this field, including the recent Eurostat “Workshop on Frontiers in Benchmarking Techniques and their Application to Official Statistics” held in April 2005 that attracted over 150 participants¹. Whether the task force wants to pursue this line will need to be discussed and decided after all issues and various opinions have been taken into consideration.

9. Three papers provided that will certainly need to be looked into to see how they can contribute to the project directly are:

- Eurostat – Use of Administrative Sources for Business Statistics Purposes: Handbook of Good Practices (http://forum.europa.eu.int/irc/dsis/bmethods/info/data/new/adso/handbook_en/adso_hbk.htm);
- Eurostat – Business Register Recommendations Manual, Chapter 20 The use of Administrative Sources (Sixth Draft – December 2002); and,
- Statistics Finland – Use of Registers and Administrative Data Sources for Statistical Purposes Handbook: Best Practices of Statistics Finland.

10 The first Eurostat paper gives a good overview of the strengths, weaknesses, opportunities and threats facing the EU member states using administrative data. The paper also provides seventeen² EU country profiles in regards to their use of administrative data, these country profiles focus more on the history of the development of administrative data use, for example, including when the necessary legislation was enacted and how this impacted upon the agency. The country profiles also cover the issue of how statistics are collected in the country and the type of NSO (centralised or decentralised) and how this affects their statistical compilation efforts.

¹ See http://epp.eurostat.cec.eu.int/portal/page?_pageid=1853,1,1853_20488614&_dad=portal&_schema=PORTAL

² Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, United Kingdom.

11. There could be a case made for the task force to try and extend this paper to include OECD member countries and in the process update and expand the Eurostat paper. While the paper isn't dated, it does reference a 1997 seminar – "Seminar on Use of Administrative Sources for Statistical Purposes, Luxembourg, 15 – 16 January 1997". Is it time to organise another seminar to update this work also? The second Eurostat paper is written in the context of Business Registers, but still provides some good information and covers a number of relevant topics.

12. In 2004 Statistics Finland released its 63 page Best Practices Handbook. The handbook mostly focuses on: administrative registers and their uses; employments statistics compiled with the aid of administrative data; and some business statistics. The handbook covers some important topics such as cooperation with suppliers of the administrative data, statistical legislation, opportunities etc. The handbook is very comprehensive in covering Statistics Finland's use of administrative data and could provide direction to or be used as a guide for the task force's proposed handbook.

13. A number of the remaining papers deal with how administrative data has been used with regard to generating a specific short-term economic statistic. In some cases there was no statistic before the administrative data was used, but in most cases the administrative data has been used to improve the quality, timeliness or coverage of the statistic concerned, and the paper is written as a case study. For some of these statistics the changes in the use of administrative data have been performed to meet the requirements of the EU short-term statistics (STS) regulation.

The Task Force and STESEG

14. The Task Force has six members: Finland, Germany, Italy, New Zealand, United Kingdom, and the United States; with the OECD acting as the chair at present. Its inaugural meeting is proposed for Monday 27 June, during which the above issues will be discussed and a course of action decided upon. The papers received to date have raised a number of points that the task force will be required to deal with. Together with comments received from STESEG representatives on this paper, the expected outcome of the meeting is to agree on the desired contents of the proposed handbook.

15. It is therefore requested that STESEG meeting in June focus on the following points. We ask that STESEG representatives provide written comments on these points to aid the discussion at STESEG and for the Task Force meeting to be held on 27 June.

- Can the handbook meet the needs of both centralised and decentralised NSOs in different stages of development in regards the use of administrative data or should it focus on a specific group, e.g. NSOs that are starting to use administrative data for the first time?
- Should the handbook try and deal with specific complex statistical methods employed to deal with administrative data especially in regards to the sheer size and sampling techniques involved? Alternatively the handbook could just focus on the general models and underlying assumptions for dealing with administrative data and only provide references to detailed technical papers which explain complex statistical methods that have been applied in some countries.
- Should the proposed handbook look at expanding on the work done so far by Eurostat and Finland (see paragraph 9)? This could involve looking at forming a partnership with Eurostat or using the Eurostat work to date as a starting point?
- Should the proposed handbook place more emphasis on presenting case studies or focus more on the issues presented in the scope document such as; quality, timeliness, and coverage – in a general overall sense?

- Should the proposed handbook be an actual publication or be presented in a more dynamic format such as the timeliness framework (<http://www.oecd.org/std/research/timeliness>). This later approach would gather together papers and documents, logically structured and presented, and place these in a usable and easily accessible format together with the appropriate content information on the internet?
- Additional papers; STESEG members are encouraged to provide administrative data papers of interest to the task force to assist them with the production of the handbook.

Short-Term Economic Statistics Division
Statistics Directorate
OECD

29 April 2005

Appendix 1

| Providing Country | Paper |
|-------------------|---|
| Finland | Turnover in Other Services and Production in Construction – Improvement of Timeliness |
| Finland | Use of Registers and Administrative Data sources for Statistical Purposes – Best Practices of Statistics Finland |
| Germany | Register-based implementation of the EU regulation concerning short-term statistics in the service sector (NACE sections I and K) |
| Germany | Impact of the Use of Administrative Data and New Surveys on Short-term Business Statistics in the Services Sector in Germany |
| Italy | Use of Administrative Data for Short Term Statistics on Employment, Wages and Labour Cost |
| United Kingdom | Eurostat – Business Register Recommendations Manuel – Chapter 20 The Use of Administrative Sources |
| United Kingdom | A Framework for Assessing and Reporting the Quality of Information provided by the Neighbourhood Statistics Service |
| United Kingdom | Administrative Sources Strategy |
| United Kingdom | Eurostat – Use of Administrative Sources for Business Statistics Purposes – Handbook of Good Practices |
| USA | Statistical Operations and Studies in the SOI Program of the IRS |
| USA | The assessment of Survey of Income and Program Participation (SIPP) benefit data using longitudinal administrative records |
| USA | Monthly Output Index for the US Transport Sector |
| USA | Vital and Health Statistics – Design and Operation of the National Hospital Discharge Survey: 1998 Redesign |
| USA | The Changing Impact of Social Security on Retirement Income in the United States |
| USA | An Assessment of the Current Population Survey and the Survey of Income and Program Participation Using Social Security Administrative Data |
| USA | Exploring Social Security Payment History Matched with the Survey of Income and Program Participation |
| USA | Measuring labor dynamics: the next generation in labor market information |
| USA | Business employment dynamics: new data on gross job gains and losses |
| USA | Developing Business Demographic measures of Industry and Size – 18 th International Roundtable on Business Survey Frames |