



STATISTICS DIRECTORATE

STD/STESEG(2004)12
For Official Use

OECD Short-Term Economic Statistics Expert Group

Administrative data used for Short Term Business Statistics at Statistics Denmark

28-30 June
Room C
Château de la Muette, starting at 9:30 a.m. on the first day

Prepared by Bente Dyrberg
Statistics Denmark

Submitted to the Expert Group under item 5 of the draft agenda

Contact: Bente Dyrberg (bdy@dst.dk)

JT00165560

ADMINISTRATIVE DATA USED FOR SHORT TERM BUSINESS STATISTICS

0. Introduction

Statistical institutes have a large interest in replacing questionnaire-based surveys with surveys based on administrative data, wherever it is possible.

This is sensible in the light of the huge potential advantage derived from the use of administrative registers. The two main advantages are: administrative registers are cost-saving for statistical institutes and the response burden for the statistics in question is at the same time removed or almost diminished.

However, the question that needs to be discussed is: where is it possible to use administrative sources and what basic conditions have to be fulfilled to replace a questionnaire-based survey with statistics based on administrative sources? Different aspects of quality are general for all statistics and are essential to consider. For short-term statistics the question of timeliness is of special importance.

In this paper, quality issues will be discussed on the basis of the experiences of Statistics Denmark, where there is a long tradition of using administrative data in the statistical production process. There is special focus on the use of administrative data for short-term statistics.

The paper is divided into the following paragraphs:

1. The Danish background
2. The use of administrative registers
3. Quality aspects in relation to the use of administrative data for short-term statistics
4. Achieving data quality when using administrative registers
5. Two examples of the usage of administrative registers - a good and a not so good case
6. Conclusions

1. The Danish background

Statistics Denmark started early to utilise administrative registers.

Two main factors affected this development:

1. Central public registers with unique identifiers were established in the 1960's containing information about people and enterprises respectively.
2. The Act on Statistics Denmark from 1966 included a paragraph saying:
Statistics Denmark...

...will supervise or assist in the establishment and utilisation of central public registers which serve to perform administrative duties for the public sector, business and industry and which can be used for statistical purposes

For example, the Population and Dwelling Census has been based on administrative registers since 1981.

For a long time Statistics Denmark has also used administrative sources for short-term statistics and still continues to increase its usage of registers.

Several important short-term indicators are based on administrative sources, e.g.:

Turnover for most sectors (based on VAT)

Employment (based on The Labour Market Supplementary Pension Scheme)

Construction production and construction building permits (based on The Register of Building and Dwelling Statistics)

Most short-term indicators are, however based on questionnaire surveys. The reason is that there is no administrative data covering these indicators or that the administrative data is not timely enough.

Since 1995, Statistics Denmark has had special focus on the quality dimension of timeliness, and there has been a substantial reduction in release times both for questionnaire-based indicators and indicators based on administrative registers.

2. The use of administrative registers

There are several advantages of using administrative data;

- for the National Statistical Institutes (NSI) it is cheaper to reuse data already collected for other purposes outside the NSI's
- the response burden is reduced
- the data can be used in several different contexts

It is efficient and cheaper to produce statistics when National Statistical Institutes (NSI) use data already collected for other purposes outside the NSI. However, the effort to get the data into a usable form for statistics should not be underestimated. For some registers editing of data is needed to a great extent and consequently frequent contact to the responsible authority is also necessary.

Also the use of administrative registers lowers the response burden, which is an extremely important aspect for politicians and subsequently also for the NSI's today. In Denmark, there has been growing pressure on reducing the response burden for the last 8-10 years, which has increased consciousness to use administrative sources, where it is possible.

Administrative registers are used for different purposes:

Direct - as the only source for a specific published short term statistics e.g.

Monthly Turnover

Quarterly Employment Statistics

Quarterly Construction Activity

or *indirect*- used for example when

Editing data collected from questionnaires

Grossing up data from survey samples to cover the whole population

Reducing samples for questionnaire-based indicators

3. Quality aspects in relation to the use of administrative data for short-term statistics

The advantages of using administrative registers are clear.

It is especially worth noticing the ideal situation where one register can be used as the only source for one particular short-term statistic. In that situation, you achieve *total* coverage of, for example, turnover and employment. That gives you a solid foundation and wide range of possibilities for producing short-term statistics. For example, ideally seen it means that:

- It will be possible to produce good preliminary statistics of a high quality because you easily can impute data which is still lacking. The imputation can be based on individual data for the lacking units.
- You have a solid frame of reference for all short-term statistics e.g. also for questionnaire-based surveys.
- You can form tables as you want, because you have information on unit level. So it is possible to produce/sell individual service tasks.
- You can combine data with information from the Business Register because of the unit level information

Unfortunately, administrative registers are often not ideal for statistical purposes. The administrative registers have been established to cover certain non-statistical purposes, and therefore very seldom completely fulfil the demands for statistics, which are defined very precisely in EU Regulations, international recommendations and so on.

Before using administrative registers, it is therefore necessary to investigate the aspects which affect the quality of the statistics. It concerns central questions like:

What are the definitions of the variables in the register compared to what should be covered by the statistics?

Even if the statistical definition does not completely fit the definition of the administrative concept, collected information can often be used as an indicator anyway (for example: information on sales can be used as an indicator of turnover). In general, it is an advantage that the main focus of short-term statistics is on the development and not on the level. Also, short-term indicators are of a preliminary character and the demands on accuracy are not as high as, for example, on structural statistics.

Which units are used?

Solutions have to be found to form statistical units, if special administrative units are used in the registers. At Statistics Denmark these relations are established in the Business Register.

What about timeliness?

Especially for short-term statistics timeliness is an aspect which is very important. Negotiations with the relevant authorities can, with a bit of luck, result in quicker routines. For example, on-line solutions (instead of magnetic tapes) can speed up the delivery time of data, and consequently create more possibilities of producing preliminary versions of the relevant statistics.

How interested is the responsible authority in keeping an updated register?

If the authority has its own economic interest in collecting timely data from their respondents, there is very great chance, that the register has been updated (for example: the Tax authorities collection of VAT-register), and consequently the data is of good quality. On the other hand if, for example, a municipality has a pure administrative task for the state to update a register, the quality of the updating can be very low and therefore result in poor quality of the register (for example: that is the case for part of the updating of the Building and the Dwelling register)

Other aspects which can affect the quality are

- changes in the administrative rules (for example: change of thresholds)
- summer holidays (which in some cases can mean that there is no delivery of data for that month)
- reporting periods in the administrative register are different from what should be covered by the statistics
- delays in data delivery

To sum up:

If you can use administrative registers also for short-term statistics, there are clear advantages – direct or indirect..

But as a producer of statistics you should be aware of the following aspects :

By using administrative registers the statistical institute is dependent on the administrative procedures. When the administrative systems change a reconstruction of the statistical systems to handle the data is needed or in very seldom - worst cases - the source has to be given up.

In general, a good measurement of the quality of the register is how much interest the responsible authority has in keeping a continuously updated register.

The quality can be so poor that the information cannot be used as the only source for certain statistics.

Especially for short-term statistics, timeliness can be difficult to achieve and other sources have to be used.

4. Achieving data quality when using administrative registers

If administrative sources are chosen as the basis for statistics, work has to be carried out to process the data into a form where it can be used. Before publishing a short-term indicator based on administrative sources, quality checks are needed.

In general, it is very important to make sure that the administrative data looks sensible. Experience shows that when data is delivered almost ready-made for use, there is the risk of being blind to detecting errors.

First, a macro check is needed. The question is: do the overall figures (for example: the number of records or the total value) have the right level if compared to former periods? If something is wrong, the responsible authority should be immediately contacted. Especially when we focus on short-term statistics it is important that no time is wasted on more detailed editing before recognising that something is wrong.

If the macro check is all right, further detailed checks can be implemented. Often the register consists of large amounts of records and the check also has to be performed very quickly. Therefore, it is important that editing procedures are as automatic as possible.

Unlikely developments or unlikely figures for large enterprises which influence the publication cells should be selected for manual handling (micro check). Other very typical errors can be solved automatically.

It is necessary to go through logical and plausibility checks – but only on an overall branch level. For later versions of the indicator more detailed checks could be performed.

Sometimes, systematic errors appear and then the responsible authority should be contacted.

On an overall level, the figures should also look sensible when compared to other related indicators, although it can be difficult to compare data from different sources.

If the qualitative checks mentioned above imply that overall figures or some figures do not look plausible at all, and you cannot find the reason why the data looks atypical, then publication of the data has to be postponed until you either find the error or an explanation. This risk is especially connected to short-term indicators. It could be devastating for a statistic, if this happens too often.

So, a large number of quick quality checks have to be carried out when administrative data is received. But most of all, it is important to use common sense throughout the whole process - do the figures look sensible?

5. Two examples of the usage of administrative registers. A good and a not so good case

The success of using administrative sources will greatly depend on the quality of the incoming data. Some errors can be found in the following working processes in the NSI's, but the updating has to be very quick. In this paragraph, two examples from Denmark are shown, a good one and one that is not quite so good.

The good case:

Indicator	Administrative Register
Monthly/quarterly turnover for most market activities	VAT

Since 1978 this turnover indicator has been published, but up to 2002 it was only on a quarterly basis. But from 2002 it has become a monthly indicator based primarily on the big enterprises which must submit VAT returns monthly to the Danish Customs and Tax Authorities. These big firms make up only 5 per cent

of the total number of firms, but they account for 82 per cent of the total turnover. The sales figures obtained from the VAT returns are very reliable, because the tax authorities have a big interest in collecting these data for tax reasons. The purchase data are influenced by investments, which are included in the purchase, and therefore this variable is not quite so useful.

On-line delivery of data from Customs and Tax to Statistics Denmark has just started, and this has improved the timeliness of the statistics from $t + 2$ months to $t + 1\frac{1}{2}$ months. Moreover, Statistics Denmark is investigating the possibility of issuing a preliminary release which will be published one month after the reference period (i.e. $t + 1$). The preliminary version will be based on the early available big firms and it is expected to include 70 per cent of the total in terms of turnover, the rest will be imputed.

So, this preliminary version could become a very quick indicator in the future, and could also be used as an important benchmark for other short-term statistics. However, comparisons with other statistics are affected by administrative units that are used for this turnover indicator.

The case that is not quite so good :

Indicator	Administrative Register
Quarterly Construction : started and completed buildings (building activity)	The Register of Building and Dwelling Statistics

For the indicators for construction based on administrative data, it has been necessary to change the publication of the statistics several times due to the very poor quality of the administrative register.

Data for a quarter is published about 2 months after the reference period. At that time we only have received 40-50 % of the expected data. This means that a vast amount of data has to be imputed on the basis of knowledge of earlier patterns of reporting from municipalities. But the main problem is that the reporting pattern changes all the time. Consequently, the development is typically underestimated to a different degree from quarter to quarter.

Because of these major quality problems, the statistics for construction activity were changed from a monthly statistic to a quarterly statistic four years ago. And this year we have changed the way we publish the data, so the users do not only obtain information about the latest calculated data, but also get information about the first publication of the same quarter one year ago. All figures are revised in every publication.

The register also is used to the calculation of the production in the construction sector.

The responsibility for the register has changed and is now in fact under the same Ministry as Statistics Denmark. The register has repeatedly come under strong criticism during the past years (also from Statistics Denmark), but nothing has changed yet.

The fundamental problem is the big delay in the reporting from the municipalities because for some parts of the register, namely the part which concerns building activity the task of updating is not prioritised. The municipalities have no special interest in updating these data quickly.

However, it will be very costly for Statistics Denmark not to use the Building and Dwelling Register. The alternative is to start our own collection from the municipalities, but that will be very costly and definitely cause an outcry from the municipality due to the large response burden. So the strategy for

Statistics Denmark is to put more pressure on the authorities to improve the updating and make the updating problem visible.

6. Conclusions

The advantages of using administrative registers are so clear that these sources should be used where possible, also for short- term indicators.

But it is important to be aware of any quality problems and also establish solutions to achieve good quality.

The quality of a register is widely affected by how much interest the responsible authority has in keeping a correct and updated register. But, also differences in definitions, units and so on can create problems. Timeliness can cause special problems for short-term statistics.

However, all in all the Danish experiences give an unambiguous support to the utilisation of administrative sources for producing both structural and short-term statistics.