

Frameworks for the Access and Use of Administrative Data, With the Example of Current Practice in the United Kingdom

**Steven Vale, Statistical Framework Division,
Office for National Statistics, UK**

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 paper describes the various legal, policy, organisational and technical frameworks needed to facilitate access to administrative data for the purposes of compiling economic statistics.

Section 2 of this paper considers the general principles for such frameworks based on experiences from a range of countries, whilst recognising that the best solution for one country might not be suitable for another. Section 3 gives more specific information on the current frameworks in the UK, and how they might change in the future.

2. General Principles

A detailed set of frameworks is necessary to facilitate access to administrative data for statistical purposes. 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.

2.1 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 the international level. An example of this 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 institutions have legal texts defining their roles and responsibilities, typically in the form of a statistics act. In many countries, these legal texts have been revised in recent years, and now include specific provisions for the access to administrative data. Examples include the statistics acts of Ireland¹, Norway² and Germany³.

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, and thus improve the comparability of statistics derived wholly or partly from administrative sources, the European Union is putting in place an international legal framework for access to administrative data. The cornerstone of this framework is Regulation 322/97 of 17 February 1997 on Community statistics, commonly known as the "Statistical Law"⁴. This Regulation gives the national statistical offices 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 regulations in specific areas of statistics go further, and remove this dependency on national limits and conditions. An example of this is the new draft 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.

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.

There may also 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. For example, for many small businesses, 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 be clear what is considered to be business data, and what is considered to be personal data in such cases.

¹ www.irishstatutebook.ie/ZZA21Y1993.html (See sections 30 and 31)

² www.ssb.no/english/about_ssb/statlaw/statlov_en.html (See chapter 3-2)

³ www.destatis.de/download/ueber/bstag03_e.pdf (See article 8)

⁴ <http://forum.europa.eu.int/irc/dsis/bmethods/info/data/new/322-97en.htm>

2.2 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 offices 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 recently agreed code of practice for the European Statistical System⁶ 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

⁵ <http://unstats.un.org/unsd/methods/statorg/FP-English.htm>

⁶ http://epp.eurostat.cec.eu.int/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/CODE_OF_PRACTICE-EN.PDF#search=%22ess%20code%20of%20practice%22

“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 institute.

2.3 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 and contact details of the key people involved in the supply of data in both the administrative and statistical organisations should be recorded. In some cases, the names of all persons within the statistical organisation who are entitled to use or view the data may also be included.
- 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 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 will normally be for a fixed period, but should include provisions to renew or extend them if necessary.
- 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.

2.4 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 GESMES, 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. Current Practice in the UK

The current practice regarding the frameworks for access to administrative data are described here for information, and to allow the reader to draw comparisons with practices in his or her own country. It is not the intention of this paper to promote the UK approach as an example of best practice that should be followed by other countries. UK practice has evolved to meet specific national requirements and constraints, which will be different in other countries.

By way of a general introduction to UK practices, it may help to outline the main features of the UK statistical and administrative systems. The statistical system is partly de-centralised. The Office for National Statistics (ONS) is currently an executive agency linked to the Treasury (finance ministry), and is responsible for most economic and social statistics. However, there are government statisticians in many other departments, who, together with ONS staff form the professional grouping known as the "Government Statistical Service".

Statisticians in other government departments often have access to administrative data collected by their department, but mechanisms for data sharing are limited. Most statistics produced by members of the Government Statistical Service are published under the brand name "National Statistics".

Although most economic statistics are produced by the ONS, construction statistics remain the responsibility of the Department for Trade and Industry, and some economic statistics for Northern Ireland, Scotland and Wales are produced by the devolved administrations covering those parts of the UK. However, virtually all UK economic statistics are linked to the statistical business register maintained by the ONS, which guarantees a certain level of consistency.

One particular feature of administrative sources in the UK is the lack of common identification numbers. The UK does not currently have unique identifiers for persons or businesses. Each administrative source for business data has its own numbering system, and the units used in one system are not necessarily the same as those in another. An enterprise (defined in the statistical sense) may correspond to two units in one administrative source, and three in another, with no correlation between those administrative units. This is particularly true for larger multi-activity businesses. Thus matching business records using variables such as name and address is a key requirement before administrative data can be used for the purposes of economic statistics.

Until recently there were two main tax agencies in the UK, each keeping separate records of businesses for the purposes of different taxes. These two agencies have now merged, and there are plans to improve the coherence of different tax systems, but it is still likely to be many years before their business data systems are fully integrated.

Public and business opinion in the UK is generally in favour of limited sharing of business data between government departments and agencies, particularly if this increases efficiency and reduces costs, including those relating to the administrative burden on businesses.

3.1 The Legal Framework in the UK

The UK does not have a single statistical act of the type found in most other countries. Business surveys are mainly conducted under the Statistics of Trade Act, which dates back to 1947, but this Act does not refer to the use of administrative data. The use of administrative data to maintain the statistical business register is governed by source-specific legislation such as the VAT Act of 1994, and the Finance Act of 1969. These acts also permit data from the specific sources they cover to be used for wider statistical purposes.

As a member of the European Union, the UK is also subject to the provisions of the European legislation relating to the use of administrative data described in section 2.1 above. These provisions have not, however, been used to gain access to any particular source so far.

There are currently plans to make the ONS more independent of central government, which will require new enabling legislation. It is not clear yet exactly what will be included in this legislation, and whilst it is unlikely that the legislation will amount to a full statistics act, as in other countries, it is possible that it will include some limited provisions on expanding data access.

3.2 The Policy Framework in the UK

UK government policy is moving towards greater sharing of administrative data between government departments and agencies. A recent review on regulatory inspections and enforcement by Philip Hampton⁷ recommended a number of actions, including:

“... regulators should look to improve cooperation and data sharing to reduce the need for businesses to submit the same data more than once.”

In response to this review, work has recently started on developing a business index, and a prototype system of single business identifiers to facilitate data sharing. Both areas of work are still at the conceptual stage, but have aroused considerable interest across government.

The main drivers for this policy are the desire to make efficiency savings within government, and to reduce the administrative burden on businesses. The ONS is

⁷ http://www.hm-treasury.gov.uk/budget/budget_05/other_documents/bud_bud05_hampton.cfm

playing a full part in all such development activities, in order to secure the maximum benefits for the statistical system.

The ONS supports and applies both the United Nations Fundamental Principles of Official Statistics and the European Statistical System Code of Practice. In addition, there is a national code of practice⁸ for members of the Government Statistical Service. The key provisions of this code of practice in relation to the use of data from administrative sources 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 a number of protocols, the most relevant of which is the Protocol on Managing Respondent Load⁹. This includes the following statements:

- *“2. New statistical surveys will not duplicate existing sources. The costs of compliance will be taken into account in survey design.*
 - *a) Producers of National Statistics will consider using existing survey data, administrative data and other non-survey sources before introducing a new survey. They will examine existing data and document the consideration given to other potential sources. 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.*
 - *a) 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.*
 - *b) 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.”*

It is also ONS policy that the statistical business register should form the gateway into the statistical system for all administrative micro-data concerning businesses. In this

⁸ http://www.statistics.gov.uk/about/national_statistics/cop/default.asp

⁹ http://www.statistics.gov.uk/about/national_statistics/cop/downloads/respondentload.pdf

way administrative units and variables can be converted into their statistical equivalents, which facilitates the integration of administrative and statistical data. This does not mean that the register should store all of the variables from administrative sources, just those needed to allow the necessary linkages and transformations.

3.3 Organisational Frameworks in the UK

The organisational frameworks for the transfer of business data between government departments and agencies in the UK tend to be incorporated in “service level agreements”. These are signed at a senior level, but are not legally binding. The contents of service level agreements in the UK comply with the recommendations set out in section 2.3 above. They contain general provisions in the main part of the agreement, and have specific data requirements and specifications in annexes. No payments are involved, but in some cases, statistical analyses or tools are provided in return.

Government departments and agencies that supply administrative data for the ONS business register are represented on the management committee for that register. This committee meets every three months, and includes users of the register. It receives an annual report on the service level agreements regarding transfers to and from the register.

The company registration agency in the UK (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, Dun and Bradstreet, under a multi-year contract.

3.4 Technical Frameworks in the UK

The ONS currently has to take a fairly flexible approach to data and metadata formats and standards, as these are constrained by what is feasible for both the supplying organisation and the ONS itself. Most data transfers tend to be via flat format 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 quarterly copies of administrative databases, sent on CD or DVD-ROMs, by post. However, for Value Added Tax data, it has been possible to set up a system of daily updates using transaction files sent via the government secure intranet. The intention is to migrate the other data flows to this model wherever possible.

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

4. Summary

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 also help in terms of providing guidance, and compliance with them is often seen as a political goal, so they should be quoted wherever possible in discussions with administrative departments.

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.