

Use of Administrative Data to Maintain the Statistics New Zealand Business Frame and the Economic Statistics Programmes

1. Executive summary

Statistics New Zealand has a comprehensive and robust Business Frame (BF) that underpins the quality and integrity of New Zealand's business statistics and national accounts programmes. Tax data is the primary administrative data source used to maintain the BF.

Tax data related to individual businesses, together with a small range of other administrative data sources, provide the breadth of coverage to support the core functions of a business register (or Frame) which comprehensively and accurately represents the population of economic units in the economy. The BF is used as a frame for business surveys, as a source of standard classifications and as an integration framework between data sources. The BF, which maintains a link to the unique tax system number for each enterprise, has been a key facilitator in the integration and use of tax data within the official statistics system.

The use of individual business tax and other administrative data within the official economic statistical programme has major benefits in terms of the range and depth of statistical outputs that can be efficiently produced without imposing additional compliance costs on providers. Significant benefits are also gained by utilising existing information collected by these government organisations which have infrastructures of systems, people and processes that the statistics office could never realistically replicate.

The use of administrative data by Statistics New Zealand to maintain the BF and as a source of data for economic statistical outputs is supported by Central Government initiatives to reduce compliance costs on businesses and by the organisations responsible for the administrative data. Statistics New Zealand, working with the tax administration and others, has however had to address the challenges of transforming data collected for administrative purposes into data that fits with statistical model requirements.

Consideration of the legal, confidentiality and privacy issues associated with administrative data are also very important for Statistics New Zealand. Not only must Statistics New Zealand meet the requirements of the Statistics Act 1975, it must ensure that the security and other requirements of legislation such as the Tax Administration Act relating to access and use of individual administrative data are also met.

Fundamental to the success of the BF strategy in New Zealand is provision for the supply of individual business data to Statistics New Zealand in other legislation such as the Tax Administration Act. In contrast, the Statistics Act provisions regarding the security and confidentiality of individual business data used in the production of official statistics preclude Statistics New Zealand from providing any other government department or other organisation, including the Courts, access to individual data about businesses or individuals.

2. Business Frame

The Statistics New Zealand's Business Frame (BF) is a list of private and public sector businesses and organisations ('enterprises') engaged in the production of goods and services in New Zealand. The BF is comprehensive and is continuously maintained to reflect real world changes (eg business start ups, businesses closures etc). The coverage of the BF is estimated to be 99 percent of the economic activity of the business sector covered by the tax system. There are approximately 420,000 live enterprises recorded on the BF. The BF is used by Statistics New Zealand as a:

- Frame for the business surveys operated by Statistics New Zealand. Using the one centralised frame for all business surveys ensures that:
 - the same standard classifications are used in all outputs

- data from the different surveys can be integrated
 - efficiency gains are achieved from having one frame, not many with duplicated maintenance activity
 - quality gains for all surveys in sharing survey feedback via the central frame are achieved
 - one picture of businesses in terms of managing overlaps in Statistics New Zealand survey designs and monitoring compliance costs resulting from the surveys.
- Framework for the integration of administrative data with Statistics New Zealand data. Each enterprise on the BF is linked to a tax unit number. All enterprises that are limited companies are linked to a Companies Office number. The Companies Office is the government organisation in New Zealand responsible for the registration of limited liability companies. These links from the BF enterprise unit to administrative system numbers are the key to integrating these administrative data sources with other data collected by Statistics New Zealand.
 - Source of standard classifications for household surveys that collect details on people's employers. By matching the name of the employing business to the BF, the standard classifications of these employing businesses (eg industry code) from the BF can be used. When this occurs there will be consistency in classifications across business and household surveys.
 - Source of statistics on business numbers and business demographics.

To produce high quality economic and financial statistics, Statistics New Zealand aims to ensure that the BF is fit for use by being:

- Complete in terms of coverage
- Free from duplicates and ceased units (for example, enterprises that are no longer in existence)
- Accurate in terms of classifications and size variables
- Timely in terms of reflecting real world changes on the BF.

3. Use of administrative data by Statistics New Zealand

Statistics New Zealand has a long history in the use of administrative data in the production of official statistics. Statistics New Zealand has used tax data from Inland Revenue (the tax office) in the production of statistics since the 1920s. Up to the mid-1980s, the statistics were derived solely from annual tax return data. They included incomes of individuals, self-employed workers and companies. There was no integration with other data sources.

From the mid-1980s, Statistics New Zealand incorporated tax data with other Statistics New Zealand data. This change was made possible by Statistics New Zealand reconciling the BF with GST¹ registered businesses on the Inland Revenue Client Registration database. As part of the reconciliation (of the BF with the Client Register information), a concordance from the BF to a business tax number has been maintained. This link from the BF enterprise statistical unit to the businesses tax number has been the key to the expanded use of tax data within the official statistical system.

¹ GST means Goods and Services Tax. This is a VAT tax that covers almost all business activity.

The maintenance of the BF (as detailed below in section 8) is used:

- To supplement Statistics New Zealand collected data in a wide range of business statistics outputs. A common model is Statistics New Zealand sourced data (eg via a survey questionnaire) for large and complex businesses with tax data for small and medium businesses
- As the main data source for a statistical output. With standard classifications (eg industry and sectors codes) from the BF for business tax units. Examples include a quarterly business activity indicator series derived from GST tax data and a linked employer-employee database derived from employment tax data. Both of these series use the link to the BF to access the standard enterprise classifications from the BF. The link to the BF also allows integration of these series with other statistical outputs based on the BF.

Tax data is the prime administrative data source used to maintain the BF. Other administrative data sources are also used from the Companies Office, which is responsible for the registration of limited liability companies in New Zealand; the Government Actuary Office, which is responsible for the registration of superannuation schemes; and we are investigating the use of Government Accident Compensation scheme information.

The integration of tax data into the BF enables tax data to be used more extensively in the broader programme of economic statistics:

Annual Enterprise Survey (AES) – this survey measures the financial performance and financial position of New Zealand businesses by industry and is used in the compilation of the Annual National Accounts. Tax data from New Zealand's tax agency, Inland Revenue, is used in a variety of ways. Annual accounts summary information (IR10) provides data for all un-incorporated businesses in all surveyed industries and for all businesses in the agriculture industries. Goods and services tax (GST) and employment counts from the Employer Monthly Schedules (EMS) are used to improve the efficiency of the sample design and the statistical calculation methods.

Sub-annual financial statistics – tax data is also used in the redesigned monthly retail trade survey (RTS), quarterly wholesale trade survey (WTS) and quarterly manufacturing survey (QMS). GST and EMS data is used in bivariate stratification in a similar fashion to AES. GST is used to replace direct surveying of small businesses. Currently the threshold for the contribution of modelled administrative data to total industry estimates is 10–15 percent.

Business Activity Indicator – this is a monthly experimental series that was developed in the late 1990s and measures changes in activity of the economy for businesses registered GST. It maps GST data to the statistical units on the Statistics New Zealand BF to model monthly GST sales by industry and region.

National Accounts Data Confrontation – the availability of the tax population enables us to compare the estimates from surveys to aggregates in the tax system. For example, estimates of total profit from surveys can be compared to total tax paid by businesses to assess the coherence of both estimates.

In the wider Statistics New Zealand perspective, the use of administrative data in the production of official statistics is wide-ranging and expanding. Examples include:

- Tax data as described above
- Customs documentation of exports and imports of goods
- Arrival and departure information from airports and ports

- Births/deaths
- Marriages
- Student loans.

In terms of volumes of data used by Statistics New Zealand, it is estimated that the volume of administrative data used significantly exceeds the volume of data directly collected by Statistics New Zealand survey activity.

4. Political and legal environment associated with the use of administrative data in the production of official statistics in New Zealand

Over recent years, there has been a strong direction from Central Government and from the business community in New Zealand on the need for all government departments and agencies to minimise the compliance cost imposed on business by regulations and requests for information. A government compliance cost committee, which included elected members of Parliament, has been very active in facilitating and promoting constraint and minimisation of government compliance cost.² This has required Statistics New Zealand to consider all its information requests in terms of the compliance costs imposed and the use of alternative existing data sources. This has strengthened and reinforced the need for Statistics New Zealand to use administrative data, where it exists, for the production of official statistics.

The Statistics New Zealand published strategic directions³ specifically noted this focus on the use of administrative data. Statistics New Zealand Strategic Outcome 3 is to maintain the cooperation of data providers. A key initiative to achieve this outcome is to "more effectively utilise existing data sources, particularly administrative data."

Statistics New Zealand's use of administrative data in the Business Frame (BF) and other areas needs to fit with the legislative rules controlling the organisations that collect the administrative data. With the tax example, in New Zealand the Inland Revenue Department (IRD) is responsible for the operation of the tax system. The work of IRD is largely governed by the Tax Administration Act. This Act gives the head of IRD the authority to supply data to Statistics New Zealand for statistical purposes.

The Statistics Act provisions regarding the security and confidentiality of individual business data used in the production of official statistics preclude Statistics New Zealand from providing any other government department or other organisation, including the Courts, access to individual data about businesses or individuals.

A memorandum of understanding (MOU) between IRD and Statistics New Zealand defines:

- The relationship between IRD and Statistics New Zealand, including the relevant legislative requirements related to the supply and use of the tax data
- The tax data supplied to Statistics New Zealand
- The conditions under which Statistics New Zealand can use the tax data
- A joint working committee that meets quarterly to facilitate the data supply and to keep both organisations up to date with developments on the use of tax data by Statistics New Zealand and the collection of tax data by IRD.

² Compliance Cost Committee 2003 Report
<http://www.med.govt.nz/bustl/compliance/report-back-2003/overview/index.html>

³ Statistics New Zealand Strategic Directions 2002 and beyond
<http://www.stats.govt.nz/about-us/corporate-reports/strategic-plan/default.htm>

Similar arrangements exist between Statistics New Zealand other administrative data suppliers. The BF also uses administrative Companies Office information as a maintenance source. The MOU with the Companies Office is much simpler than the IRD example as all the Companies Office information is publicly available and does not have the confidentiality and privacy issues associated with tax data.

The enabling conditions in the administrative data collecting organisation's legislation/ governing rules (that allows for the information to be supplied to Statistics New Zealand for statistical purposes) are important. Some potential administrative data sources in New Zealand do not have this enabling legislation and this creates a barrier to the data being used for statistical purposes.

5. Benefits of using administrative data

The benefits of using administrative data are significant for Statistics New Zealand.

The IRD and other government departments that collect administrative data carry out their legislated tasks in a professional, quality and efficient manner. They have a typically large infrastructure of systems, people and processes to carry out their work that the statistics office could never hope to replicate. So by leveraging off the data collected by these other organisations, significant benefits are obtained.

- Data that would not otherwise be available to the statistics office. The administrative data in these examples provides:
 - Statistical outputs that could not otherwise be produced, in terms of both topics and fine desegregation. Administrative sources often give complete, or almost complete, coverage of the target population, whereas sample surveys often only directly cover a relatively small proportion. The use of administrative sources therefore eliminates survey errors, removes (or significantly reduces) non-response, and provides more accurate and detailed estimates for various sub-populations, eg businesses in small geographic areas. Coverage is often of particular interest from the point of view of statistical business registers.
 - Efficiencies for the statistics office. The respondent contact and data collection costs are not incurred. But it is important to note that most uses of administrative data will still require substantial resourcing to cover the development and ongoing work (ie transforming administrative data to the required statistical outputs), as well as the data management costs.
- Reduced compliance cost. Using data from administrative sources helps to reduce the response burden created by surveys. While businesses may usually understand the reasons for supplying data for registration and taxation purposes, even if they do not like doing so, they may see statistical data requests as an extra, less necessary, burden. If they have already provided details to other government departments, they may become annoyed at receiving similar requests from the national statistical institution. Statistics New Zealand has made significant reductions in business compliance costs over the past five years by using tax data in business survey outputs.

The BF has all these benefits from the use of administrative data. The size and scope of statistical business registers makes it very unlikely that they can be satisfactorily populated and maintained solely by survey data. Progressively over the past 20 years, as the use of tax and other administrative data to maintain the BF has been extended:

- BF quality has improved. Coverage, timeliness and accuracy have all improved
- BF operational costs have been reduced

- Business compliance costs resulting from BF frame update surveys have been reduced.

An example of these improvements is the outcomes of a project implemented over 2002 to 2004 that significantly extended the use of tax data to maintain the BF.

Item / Process	Benefits resulting from the 2002-04 project to extend the use of tax data to maintain the BF	Comments
BF coverage.	Improved. Gaps in the BF coverage fixed. Approximately 30,000 units added to the BF.	These were all small businesses within scope of the BF coverage. In total they contributed less than 1 percent of total estimated economic activity.
BF enterprise employment size measure.	Timeliness and quality improved. Previously only updated annually and not for all enterprises. Now updated monthly for all enterprises using tax data.	
Compliance cost from frame update surveys.	Reduced by over 20 percent.	An extended use of tax data reduced the need for Statistics New Zealand surveys.
Statistics New Zealand resources required to maintain the BF.	Reduced by 10 percent.	A result of more automated use of tax data to maintain small businesses on the BF.
Tier approach to frame maintenance. (Stratification of the BF into large/medium/ small businesses.) Tax data used automatically to update small businesses. Person effort focused on large businesses.	Improved BF quality for large and small businesses. Previously some of the small businesses did not receive any maintenance activity except for life cycle events (eg ceasing).	Allowed the focus for person resources used for frame maintenance to be on large / complex businesses that dominate the economy.

6. Confidentiality and privacy issues related to the use of administrative data

Consideration of the confidentiality and privacy issues related to the use of administrative data are very important in terms of providing confidence to the wider community that the statistics office and the administrative data collecting organisations are complying with the law and information privacy requirements. All uses of administrative data by Statistics New Zealand are in concordance with the conditions of the Statistics Act 1975. As well any additional 'inherited' requirements from the organisation that originally collected the administrative data are recorded in the MOU documents noted above. It is very important to note that Statistics New Zealand supplies no unit record data to the organisations that supply the administrative data. It is a one way flow from the suppliers to Statistics New Zealand.

Some of the Statistics New Zealand uses of administrative data include the use of administrative unit records about individuals (eg employees). The use of data related to individuals in New Zealand is governed by the Privacy Act 1993. Statistics New Zealand has developed specific protocols to address the requirements of this Act when dealing with these datasets and related data integration projects.

Communicating the benefits of the increased use of existing data sources can help to enhance the prestige of a national statistical office by making it more efficient and cost-effective and extending the range of official statistics.

7. Challenges in the use of administrative data for statistical purposes

The organisation responsible for the collection of administrative data (eg tax) operate under the rules and regulations defined for the administrative processes (eg collection of tax). To use this administratively sourced data in the statistical system, the statistician needs to:

- Clearly understand the rules and processes that define the administrative data. These rules and processes will influence the coverage, timing, quality, and completeness of the administrative data.
- Understand the differences between the model defining the statistical data requirements and the administrative data that is available, eg statistical units on the BF versus units defined in the tax system.
- Develop appropriate methodology and processes for the statistics office to transform, model or adapt the administrative data to the statistical model. For business statistics in New Zealand, outcomes from this process have included:
 - Use of tax data for small and medium-sized businesses with a simple structure where the statistical unit directly matches the tax unit structure. For large and complex businesses, Statistics New Zealand collected data is used
 - Using models to derive the required statistical outputs from tax data. For example modelling two and six monthly GST tax returns to produce quarterly (three-monthly) data
 - Using statistical techniques such as estimation/forecasting to address timeliness issues. This is done in New Zealand for quarterly economic surveys when only two months of tax data are available in the required time frame
 - Using tax data that is correlated to a required statistical variable (not available in the tax system) to model the required variable
 - Accepting the administrative data as it is and being very clear with users on defining the statistical outputs produced. Also providing overlaps between series on the old and new definitions to help users make the transition. This is occurring now in Statistics New Zealand with the move of some business employment series to data that is based on the tax system definitions which differ from previous concepts used in Statistics New Zealand surveys

- Requesting changes to administrative rules so that the administrative data better fits with the statistical requirements.

The administrative data supply could change or stop with little notice, for example with a legislative change influencing the tax system. This is a risk and needs to be mitigated by a strong relationship with the supplying organisations; agreement that Statistics New Zealand will receive warning about impending changes and will be included as a stake holder in planned changes; and recognition from a whole of government approach of the benefits of the official statistics produced from the administrative source.

8. How is administrative data used to maintain the Business Frame (BF)?

The BF is maintained on a continuous basis using information from a range of sources. These sources include tax data, Companies Office data, survey feedback, media reports, company reports, etc. Tax data is a key source and provides the foundation of the BF. The BF holds the tax registration number for every enterprise on the BF. This link from the BF to tax data is an essential facilitator in Statistics New Zealand's use of administrative tax data to maintain the BF and also its use of tax data in statistical outputs.

Units on the BF are stratified into three tiers on the basis of their structure, GST (Goods and Services Tax, a VAT type tax) activity and employment numbers. These tiers are used in the BF maintenance strategy to ensure that the person resources used to maintain the frame are focused on large businesses that dominate the economy.

Tier	Description	Number of enterprises	Contribution to total GST activity
1	Large and/or complex enterprises	60,000	85 percent
2	Medium sized enterprises	110,000	10 percent
3	Small sized businesses	250,000	5 percent

All units are updated using tax data as described below. For tier 3 enterprises, tax data is the prime source of update information and most updates are automated. For tier 1 and 2 enterprises, frame update survey and other information is used to supplement and confirm the tax information.

BF attribute / process	Use of administrative data	Comments
Adding new enterprises to the BF.	<p>Tax data is the main source for all tiers.</p> <p>Each month new business tax registrations that meet the BF minimum economic significant criteria are added to the BF. Standard BF classifications of industry, sector and geo-code are derived from tax registration information. Enterprise contact details (eg postal address etc) on the BF are populated from tax registration information.</p>	<p>New tier 1 and 2 enterprises are also surveyed with a Statistics New Zealand frame update survey questionnaire to confirm the classification information derived from tax data. Also to collect frame information not available from tax data:</p> <ul style="list-style-type: none"> ➤ Ownership relationships between enterprises ➤ Enterprises with more than one business location ➤ Overseas ownership and financial activity indicators.

BF attribute / process	Use of administrative data	Comments
Ceasing enterprises on the BF.	<p>Tax data is a key source for all tiers.</p> <p>On a monthly basis enterprises on the BF are ceased if their tax system registration record is ceased or tax data indicates that the enterprise's level of financial activity has moved below the BF economic significance criteria.</p> <p>Companies Office information on limited companies that are ceased on the company register are ceased on the BF. This is a monthly process.</p>	
Employment size measure.	<p>Tax data is the main source for all tiers.</p> <p>The BF enterprise employment size measure is updated monthly for all enterprises using tax data.</p>	
GST \$ activity size measure.	<p>Tax data is the only source.</p> <p>The BF enterprise GST \$ size measure is updated monthly for all enterprises using tax data.</p>	
Contact details for existing enterprises (eg postal address).	<p>For tier 3 enterprises BF contact details are updated to the latest tax record information if Statistics New Zealand has had no survey interaction with the unit over the past year.</p>	<p>Tier 1 and 2 enterprises are surveyed with a frame update questionnaire to confirm and update all their BF information. Tier 1 units every year. Tier 2 units every third year or more frequently if they display unusual levels of tax activity.</p>
Legal name changes to limited liability companies on the BF.	<p>Companies Office information is used monthly to update the legal names of companies on the BF that register a name change with the Companies Office.</p>	
Ownership relationships between companies.	<p>From late 2005 this BF information will be updated annually from administrative Companies Office information.</p>	<p>This will reduce the information collected in the frame update surveys. Currently the main source of this ownership relationship information is from the frame update surveys used for tier 1 and 2 enterprises.</p>

The BF has a quality monitoring programme to understand and report on the quality being obtained from these uses of administrative data. Research also identifies opportunities for improving quality and extending the use of administrative data.

References

Eurostat Business Register manual,

http://forum.europa.eu.int/irc/dsis/bmethods/info/data/new/embs/registers/embs1_5.html

Compliance Cost Committee 2003 Report,

<http://www.med.govt.nz/buslt/compliance/report-back-2003/overview/index.html>

Statistics New Zealand Strategic Directions 2002 and Beyond,

<http://www.stats.govt.nz/about-us/corporate-reports/strategic-plan/default.htm>