Frameworks for Access to Administrative Sources

3.1 Introduction

The access to data from administrative sources is one of the key barriers to the wider use of such data for statistical purposes. This chapter describes the various frameworks needed to facilitate access to administrative sources, drawing on examples and experiences from several countries. These frameworks typically have several dimensions: legal, policy, organisational and technical, each of which is considered below. It is necessary to reach agreement in all of these areas before the benefits of the use of administrative data can be realised.

3.2 Legal Frameworks

Legal frameworks are normally constructed at the national level, and are specific to national sources and circumstances. In some cases, however, there may also be relevant legislation at either the sub-national (e.g. state) level, or the international level. An example of the latter is the statistical legislation of the European Union, which is binding on Member States. In such cases, it is possible that there are two or more alternative legal gateways to administrative data.

Most national statistical organisations have legal texts defining their roles and responsibilities, typically in the form of a statistics act. In many countries, these legal texts include specific provisions for the access to administrative data. Examples include the statistics acts of Ireland[1] and Norway[2].

Box 3.1 – Extracts from the Irish Statistics Act of 1993

Section 30. (1) For the purpose of assisting the [statistical] Office in the exercise of its functions under this Act, the Director General may by delivery of a notice request any public authority to –

(a) allow officers of statistics at all reasonable times to have access to, inspect and take copies of or extracts from any records in its charge, and

(b) provide the Office, if any such officer so requires, with copies or extracts from any such record,

and the public authority shall, subject to subsection (2) of this section, comply with any such request free of charge.

Section 31. (1) The Director General may request any public authority to consult and co-operate with him for the purpose of assessing the potential of the records of the authority as a source of statistical information and, where appropriate and practicable, developing its recording methods and systems for statistical purposes, and the public authority shall comply with any such request, in so far as resources permit.

(2) If any public authority proposes to introduce, revise or extend any system for the storage and retrieval of information or to make a statistical survey it shall consult with the Director General and accept any recommendations that he may reasonably make in relation to the proposal.

Some national legal frameworks give more powers than others for access to administrative data for statistical purposes. This is because national historical, political and cultural factors have a strong influence on these frameworks. Cultural factors can be particularly important, as some cultures are much more favourable than others to the idea of data sharing between government departments and agencies. As a result of these national differences, legal frameworks are not particularly harmonised or even consistent between countries.

To address this issue of consistency, the European Union has included provisions on access to administrative data in Regulation 223/2009 on European statistics, commonly known as the “Statistical Law”[3]. This Regulation gives the national statistical organisations of Member States the right of access to the administrative data needed to meet their obligations under European statistical legislation, but states that such access is still subject to national limits and conditions.

Individual European Union regulations in specific areas of statistics go further, and remove this dependency on national limits and conditions. An example of this is the business registers regulation, which gives unconstrained access to any administrative sources, where data from these sources are necessary to meet the requirements of the regulation[4].

As well as giving access to data from administrative sources, legal frameworks also set out the limits to such access, and to the uses of administrative data. Often there are restrictions that data can only be used for specific statistical purposes, and that the confidentiality of individual records should be maintained.

For example, there may be specific restrictions on the use of data for unincorporated businesses, particularly sole-proprietorships, where business data could be considered to be personal data relating to the owner of the business. In such cases, business profit can be seen as equal to personal income. Many countries have data protection legislation covering information about individual citizens, therefore it is important to make clear distinctions between what constitutes business and personal data in such cases.

The legislative process can take time, and statistics may often be seen as a relatively low priority by legislators, so a sustained period of lobbying and highlighting the benefits of using administrative data may be necessary. Given all the efforts that are usually needed to introduce or revise statistical legislation, it is therefore necessary to make the most of the opportunity. In particular, it is essential to avoid the mistake of proposing legislation that just meets current requirements. It may be ten or more years until the next opportunity to revise legislation, so it is necessary to have a long-term strategy for the use of administrative data, and to ensure that the legislative proposals meet all envisaged requirements for the foreseeable future. In this way, legislation can be seen as a short-term barrier, but a long-term opportunity.
Even whilst legislation remains a barrier, it does not necessarily prevent any use of administrative data. In one example, whilst waiting for a suitable legal framework for access to Corporation Tax data, a member of staff from the United Kingdom Office for National Statistics was seconded to the tax agency to explore the feasibility of using these data for statistical purposes. This person had access to the micro-data whilst seconded to, and physically working in the premises of the tax agency, but could only take non-disclosive, aggregate analyses back to the statistical office. This approach meant that various data issues could be addressed, including a proper assessment of the real value of the tax data, whilst simultaneously exploring the possible legal routes to gaining access.

It should also be noted that legislative restrictions often concern the use of micro-data, i.e. information on individual people or businesses. Although statisticians are habitually used to working with data at this level to produce aggregate results, it may sometimes be feasible to work with non-disclosive, low-level aggregates instead. In some cases, this could be done by simply re-defining the statistical unit from the individual to a small group of individuals sharing certain characteristics, perhaps with a weight equal to the number of members of that group.

3.3 Policy Frameworks

Many countries have general policies on data sharing within government, which will influence the right of access to administrative data for statistical purposes. However, it is often easier to change policies than to change laws, and policy tends to evolve over time. It is therefore important that national statistical organisations participate fully in policy development, and take an active part in any discussions within government that might lead to policy changes. In this way, any changes should be formulated in a way that gives the maximum possible benefit to the statistical system.

Policy frameworks also encompass voluntary codes of practice, the most important of which, for statistical purposes, is the United Nations “Fundamental Principles of Official Statistics”[5]. Principle 5 concerns cost-effectiveness, and suggests the use of data from administrative sources in this context:

“Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.”

An explanatory note to principle 5 also stresses cost-effectiveness, and goes on to say that:

“Statistical offices must be cost-effective, making the best choice of concepts, sources and methods by balancing quality, timeliness, costs and the reporting load of respondents … The overall cost-effectiveness of an agency is influenced by organizational planning and operation, the sound application of statistical methodology, exploitation of information and communication technology and also access to administrative records.”

The code of practice for the European Statistical System[6] contains similar provisions, but the use of data from administrative sources is encouraged in slightly different contexts. Principle 2 concerns the mandate for data collection, and states that:

“Statistical authorities must have a clear legal mandate to collect information for European statistical purposes. Administrations, enterprises and households, and the public at large may be compelled by law to allow access to or deliver data for European statistical purposes at the request of statistical authorities.”

Principle 9 is concerned with ensuring that the burden on respondents to statistical surveys is not excessive. It states that:

“The reporting burden should be proportionate to the needs of the users and should not be excessive for respondents. The statistical authority monitors the response burden and sets targets for its reduction over time.”

One of the proposed indicators to measure the application of this principle is:

“Administrative sources are used whenever possible to avoid duplicating requests for information.”

Codes of practice may also exist at the national level, and are often valuable as a way of reassuring the public that data will only be used for specific and reasonable purposes. To have any real value, it is important that these codes of practice are made available to the general public, typically via the internet site of the national statistical organisation.

3.4 Organisational Frameworks

Once the legal and policy frameworks are in place to permit the use of administrative data, it is necessary to consider the organisational arrangements to facilitate data flows. Typically this takes the form of a written agreement. This may be a contract, particularly if a private sector organisation is involved, but, if the agreement is between government departments or agencies, it is more likely to be a "service level agreement", “protocol” or “concordat”. The difference is that contracts tend to be legally binding, whereas other forms of agreement are not.

There are certain key features that should be present in any such agreement. These are as follows:

- The legal basis: A reference to the legislation permitting the access to the administrative source for statistical purposes, and to any legislation that imposes restrictions on this access.
- Names of persons transferring / receiving data: The names, contact details and job titles of the key people involved in the transfer of data in both the administrative and statistical organisations should be recorded. In some cases, this can include all persons within the statistical organisation who are entitled to use or view the data.
- Detailed description of data covered: This will include information identifying the data set and the variables contained within it.
- Frequency of data supply: This will specify when and how frequently the administrative organisation will supply data.
- Quality standards: These set the parameters for the quality of the data supplied. Examples might include requirements for addresses to meet certain standards, or for a maximum proportion of missing or erroneous variables, to ensure that the data received are fit for purpose. The priorities assigned to different variables, and hence the effort put into quality assurance, will often differ between administrative and statistical organisations, therefore agreeing common standards can be difficult.
- Confidentiality rules: It is important to set out what the data can be used for, what rules and procedures will be in place to prevent disclosure, and in what circumstances the data can be passed on to customers of the statistical organisation.
- Technical standards: These are covered in more detail in the technical frameworks section below.
• Provision of metadata: It is important that data flows are accompanied by the relevant metadata, which can include dates, descriptions for any codes used, information on the units used, etc.
• Provisions for payment for supply of data: Data transfers between government departments or agencies are generally free of charge, though in some cases, the statistical organisation may be required to contribute towards the costs of extracting and transferring the data. Data from private sector organisations are often charged for at market rates, though it may be possible to negotiate discounts, particularly if there are several users of a private sector data source within government. In some cases, it might be possible to offer statistical analyses or expertise as a form of payment for data received.
• Period of agreement: Agreements may be for a fixed period, but if so, they should include provisions to renew or extend them as necessary. An alternative approach is to have an agreement that is considered valid until one of the parties wants to make a change.
• Contingencies for changes in circumstances: It is important for the statistical organisation to have advance warning of changes affecting the administrative source. The agreement should specify that any proposed changes are communicated to the statistical organisation as soon as possible, to allow the impact of the changes on statistical outputs to be minimised.
• Procedure for resolving disputes: The agreement should specify the method to resolve any disputes between the statistical and administrative organisations, which will normally be to escalate issues to senior managers, or possibly even to the relevant ministers.

3.5 Technical Frameworks

Technical frameworks refer to the mechanisms by which data are transferred, as well as any relevant data or metadata standards. Data transfer mechanisms can take any form from paper records sent by post to real-time updates via a secure electronic link. The mechanism used has to take into account the technical possibilities open to both the sending and the receiving organisation, so is often a compromise reflecting a sub-optimal solution for at least one of these organisations.

There are a number of international standards for data and metadata transmission, including XML, SDMX and DDI, to name but a few. Some countries also have national versions, particularly for data transfers within government. It is therefore important to agree which standards are to be used.

3.6 Summary

It is essential to have a legal framework in place to permit the use of administrative data for statistical purposes. The other frameworks described above are not essential, but are very useful for assuring a smooth flow of data, and minimising any problems or misunderstandings between the data supplier and the statistical organisation. For this reason, it is helpful if they are reflected in written documents, agreed by all parties.

Comparing practices between countries can be useful for benchmarking purposes, but it should be remembered that specific national situations and issues often require specific solutions. International standards can help in terms of providing guidance (and compliance with them may be seen as a political goal) so they should be quoted wherever possible in discussions with administrative departments.

### Box 3.2 – Case Study: Frameworks in the United Kingdom

#### Legal frameworks

The Statistics and Registration Services Act 2007 provides the framework for access to administrative data, but does not give a blanket right of access as in many other countries. The conditions of access and use of administrative data are often governed by source-specific legislation such as the Value-added Tax Act of 1994. Access to new administrative sources is subject to parliamentary approval. As a member of the European Union, the United Kingdom is also subject to the provisions of the European legislation relating to the use of administrative sources.

#### Policy Frameworks

In addition to applying the United Nations Fundamental Principles of Official Statistics and the European Statistical System Code of Practice, there is a national code of practice (http://www.statistics.gov.uk/about/national_statistics/cop/default.asp) for members of the Government Statistical Service. The key provisions in relation to the use of administrative data for statistical purposes are:

- “5(f) The same confidentiality standards will apply to data derived from administrative sources as apply to those collected specifically for statistical purposes.”
- “7(c) The value of administrative data in producing National Statistics will be recognised, and statistical purposes should be promoted in the design of administrative systems.”
- “7(d) Statistical systems will be designed in ways that maximise the potential to add value through data integration.”

The code of practice is supported by various protocols, including a Protocol on Managing Respondent Load (http://www.statistics.gov.uk/about/national_statistics/cop/downloads/respondentload.pdf), which contains the following statements:

- “2. New statistical surveys will not duplicate existing sources... Producers of National Statistics will consider using existing survey data, administrative data and other non-survey sources before introducing a new survey... A survey will be conducted only where there is no suitable alternative data source.”
- “4. The value of administrative data in producing National Statistics will be recognised, and statistical purposes should be promoted in the design of administrative systems. National Statistics will, where appropriate, be derived from information supplied for the administration of government business and public services. This will be achieved, wherever possible, by direct extraction of relevant data from the systems supporting the administration. Producers of National Statistics will seek to influence those responsible for the design of administrative systems so that these systems can also capture data for statistical purposes in an economical way.”

#### Organisational Frameworks

The organisational frameworks for the transfer of data between government departments and agencies tend to be incorporated in “service level agreements”. These are signed at a senior level, but are not legally binding. They contain general provisions in the main part of the agreement, and
have details of specific data requirements and specifications in annexes. Usually there is no payment, but in some cases, statistical analyses or tools are provided in return.

Government departments and agencies that supply administrative data for the statistical business register are represented on the management committee for that register, which also includes users. This helps them to better understand how their data are used, and the implications of data quality.

The company registration agency (Companies House) operates on more of a commercial basis, so the framework for the transfer of data from that agency takes the form of a contract, with a payment. Data on business ownership and control links are also purchased from a private sector business data supplier.

- **Technical Frameworks**

Most data transfers are via text files, with either fixed-length fields or standard delimiters. One area where some standardisation has been possible is in the format of business addresses. This has been facilitated by the availability of address referencing software tools based on Post Office standards.

Most data transfers are currently via discs sent by post, or, for smaller data-sets, via secure e-mail links. However, for Value Added Tax data used in the statistical business register, a system of daily updates has been set up using transaction files sent via the government secure intranet.

Metadata are usually transmitted as reference tables, either accompanying the data, or separately, on a less frequent basis. Metadata defining codes are stored as look-up tables, whereas more general metadata are recorded in a database of standards and guidance.

[2](http://www.ssb.no/english/about_ssb/statlaw/statlov_en.html) (See chapter 3-2)
[5](http://unstats.un.org/unsd/methods/statorg/FP-English.htm)