

Business Demographic Statistics Review Report
*Based on Statistics New Zealand's
Longitudinal Business Frame*

May 2006

Acknowledgement

This report was prepared by the Statistics New Zealand's Business Financial and Structural Statistics unit and produced by the Product Development and Publishing unit.

Further information

For further information on the statistics in this report, or on other reports or products, contact Statistics New Zealand's Information Centre.

Visit our website: www.stats.govt.nz
or email us at: info@stats.govt.nz
or phone toll free: 0508 525 525

Auckland		Wellington		Christchurch	
Private Bag	92003	PO Box	2922	Private Bag	4741
Phone	09 920 9100	Phone	04 931 4600	Phone	03 964 8700
Fax	09 920 9198	Fax	04 931 4610	Fax	03 964 8964

Information Centre

Your gateway to Statistics New Zealand

Statistics New Zealand collects more than 60 million pieces of information each year. New Zealanders tell us how and where they live and about their work, spending and recreation. We also collect a complete picture of business in New Zealand. This valuable resource is yours to use. But with all the sophisticated options available, finding exactly what you need can sometimes be a problem.

Giving you the answers

Our customer services staff can provide the answers. They are the people who know what information is available and how it can be used to your best advantage. Think of them as your guides to Statistics New Zealand. They operate a free enquiry service where answers can be quickly provided from published material. More extensive answers and customised solutions will incur costs, but we always give you a free, no-obligation quote before going ahead.

Liability statement

Statistics New Zealand gives no warranty that the information or data supplied in this report is error free. All care and diligence has been used, however, in processing, analysing and extracting information. Statistics New Zealand will not be liable for any loss or damage suffered by customers consequent upon the use, directly or indirectly, of information in this report.

Reproduction of material

Any table or other material published in this report may be reproduced and published without further licence, provided that it does not purport to be published under government authority and that acknowledgement is made of this source.

Contents

Executive Summary	vi
Report outline	vi
Key results	vii
1. Introduction	1
2. Statistics New Zealand's Longitudinal Business Frame	3
2.1 Source data for the LBF	3
2.2 Longitudinal links	5
<i>The need for a clearer picture on births and deaths</i>	5
<i>Criteria applied for linking</i>	5
<i>Processes to identify longitudinal links</i>	6
<i>Unique identifiers on the LBF</i>	8
2.3 Strengths and weaknesses of the LBF	8
<i>Strengths</i>	8
<i>Weaknesses</i>	9
3. Potential Changes to Business Demography Statistics	10
3.1 Population of interest	10
<i>Current business demography population</i>	10
<i>LBF population</i>	10
<i>Population used for experimental business demography series</i>	11
3.2 Stock of active enterprises	11
<i>Current business demography stock of active enterprises</i>	11
<i>Eurostat recommendation on the stock of active enterprises</i>	12
<i>Options available for the stock of active enterprises using the LBF</i>	12
3.3 Identification of enterprise births and deaths	12
<i>Current business demography methodology</i>	12
<i>Eurostat recommendations on births and deaths</i>	13
<i>Births, deaths and restructures on the LBF</i>	13
<i>Separating entries and exits from births and deaths</i>	14
3.4 Frequency of output	17
<i>Current business demography methodology</i>	17
<i>Options using the LBF</i>	17
3.5 Business size measure	18
<i>Current business demography methodology</i>	18
<i>Business size measures on the LBF</i>	18
3.6 Updates to business demography data	19
<i>Current business demography methodology</i>	19
<i>LBF development</i>	19
4. Presentation of Experimental Statistics	20
4.1 Stock of active enterprises and geographic units	20
<i>Total number of enterprises and employees</i>	20
<i>Industry</i>	24
<i>Business size</i>	25
<i>Region</i>	27
<i>Business type and institutional sector</i>	29
4.2 Enterprise births	31
<i>Number of births and birth rates</i>	32
<i>Industry</i>	36
<i>Business size</i>	37

4.3 Enterprise deaths	39
<i>Number of deaths and death rates</i>	39
<i>Industry</i>	41
<i>Business size</i>	42
4.4 Survival and transitions	45
<i>Survival of enterprise births</i>	45
<i>Continuation rates for continuing enterprises</i>	46
<i>Growth and loss of employment 2000–2005</i>	48
5. Future Development and Call for User Feedback.....	49
5.1. Plan and timeframe for system development.....	50
5.2 Opportunity for user feedback.....	51
A. Appendix.....	52
A.1 Change in Business Frame maintenance strategy	52
A.2 Changes in employment data	52
A.3 Industry coverage	53
A.4 Glossary of terms.....	55
Bibliography	52

Figures

1. Statistical model to structure enterprises on the business frame	4
2. Possible lifespans of enterprises across annual snapshots	15
3. Graphical representation of birth identification	16
4. Graphical representation of death identification	16
5. Number of enterprises, at February	20
6. Number of employees, at February.....	21
7. LBF enterprise and employment growth, at February	22
8. BD enterprise and employment growth, at February.....	22
9. Number of enterprises by industry, at February 2005 compared with February 2000	24
10. Proportion of enterprises and employees by business size groups, at February 2005.....	25
11. Average employee count, all businesses and employing-only businesses	26
12. Number of geographic units by broad regions	27
13. Proportion of geographic units and employees by regional council, at February 2005	28
14. Proportion of total enterprises by business type, at February	29
15. Proportion of employees by sector, at February	30
16. Entries and births excluding agriculture, at February	32
17. Entries and births including agriculture, at February	33
18. Birth rates, at February.....	33
19. Identified LBF births by subcategory, at February.....	34
20. Identified LBF pure births by subcategory, at February	35
21. Identified LBF births by industry, 2001, 2003 and 2005	36
22. Identified LBF births by size group, 2001, 2003 and 2005	37
23. Employees of LBF births by size group, 2001, 2003 and 2005.....	38
24. Exits and deaths excluding agriculture, at February	39
25. Exits and deaths including agriculture, at February	40
26. Death rates, at February	40
27. Identified LBF deaths by industry, 2001, 2003 and 2005	41

28.	Identified LBF deaths by size group, 2001, 2003 and 2005	42
29.	Employees of LBF deaths by size group, 2001, 2003 and 2005.....	43
30.	Average employee count, identified LBF births and deaths	44
31.	Enterprise survival rates by size group, all births in 2001	45
32.	Enterprise survival rates by industry, all births in 2001	46
33.	Enterprise continuation rates by size group, incumbent enterprises in 2001	47
34.	Growth and decline of continuous businesses, February 2000 to February 2005	48

Tables

1.	List of LBF variables available for business demography	4
2.	Annual percentage change in enterprise numbers from LBF, including agriculture	23
3.	Annual percentage change in employment count from LBF, including agriculture	23
4.	Overview of BD population, entries and exits.....	31
5.	Overview of LBF population, births and deaths.....	31
6.	LBF births by subcategories.....	31
7.	Continuation rates of incumbent LBF enterprises, February 2000 to February 2003.....	47

Abbreviations

ANZSIC	Australia and New Zealand Standard Industrial Classification
BD	Business Demography
BF	Business Frame
EC	employee count
EMS	Employer Monthly Schedule
ENT	enterprise
FTE	full-time equivalent
GEO	geographic unit
GST	Goods and Services Tax
IBULDD	Improved Business Understanding via Longitudinal Database Development
IRD	Inland Revenue Department
ISIC	International Standard Industrial Classification of all Economic Activities
KAU	kind-of-activity unit
LBF	Longitudinal Business Frame
LEED	Linked Employer-Employee Data
NZSIC	New Zealand Standard Industrial Classification
PAYE	pay as you earn
PBN	permanent business number
RME	rolling mean employee count
STEPS	Business Surveys Tax Expanded Population Strategy

Executive Summary

Report outline

- This paper is the final report of stage 1 of the Business Population Statistics Review Project, which looks into the feasibility of producing a new range of business demography statistical outputs that better meet the needs of stakeholders.
- The report provides experimental business demography statistics for New Zealand for 2000–2005 and discusses the underlying methodologies. The data source for the experimental series is the Longitudinal Business Frame, which was developed by the Linked Employer-Employee Data project (LEED).
- The Longitudinal Business Frame is a more statistically robust data source for business demography in terms of its maintenance and the enterprises covered than the currently used Business Frame. It provides a rich panel dataset of monthly information on all active business units.
- The Longitudinal Business Frame holds historical data back to April 1999 and is updated monthly. It facilitates the creation of a consistent time series for business demography from 2000–2005 without methodological breaks, and allows for updates of previously published data if necessary.
- Basing business demography statistics on the Longitudinal Business Frame addresses current limitations in the identification of enterprise births and deaths as well as in the scope of enterprises included in the population.
- Newly developed methods make it possible to identify with greater certainty the level of real enterprise births and deaths, as opposed to enterprise entries and exits that include dormant enterprises, reactivations, and administrative churn. This is in line with the development of international standards for business demography by Eurostat and the Organisation for Economic Co-operation and Development, and comparable to methods used by other countries.
- The results presented in this report show that basing business demography statistics on the Longitudinal Business Frame is feasible and that the differences between the experimental and official series are explicable. The calculated birth and death rates align well with international results from Europe, the United Kingdom, Canada, the United States and Australia, where birth rates lie between 9 and 14 percent. Nevertheless, international comparisons have to be treated with caution due to differences in the underlying definitions between countries.
- It is proposed to develop a system to put the new business demography statistics into production in the first half on 2007. The first official publication of business demography statistics based on the Longitudinal Business Frame can be expected by the end of 2007.
- Data on the Longitudinal Business Frame can in principle be linked with other data sources of Statistics New Zealand. Current developments are testing the feasibility of placing business demography statistics within a wider context of business performance measures.
- Statistics New Zealand welcomes user feedback on the methods and results presented in this report. Of foremost interest is whether users feel that the methods applied to identify enterprise births and deaths and the resulting statistics on business dynamics better support their needs.

Key results

Results of Experimental Series Compared with Official Series

Statistics	Experimental series (excluding agriculture)	Official series as published October 2005 (excluding agriculture)	Difference
Number of enterprises as at February 2005	369,402	334,340	+35,062
Number of geographic units as at February 2005	402,564	366,128	+36,436
Number of employees as at February 2005	1,755,100 (LEED)	1,726,140 (Business Frame)	28,960
Annual change in number of enterprises 2004–2005	2.8%	3.1%	-0.3
Annual change in number of employees 2004–2005	3.4%	5.2%	-1.8
Percentage of enterprises with less than 10 employees as at February 2005	93%	92%	+1
Percentage of geographic units in Auckland region as at February 2005	36%	35%	+1
Identified enterprise entries/births as at February 2005	47,875 (births)	58,144 (entries)	-10,269
Identified enterprise exits/ deaths as at February 2005	37,953 (deaths)	48,097 (exits)	-10,144
Entry rate/birth rate as at February 2005	13% (birth rate)	17% (entry rate)	-4
Exit rate/death rate as at February 2005	10% (death rate)	14% (exit rate)	-4
Turnover rate as at February 2005	23%	31%	-8
Percentage of enterprises entered/birthed in 2001 that survived until 2005	51%	50%	+1
Percentage of enterprises entered/birthed in 2003 that survived until 2005	69%	66%	+3

Results of Experimental Series

Statistics	Experimental series (including agriculture)	Experimental series (excluding agriculture)	Difference
Number of enterprises as at February 2005	433,089	369,402	63,687
Number of geographic units as at February 2005	469,759	402,564	67,195
Number of employees as at February 2005	1,835,450 (LEED)	1,755,100 (LEED)	80,350
Annual change in number of enterprises 2004–2005	2.2%	2.8%	-0.6
Annual change in number of employees 2004–2005	3.3%	3.4%	-0.1
Percentage of enterprises with less than 10 employees as at February 2005	93%	93%	0
Percentage of geographic units in Auckland region as at February 2005	31%	36%	-5
Births as at February 2005	51,934	47,875	4,059
Deaths as at February 2005	42,283	37,953	4,330
Birth rate as at February 2005	12%	13%	-1
Death rate as at February 2005	10%	10%	0
Turnover rate as at February 2005	22%	23%	-1
Percentage of enterprises entered/birthed in 2001 that survived until 2005	52%	51%	+1
Percentage of enterprises entered/birthed in 2003 that survived until 2005	70%	69%	+1

1. Introduction

This paper reports on the creation of an experimental series of business demography statistics based on the Longitudinal Business Frame. It discusses the underlying methodologies developed to construct the Longitudinal Business Frame and shows how an enhanced business demographic series that aligns with international best practice can be derived from it.

The paper is the final report of stage 1 of the Business Population Statistics Review Project. This project looks into the feasibility of producing a new range of business demography statistical outputs that better meet the needs of stakeholders. The goal of Stage 1 is to ensure that established users are in agreement with the application of the new concepts and methods presented. If general support can be ascertained, stage 2 of the project will be initiated to bring the experimental series into production.

Business demography statistics are at present an annual snapshot (as at February) of the structure and characteristics of New Zealand businesses on Statistics New Zealand's Business Frame. The Business Frame is a list of individual, private and public-sector businesses and organisations that are engaged in the production of goods and services in New Zealand. It is maintained using information from Inland Revenue, such as Goods and Services Tax registrations and Employee Monthly Schedule returns, as well as Statistics New Zealand's survey information. The prime objective of the Business Frame is to provide an up-to-date and accurately classified list of the target population for designing, selecting and operating Statistics New Zealand's economic and financial survey programme.

Major limitations of the current business demography series are:

- The current methodology for identifying enterprise births and deaths from the Business Frame cannot distinguish between genuine enterprise start-ups (failures) on the one hand, and enterprise entries (exits) due to administrative changes on the other hand.
- The population base for the current series is limited to enterprises beyond certain thresholds of economic activity, and excludes the agriculture industry.
- The current business demography system does not have the functionality to update previously published statistics if information relating to past periods changes on the Business Frame.

International developments in business demography have given rise to new methods to define and recognise enterprise births and deaths. This was a key factor in the construction of the Longitudinal Business Frame as an alternative source for business demography statistics. On the Business Frame, enterprises that have undergone administrative changes (and thus changed identifiers such as tax numbers) may appear as enterprise deaths with succeeding births, if they are not recognised as ongoing by existing processes. The Longitudinal Business Frame uses statistical methods to track enterprises over time and to link – if possible – new enterprise identifiers with their predecessors. Thus it is possible to separate enterprises that are continuing, dormant or reactivated from real enterprise births and deaths, following international recommendations.

The business demography statistics presented in this report are sourced from the Longitudinal Business Frame for the first time as an experimental series. The preliminary goals for the experimental series are:

- to overcome the shortcomings of the current series
- to assess the feasibility of using the Longitudinal Business Frame as data source for business demography, and to quantify the resulting changes in statistics
- to consult with users about definitions and options for the new business demography series, before methods are finalised and it goes into production.

The main differences to note between the official and experimental series are:

- The experimental series introduces a new method to separate real enterprise births and deaths from other enterprise entries and exits.
- The population for the experimental series moves towards complete industrial coverage, including the agriculture sector (A01). This not only affects business stocks, but also aggregated business dynamics.
- The experimental series uses employee counts from the Linked Employer-Employee Data (LEED) as business size measures.

Section 2 of the report summarises the development of Statistics New Zealand's Longitudinal Business Frame, and section 3 discusses the options and potential changes for business demography statistics when sourced from the Longitudinal Business Frame. Section 4 presents the experimental series for business demography statistics based on the Longitudinal Business Frame. The range of statistics shown is wider than in typical official publications, and consultation with users is sought on which graphs and tables should be included in forthcoming publications.

The paper concludes with an outlook for further development and a call for user feedback on various issues such as the methodologies applied, the scope, frequency and timing of the series, and output formats as mentioned above. Statistics New Zealand encourages readers to provide comments on the report.

Technical details and a glossary can be found in the Appendix.

2. Statistics New Zealand's Longitudinal Business Frame

2.1 Source data for the LBF

The Longitudinal Business Frame (LBF)¹ is an offshoot of the Linked Employer-Employee Data (LEED) integration project.² The LBF contains data from two main sources: Statistics New Zealand's Business Frame (BF), and payroll tax records drawn from New Zealand's taxation system. Of these, the BF is the predominant source, as it covers all businesses that are registered with Inland Revenue and meet the criteria for economic significance. This means that employing businesses on the LBF that are economically significant exist in both data sources. The unreported economy is outside the scope of the LBF.

A business is included in the BF (is economically significant) if it meets at least one of the following conditions:

- the business has annual Goods and Services Tax (GST) turnover of greater than \$30,000
- the business has paid employees
- the business is part of an enterprise group
- the business is part of a GST group
- the business has more than \$40,000 income reported on tax form IR10
- the business has a positive annual GST turnover and has a geographic unit classified to agriculture or forestry.

Data on a small number of businesses on the LBF comes directly from taxation records, as the businesses are not found on the BF. Instead, the data is sourced from the Employer Monthly Schedule (EMS) data. The EMS payroll return is filed monthly by employing enterprises and covers all payers and recipients of income that is taxed at source, other than interest and dividends. Enterprises that are out of scope of the BF *and* non-employed (not filing EMS) are not included on the LBF.

Attributes such as the industry, size or region of businesses on the BF are transferred to the LBF, while businesses that are sourced from outside the frame are updated with coded or imputed attributes. Businesses on the BF are structured according to a three-level statistical model comprising the enterprise unit, the kind-of-activity unit, and the geographic unit (business location) in accordance with the International Standard Industrial Classification of all Economic Activities, Third Revision (ISIC rev 3; see Figure 1 below). The LBF focuses mainly on geographic units, but also records information about the enterprises that the geographic units belong to, as well as the group top enterprise, if there is one.

1 For further reference, see Seyb A (2005). "Statistics New Zealand's Longitudinal Business Frame", in *Maximising Data Value. Data Use and Reuse*, Statistics New Zealand, Wellington.

2 For more information on LEED see Statistics New Zealand's website at:
<http://www.stats.govt.nz/leed/default.htm>

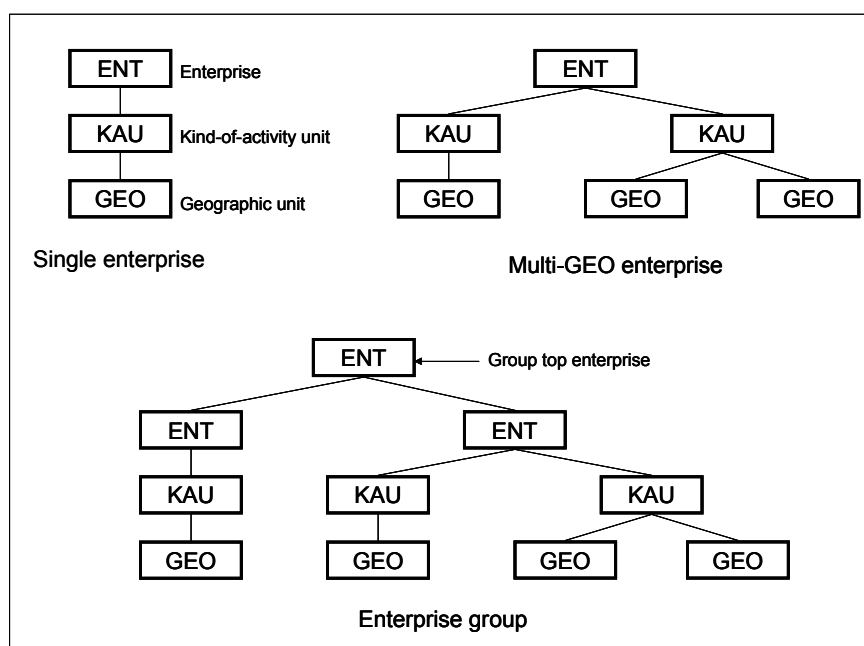
Table 1

List of LBF Variables Available for Business Demography

LBF statistical unit	Main attributes	Time attributes	Business size measure
Group top enterprise		Monthly snapshots from April 1999 to date	
Enterprise unit	<ul style="list-style-type: none"> • Predominant industry • Business type • Institutional sector • Overseas equity <p>Indicators: Economic significance</p>	<ul style="list-style-type: none"> • Birth date • Monthly snapshots from April 1999 to date <p>Indicators: Life-cycle code Live indicator</p>	<ul style="list-style-type: none"> • LBF employee count • LEED employee count
Geographic unit	<ul style="list-style-type: none"> • Ancillary industry • Primary industry • Meshblock • Area unit • Territorial authority • Regional council <p>Indicators: Economic significance</p>	<ul style="list-style-type: none"> • Birth date • Monthly snapshots from April 1999 to date <p>Indicators: Life-cycle code Live indicator</p>	<ul style="list-style-type: none"> • LBF employee count • LEED employee count
Inland Revenue only	Information of variable quality, may or may not be available <ul style="list-style-type: none"> • Industry • Meshblock 	Monthly snapshots from April 1999 to date	<ul style="list-style-type: none"> • LEED employee count

Figure 1

Statistical Model to Structure Enterprises on the Business Frame



The LBF records information taken from its data sources in monthly snapshots to construct a longitudinal dataset. Changes on the BF are stored in history tables, and for the construction of the LBF, this historical information on units and their attributes has been unwound month by month from April 1999 to present. As a result, a rich panel dataset was created, which consists of a large number of cross-sectional units observed over time.

2.2 Longitudinal links

The need for a clearer picture on births and deaths

The development of better methods to define and recognise enterprise births and deaths was a key factor in reconstructing the current business register (BF) into the LBF. The internationally accepted economic definition of an enterprise birth amounts to the creation of a combination of production factors, while the definition of a death amounts to the dissolution of a combination of production factors, both to occur with the restriction that no other enterprise is involved. The definition goes on to state that a distinction must be made between the real, observable world and its reflection in administrative files and statistical business registers. The current methodology used to register or deregister businesses on the BF cannot at first sight distinguish between genuine enterprise births or deaths on the one hand, and businesses merely changing administrative identifiers (but otherwise staying the same) on the other hand. Questionnaires have to be sent out to businesses to obtain more information (for more details, see *Transfers on the Business Frame*, below, and section 3.3.). Any change of identifiers breaks longitudinal links and makes it difficult to track an enterprise over time using methods that rely on these identifiers. The main challenge in the LBF is to repair these broken links.

Longitudinal links are broken when a business changes its unique identifier within either (or both) of the two data sources used to construct the LBF. Events in the life of a business, such as change in legal structure, restructuring (mergers, split-offs etc), or the business being sold as a going concern, may result in the business receiving a new IRD number in the tax system. Subsequently, new tax registrations appear on the BF as new enterprise structures with new reference numbers. If the enterprise is not identified as ongoing through a questionnaire, the administrative changes in identifiers and reference numbers shows on the BF as an enterprise birth for the succeeding enterprise and an enterprise death for the preceding enterprise. In the development of the LBF, better methods were found to track units over time to determine whether they are continuing (a link is then repaired) or discontinuing (a link ends).

Criteria applied for linking

In practice, criteria on which to base the combination of production factors as continuing include whether a business holds a majority of its original geographic units, keeps the same trading and legal name, continues to operate from the same location, and continues to employ most of its former employees. In contrast, indicators for a discontinuing business are whether a business formed new geographic units, has changed its trading and legal name, physically relocates to a new area, and mostly recruits new individuals.

Processes to identify longitudinal links

Three dynamic processes were developed to repair LBF longitudinal links, the first of which builds on the established process of unit transfers on the BF. Each process, by making complementary linkages to LBF units, is set up to support the others in minimising the likelihood of creating erroneous links. While the first two repair processes use the geographic unit from the BF to repair longitudinal links, the third repair process uses the Inland Revenue reporting unit from the taxation system. In general, the Inland Revenue reporting unit is the enterprise or a group of enterprises. The main objective of exploiting the repair information made by either the geographic unit or the enterprise unit is to isolate the economic birth or death of a *business*, which in business demography is defined around the enterprise or legal unit rather than around the geographic unit.³

Transfers on the Business Frame

The first process identifies geographic units that underwent a transfer from one enterprise to another on the BF, which means that the link has already been established on the BF. All births of employing enterprises that are detected on the BF receive a questionnaire to inquire if the birth is a result of a change of ownership or legal structure. Through this normal process of frame maintenance, businesses that changed their administrative identifiers on the BF, but were identified as ongoing, have their original geographic unit(s) transferred to the 'new' enterprise, and any new geographic units that have been created will be invalidated. Roughly 10,000 links between geographic units per year are established in this way.

Probabilistic matching of geographic units

Births on the BF in a particular month, which have not been linked to a predecessor on the BF, undergo probabilistic matching with all businesses that were on the LBF in the previous month. The process matches geographic units by various variables, including name (both trading name and standardised legal name), street address, and location. The tool used is Ascential's QualityStage v7.0.1 (formerly Integrity; see Seyb, 2005, pp 6–8).

Before a link is established, the probabilistic matching routine processes the data in six passes. The first two passes limit potential matches to units in the same area (meshblock) and look for an exact match on legal or trading name. The third pass takes advantage of the discriminating power of the telephone number. Later passes cast the net more widely by gradually extending the size of the region that possible pairs of units may belong to, and by using various combinations of location, industry, and name to match fields. Also, increasingly higher cut-off thresholds are used to identify links where one or more of the fields are in error. When a link is made, the LBF is updated by recording the details of the new identifiers. Roughly 1,500 to 1,900 links between geographic units per year are established by probabilistic matching.

Clerical review during the development of the LBF showed that the level of false positives (two identifiers were linked as belonging to the same business but in reality are not) was approximately 2 to 3 percent in any month, while the level of false negatives (the link has not been made to two identifiers that do belong to the same business) was no more than 5 percent of linked units. Using this method, approximately 15 percent of enterprise births in any month are shown to be continuing businesses. At the beginning of a tax year (April), the percentage of spurious births can be as high as 20 percent (Seyb, 2005, p 8).

³ OECD Working Party on Small and Medium-Sized Enterprises and Entrepreneurship (2002). "OECD Small and Medium Enterprise Outlook" (DSTI/IND/PME(2002)6/PART2), OECD, Geneva, p 5.

Linking by tracking employees

The final longitudinal link-repair process identifies continuing employers by tracking groups of common employees between enterprises, using the EMS data. The EMS payroll return is filed monthly and covers all payers and recipients of income that is taxed at source, other than interest and dividends. Two types of recipients of income are covered by the EMS: those who pay 'pay as you earn' (PAYE) income tax, and those who have withholding payments deducted. Generally, those individuals who have PAYE deducted are employees, while those who pay withholding tax are a subset of the self-employed.⁴

Each month, employers report their tax using the same unique identifier, the employer IRD number – a number that in general identifies the enterprise unit or a group of enterprises. Employers can be linked over time using the employer IRD number. Spurious births and deaths of reporting units in the EMS data occur when employees cease to be reported under a given employer IRD number, but in reality continue to work at the same location. Links can be established with a high degree of certainty between the predecessor and successor employer IRD numbers by tracking employees within LEED.

To establish links, the process compares new IRD numbers for a particular month to IRD numbers that filed in the previous month and to IRD numbers that have not filed for the relevant month. The method is effective when the predecessor business size is greater than three employees, and at least 70 percent of employees 'move' to the new IRD number. Additional restrictions are placed on the predecessors: the predecessor IRD number must cease to be used; and a majority of employees must move to the successor in the month immediately following the 'death' of the predecessor IRD number. In addition, certain industries are treated differently. In agriculture, for example, it is common for groups of employees to move en masse between farms, so the predecessor enterprises in agriculture are restricted to a minimum of 20 employees. The aim is to minimise the number of false negative links – where the process fails to recognise that two employers in different periods are in fact the same – without introducing excessive numbers of false positive links between employers – where two records are linked but in fact refer to different employers (Seyb, 2005, p 8). If a reporting unit on the EMS represents more than one enterprise on the LBF, restrictions placed around the predecessor IRD number and information sourced from the other two alternative repair processes are used to ensure a genuine match is not missed.

Between 600 and 900 links between geographic units per year are established using employee movements. Clerical review during the development of the LBF showed that the level of false positives was approximately 2 percent, while the level of false negatives was at most 5 percent of linked units. Using this method, approximately 10 percent of enterprise deaths in any month are shown to be continuing businesses; this corresponds to up to 37 percent of job and worker flows due to spurious enterprise deaths (Kelly, 2003, p 16). At the beginning of a tax year (April), the observed percentage of spurious births is as high as 17 percent. The employer repair is carried out only for employers that are a subset of the universe of businesses on the LBF, and that have not yet been repaired by the other two processes (Seyb, 2005, p 8).

4 Kelly N (2003). "Repairing EMS Employer Longitudinal Links", Statistics New Zealand, Wellington, p 4.

Unique identifiers on the LBF

The three repair processes culminate in the creation of the unique permanent business number (PBN), which is assigned to all geographic units and units sourced from tax data on the LBF. A PBN is created on the LBF after following LBF businesses from both data sources over time. It is the unique identifier of a geographic unit (or payroll unit) on the LBF: it remains the same over time, even if enterprise identifiers or IRD identifiers representing the unit change in the data sources.

The PBN is the smallest unit on the LBF. A new PBN is assigned to a geographic unit or payroll unit only if it is a new 'plant' as defined in the economic literature or 'establishment' as defined in the System of National Accounts 1993. The creation of the PBN is essential for identifying real births and real deaths of the enterprise unit on the LBF. In practice, when units sourced only from tax data are excluded from the population, PBN and geographic unit are synonymous.

2.3 Strengths and weaknesses of the LBF

Strengths

The most obvious advantage of the LBF is that the data is longitudinal and frequent (monthly). There are no restrictions in industry or geographic scope. Statistics on enterprise births and deaths, the survival and growth of enterprises, and business size (in terms of persons employed) can be produced by industry according to the Australia and New Zealand Standard Industrial Classification (ANZSIC), business type, institutional sector, percentage of overseas equity, and region. Split-offs, mergers, takeovers and restructuring are all visible in the database, which means analysts are enabled to better identify real enterprise dynamics.

The analysis of panel or longitudinal data is the subject of one of the most active and innovative bodies of literature in econometrics. This is partly because panel data provide such a rich environment for the development of estimation techniques and theoretical results. Basically, it enables the researcher to use time series cross-sectional data to examine issues that could not be studied in either cross-sectional or time series settings alone.⁵

The LBF population is larger than the population of employers in the LEED database and also larger than the population of enterprises on the BF that is the basis for the current business demography series. The advantage of thus having almost all businesses in New Zealand represented in the database, not just those beyond certain thresholds, is that it allows researchers to study business growth from the very beginning of an enterprise's life cycle. The accuracy of the information about enterprises, particularly those changing structure or legal character, is enhanced by using information about employees along with sophisticated matching techniques to confirm the nature of the changes.

Within Statistics New Zealand, the LBF forms a strong foundation for data integration across data sources. With the completion of the LBF, it is now possible to link its data with other administrative data or business surveys. Linking the LBF to other enterprise data of Statistics New Zealand is relatively straightforward, making it possible to extend the breadth and depth of information for LBF enterprises.

⁵ Green W (1993). *Econometric Analysis*, 2nd edition, Prentice Hall, Upper Saddle River, New Jersey, pp 464–465.

For example, the Improved Business Understanding via Longitudinal Database Development (IBULDD) feasibility study is currently underway at Statistics New Zealand, with the aim of linking annual enterprise data to the LBF to measure enterprise productivity and performance.⁶

Data integration is seen by policy and official statistics agencies as a growth area. It is viewed as a way of creating information-rich datasets with fewer resources, and with fewer burdens on respondents, than are required for censuses or surveys. The use of existing administrative data to create the LBF has the advantage that respondent burden is reduced, while the level and complexity of information available to researchers, analysts and policy makers is increased.

Weaknesses

While the data quality of the LBF varies by source, it can be expected that enterprises represented in both data sources have accurate and timely information. However, the LBF has limited information on industry and region for enterprises that are found only in the payroll data. Alternatively, small enterprises found in the BF and not in the payroll data may have attributes such as industry codes updated only once every three years from a questionnaire, if at all. The database also has no information about other characteristics of enterprises such as sales, purchases, or expenditures on research and development.

Births and deaths are observed at an enterprise level, since the emergence and disappearance of geographical units does not necessarily indicate a change in the enterprise life cycle. However, the longitudinal links on the LBF are constructed for geographical units, which are tracked over time. The LBF treats the enterprise identifier associated with a geographical unit merely as an attribute which may change at any time, reflecting administrative changes. It is therefore necessary to find suitable methods for deriving information on enterprise dynamics from the geographic unit level.

Statistics New Zealand plans to further improve the LBF according to user needs. Additional research into businesses that undergo extended periods of inactivity and fine-tuning of the repair processes may yield further improvements in quality. Future work includes the addition of approximately 250,000 self-employed individuals, who are not captured on the LBF at present. Information on the self-employed comes from annual tax returns, and so will be less timely than the monthly payroll data.

⁶ For more information see Statistics New Zealand's website at:
<http://www.stats.govt.nz/developments/longitudinal-database-development.htm>

3. Potential Changes to Business Demography Statistics

3.1 Population of interest

Current business demography population

Business demography statistics are an annual snapshot (as at February) of the structure and characteristics of New Zealand businesses on Statistics New Zealand's Business Frame (BF). As set out in section 2, the BF is a database of individual, private and public-sector businesses operating in New Zealand that has been maintained since 1986. To be included in the BF, a business must be registered with Inland Revenue and be economically significant (see section 2.1).

An enterprise is selected from the BF into the current business demography population if it meets at least one of the following conditions:

- it has annual Goods and Services Tax (GST) sales or expenses of more than \$30,000
- it has a rolling mean employee count (RME) of greater than three
- it is active in a GST-exempt industry (except residential property leasing and rental)
- it is part of a group of enterprises
- it has a new GST registration that is compulsory, special or forced
- it is registered for GST and involved in agriculture or forestry.

At February 2005, there were approximately 433,000 active enterprises on the BF, of which just short of 335,000 were selected as the population for business demography statistics. This population did not include enterprises in the agriculture industry (Australia and New Zealand Standard Industrial Classification (ANZSIC) A01)⁷ or enterprises introduced through the Business Surveys Tax Expanded Population Strategy (STEPS maintenance strategy) on the BF in 2003.⁸

For information on the change in BF maintenance and its implications for business demography statistics, see Appendix A.1. For a summary of changes to the business demography population since 1987 and its industry coverage, see Appendix A.3.

LBF population

As discussed in section 2, the Longitudinal Business Frame (LBF) obtains information about businesses from two data sources: the BF and employer monthly schedule (EMS) returns from Inland Revenue. In its present form, EMS data has been collected by Inland Revenue since April 1999; and collection and maintenance practices, as well as population inclusions and exclusions, have not changed substantially during that time. Statistics New Zealand receives EMS data from Inland Revenue monthly.

7 The agriculture industry (A01) is a discontinuous series in business demography statistics. Agriculture (A01) was included in 1998 and from 2004 to date. The recent inclusion of the industry was due to improved maintenance of agriculture (A01) on the BF from Statistics New Zealand's 2002 Agriculture Production Census and 2003 Agricultural Production Survey. Published business demography statistics continued to exclude the agriculture industry for the sake of comparability with previous publications.

8 The business demography methodology excludes the following: GST-exempt enterprises reporting IR10 tax returns, previously ceased BF units that have been reactivated solely due to the STEPS maintenance strategy, and the addition of employing units onto the BF that do not meet the business demography selection conditions.

Information about employers on the LBF generally comes from both the EMS payroll data and the BF, with the exception of businesses found only in the EMS. There are approximately 17,000 (4 percent) 'Inland Revenue-only' businesses on the LBF in any month (see section 2.1). More than 65 percent of businesses on the LBF are non-employing and only sourced from the BF.

Population used for experimental business demography series

The population used for the development of the experimental business demography statistics from the LBF has been limited to economically significant businesses on the LBF. Conditions for economically significant businesses are set out in section 2.1.

There are three main reasons for this:

1. The limitation to only economically significant businesses allows for a greater level of comparison with previously published business demography statistics (as outlined in the next section), as well as with other official economic statistics.
2. Inland Revenue-only businesses tend to be small, with limited information. For instance, only around half of these businesses have an industrial classification available.
3. International evidence shows that expanding the population base to include small enterprises below the economic significance threshold can subsequently result in higher birth and death rates. Research also shows that as long as the business population base is consistent over time, emerging trends should be internationally comparable.

However, the potential does exist to expand the current business demography population in the future as required.

Apart from economic significance, some additional criteria were applied to the population selection from the LBF, which correspond to restrictions on the current business demography population:

- enterprises must have at least one live geographic unit (live means not dormant, ceased or invalid)
- enterprises must be in New Zealand (units in institutional sector 6111 'rest of the world' are excluded)
- enterprises must be trading (exclusion of all units marked as non-trading on the BF).

The industry attribute of enterprises was recalculated as the predominant industry derived from an enterprise's geographic units, weighed by their respective sizes as given by the Linked Employer-Employee Data (LEED) employment count (see section 3.5). In some cases this derived industry was found to differ from the industry recorded for an enterprise on the BF. The predominant industry was also used to split the population into the agriculture sector and non-agricultural enterprises.

3.2 Stock of active enterprises

Current business demography stock of active enterprises

The current business demography stock of active enterprises consists of enterprises that are alive on the BF in the February month of the relevant reference period, and that meet the population selection conditions outlined in section 3.1.

Enterprises (and associated geographic units) that are birthed in the month of February are included in the current population selection, while enterprises (and associated geographic units) which cease in February are excluded.

Eurostat recommendation on the stock of active enterprises

Eurostat defines the stock of active enterprises as follows:⁹

The population of active enterprises consists of all enterprises that had either turnover or employment at any time during the reference period. If there is insufficient information on turnover or employment to determine whether or not an enterprise is active, then national methods leading to this aim will be accepted.

Options available for the stock of active enterprises using the LBF

As the LBF is updated on a monthly basis, the opportunity exists to either implement the current business demography methodology of annual snapshots, or potentially move closer to the Eurostat recommendation to observe a whole period of time. However, the latter presents additional issues that would need to be resolved.

The current business demography methodology is to extract active enterprises at February of each reference period – that is, take a snapshot of the data at one point in time. If the view on the data is to be extended to the whole reference period, it has to be agreed how to handle enterprises which change industry classification, business structure, or location during this time. Also, the results of this approach may not be as readily comparable to previous business demography statistics as when implementing a February snapshot.

A recent report by the Organisation for Economic Co-operation and Development (OECD) comparing methodologies in various countries shows that while Eurostat's business demography statistics are compiled using 'live during period' populations, almost all other sources favour the 'point in time' approach.¹⁰ For the experimental business demography series from the LBF, Statistics New Zealand reached a consensus with stakeholders in June 2005 to use an annual stock of active enterprises at the February month from the LBF.

3.3 Identification of enterprise births and deaths

Current business demography methodology

Births and deaths are currently identified in business demography by matching identifier variables associated with enterprises in a reference period (period T) with those of enterprises in the previous reference period (period T-1). The new enterprises in year T are considered births, while enterprises that have ceased to exist in year T are considered deaths.

The weakness of the current approach is that it fails to identify spurious births and deaths due to broken longitudinal linkages. Sometimes, businesses change administrative numbers within the tax and BF systems *without* undergoing substantial change in the real world – for example, when a business changes legal structure, or is sold as a going concern. However, if the identifiers have changed on the BF, then

9 Eurostat (2004). "Business Demography Recommendations Manual, Status August 2004", unpublished paper, Eurostat, Luxembourg, p 11.

10 Vale S (2006). *The International Comparability of Business Start-up Rates, Final Report*, OECD Statistics Directorate/Office for National Statistics, London, United Kingdom, pp 93–100.

the longitudinal links are broken (this is further discussed in *The need for a clearer picture on births and deaths* in section 2.2 above and *Births, deaths and restructures on the LBF* below). Some administrative changes are identified on the BF, because employing enterprise births that are detected on the frame are sent a questionnaire to collect their details and are asked if the birth is a result of a change of ownership or a change in legal structure. A link between the old and new units is established on the BF for such cases. The historical information available on the LBF shows that each year around 9,000–11,000 such links are established for geographic units on the BF.

Entries and exits observed in the current business demography population may also result from periods of dormancy, or enterprises temporarily moving out of scope of economic significance. A dormant enterprise is generally deactivated on the BF when it ceases to record economic activity, and reactivated when activity resumes. Enterprises that are dormant or not significant at the time of the annual snapshot will not be included in the population.

Eurostat recommendations on births and deaths

The *Eurostat Business Demography Recommendations Manual* states (Eurostat, 2004, p 20):

Enterprise births are defined (in Commission Regulation (EC) No 2700/98 of 17 December 1998 concerning the definitions of characteristics for structural business statistics) as follows: “A count of the number of births of enterprises registered to the population concerned in the business register corrected for errors. A birth amounts to the creation of a combination of production factors with the restriction that no other enterprises are involved in the event. Births do not include entries into the population due to: mergers, break-ups, split-off or restructuring of a set of enterprises. It does not include entries into a sub-population resulting only from a change of activity.”

The European Commission Regulation defines enterprise deaths as (Eurostat, 2004, p 35):

A count of the number of deaths of enterprises registered to the population concerned in the business register corrected for errors. A death amounts to the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event. Deaths do not include exits from the population due to mergers, take-overs, break-ups and restructuring of a set of enterprises. It does not include exits from a sub-population resulting only from a change of activity.

Births, deaths and restructures on the LBF

To observe enterprise dynamics over time from administrative data sources, it is crucial to be able to link businesses if their identifiers in the source change. A business may undergo several changes in its lifetime, in addition to birth and death. For example, legal or administrative entities may close down or emerge due to break-ups, mergers, split-offs, takeovers, or restructuring. Any of these events can result in the business obtaining a new identifier (eg new IRD number) in the BF or tax reporting system. A business would then appear as a birth or death in these systems. However, neither administrative changes nor the events mentioned above necessarily indicate the occurrence of a birth or death in the real world.

The following examples are possible scenarios resulting in a change of identifiers, but not amounting to enterprise births or deaths:

- An existing enterprise A assumes a new legal status as enterprise B with a new tax registration; for example, when changing from a partnership to a company. The enterprise continues to use the existing factors of production.
- Existing enterprise A ceases. Some or all of its functions continue as new enterprise B.
- An existing enterprise A splits off some of its functions to a new enterprise B. Enterprise A continues trading.
- Existing enterprises A and B amalgamate to form a new enterprise C.
- Existing enterprise A acquires all of the production factors of enterprise B.

While all of the situations described above involve the birth of new administrative identifiers, *none* of them are regarded as the birth of a new enterprise according to the Eurostat guidelines; these are all cases of ‘administrative churn’. However, enterprises cannot be linked over time by their IRD numbers (or BF identifiers), because either new numbers have come into use or existing numbers have ceased to be used.

The LBF system is designed to set up longitudinal links between businesses, and thereby allows real economic births and deaths to be separated from administrative reshuffles, which brings the methodology closer in line with the Eurostat recommendations.

During the development of the LBF, processes were put in place that aim to recognise when enterprises in different periods are actually the same business. The LBF repairs such broken longitudinal links by using a combination of probabilistic matching techniques and tracking employee movements between enterprises. For a detailed discussion of the linking processes, refer to section 2.2.

It should also be noted that the LBF has been constructed in such a way that the links resulting from the administrative process are not lost in the repair process. Preceding and succeeding enterprises can be identified.

Separating entries and exits from births and deaths

The current business demography methodology identifies enterprise entries and exits, of which births and deaths are a subset. The challenge for the experimental series from the LBF was to separate entries into real births and other entries; and similarly separate exits into real deaths and other exits. Once real births are identified, they can be classified further into:

- recent births versus reactivations of enterprises dormant for more than two years
- births surviving for at least one year versus births disappearing in the next period.

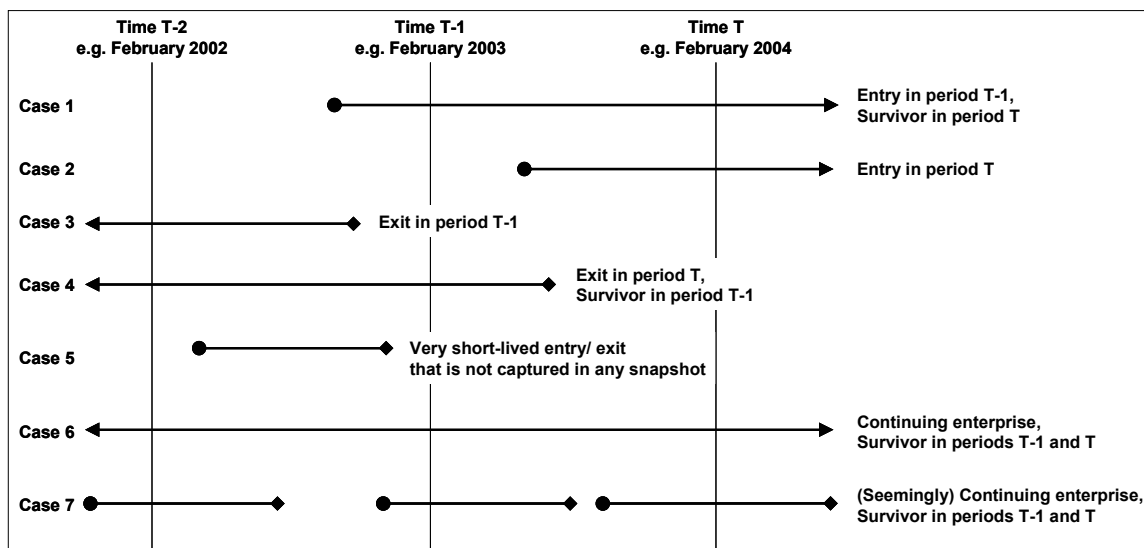
According to the recommendation of Eurostat for enterprise births and deaths, a reactivation after less than two years of inactivity is not counted as death and subsequent birth. To identify births at time T, it is therefore necessary to check movements in the enterprise population over more than one period – that is, at least back to time T-2 (see Figure 2). Looking back in time further than just one period to determine the status of an enterprise also helps to filter out temporary movements in and out of scope (as determined by the economic significance of an enterprise, which may change from one period to the next). Of course the number of periods we can

look back for births, or forward for deaths, is limited by the start and end points of the available data (the LBF holds data from April 1999 to the current month).¹¹

Figure 2 shows examples of entries and exits on a time-line. Annual snapshots are shown as vertical lines, and the lifespans of enterprises as horizontal lines.

Figure 2

Possible Lifespans of Enterprises Across Annual Snapshots



Proposed method to identify births on the LBF

- Total entries¹² of period T are all enterprises whose identifiers exist at time T but not at time T-1 (same methodology as for current business demography series).
- Of these, *real births* are all enterprises whose geographic units existed at neither time T-1 nor time T-2.
- If an enterprise consists of more than one geographical unit, it is only considered a birth if *none* of its units existed in the previous two years.
- Entries other than real births are enterprises that experience administrative changes or movements in and out of scope.

Proposed method to identify deaths on the LBF

- Total exits of period T are all enterprises whose identifiers exist at time T-1 but not at time T.
- Of these, *real deaths* are all enterprises whose geographic units exist at neither time T nor time T+1.
- If an enterprise consists of more than one geographical unit, it is only considered a death if *all* of its units disappear in the following two years.

¹¹ For enterprise births in 2001, the snapshots of April 1999 and February 2000 were used as reference points. For all other years only snapshots as at February were used.

¹² It should be noted that the population is selected in annual snapshots and not surveyed over the whole reference period. With this approach, very short-lived entries that occur after a snapshot is taken, but exit again before the next snapshot, will not be captured.

- Exits other than real deaths are enterprises that experience administrative changes or movements in and out of scope.
- If data for time T+1 are not available, the number of real deaths will be preliminary until it can be revised after the next snapshot is made.
- Review of the identified real deaths for 2001–2004 showed that they would have been overestimated by 7–8 percent if the next snapshot had not been available.

Figure 3

Graphical Representation of Birth Identification

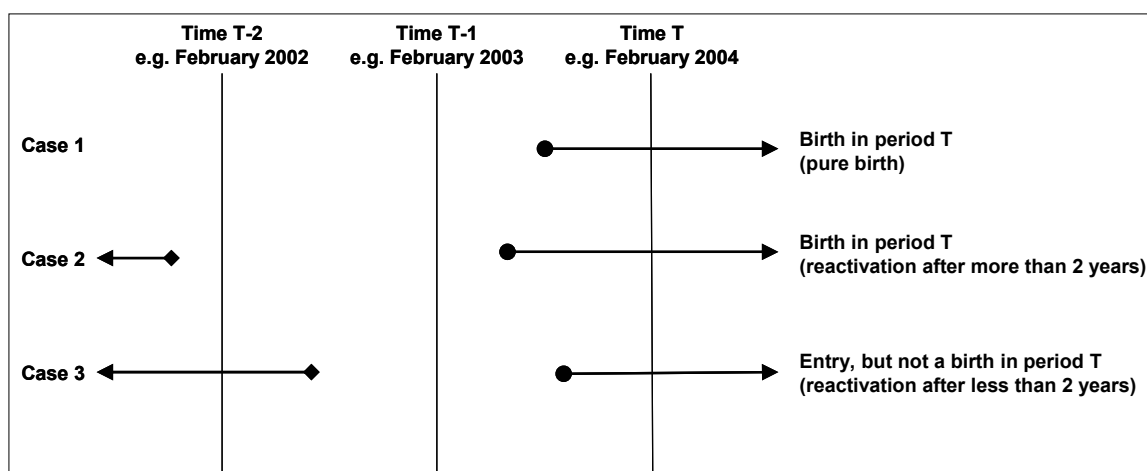
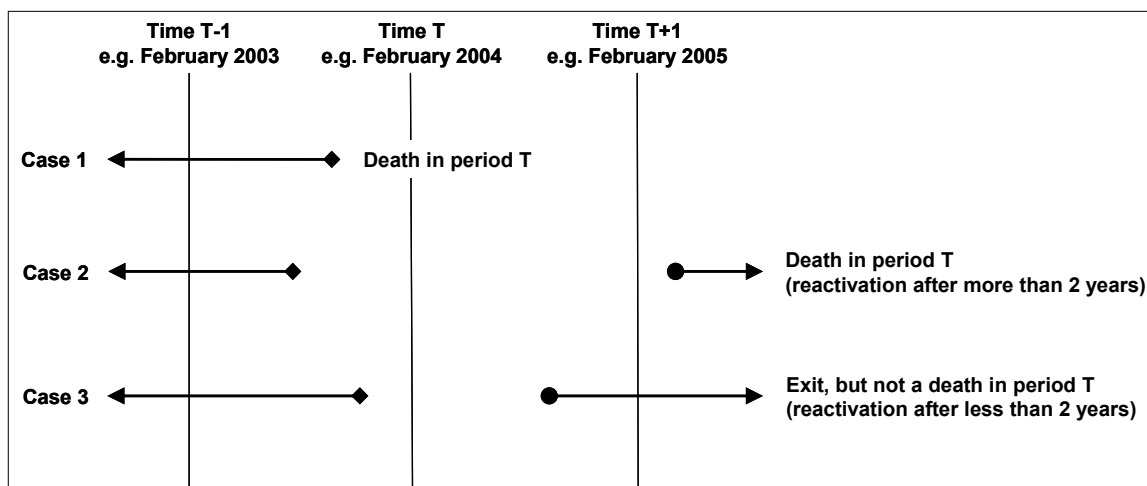


Figure 4

Graphical Representation of Death Identification



Proposed further classifications

Once real births have been identified on the LBF using the methods above, they can be analysed further. We suggest splitting real births of period T into:

- pure births (birth dates of all geographic units and the enterprise are more recent than the February snapshot of time T-2)
- other births (birth dates are not recent, therefore these are likely to be reactivations)

and also into:

- surviving births (survive at least one period until time T+1)
- short-lived births (disappear by time T+1, either due to death or dormancy).¹³

International comparability

The OECD study on international comparability of business start-up rates found that although enterprise birth rates are considered key economic indicators, their availability and definition varies considerably from country to country. Therefore, comparisons of birth or start-up rates between countries should be treated with caution. Eurostat and the OECD are currently working on standard models for business populations and standardised definitions for key indicators. However, it might take some time for countries to adopt such standards, as “data producers are often more influenced by national data requirements than international comparability” (Vale, 2006, p 6).

The definitions and methods set out in this section align well with the best practice models presented in the OECD study, especially with the standard model proposed therein (Vale, 2006, pp 93–100). With a better understanding of the various categories of enterprise entries (pure births, reactivations, other entries) it should also be possible to adapt the methods developed by Statistics New Zealand to international standards, once the latter are finalised.

3.4 Frequency of output

Current business demography methodology

Business demography statistics is an annual snapshot (as at February) of the structure and characteristics of New Zealand businesses, taken from the BF. Prior to 2003, the BF relied primarily on the annual BF update survey to maintain attributes on the BF. Since the introduction of the STEPS maintenance strategy to the BF in 2003, much more use can be made of tax data to update information. There are a number of discontinuities in the business demography series due to BF changes over time (for details see the Appendix), but annual data is available to users from 1987.

Options using the LBF

Data on the LBF is available at a monthly frequency back to April 1999, as this was the first month Inland Revenue collected EMS data.

The potential exists to create a business demography series more than once a year. There would be scope for Statistics New Zealand to investigate the feasibility of a more frequent series (eg quarterly) if there was sufficient user demand at any time in the future. Consultation with stakeholders in 2005 showed that users considered an annual output to be sufficient given current usage of the data.

¹³ Newly birthed multi-geo enterprises are classified as surviving if at least one geo survives, and short-lived if at least one geographic unit is short-lived (while the others are dying).

3.5 Business size measure

Current business demography methodology

The current indicator of business size (employment levels) in business demography statistics is the employee count (EC) of the February month. The EC recorded on the BF is sourced primarily from the EMS, which allows for the number of wage and salary earners to be derived.

There are a small number of enterprises whose employee count is collected by Statistics New Zealand surveys. These include group EMS filers, where one enterprise in the group is either employing employees of other members of a group, or is filing a return for some members of a group. In addition, there may be timing differences between the EMS data sourced from Inland Revenue and the information held on the BF. This can happen when a group of enterprises undergoes a restructuring and a number of business locations, along with their employees, are transferred between enterprises.

The current business size measure only includes people who are paid a salary or wage through the EMS, which comprises mostly employees. It does include working proprietors who pay themselves a salary or wage through the EMS; however, a large proportion of working proprietors choose to remunerate themselves through other means such as drawings or a shareholder salary. Therefore, a business size of zero does not necessarily indicate that there is no labour input into the enterprise.

Enterprises are divided into the following EC size groups:

- 0 employees
- 1–5 employees
- 6–9 employees
- 20–49 employees
- 50–99 employees
- 100 and more employees.

Since large enterprises are sparse in the New Zealand economy, the EC size group of 100 and more cannot be split further for most business demography statistics for reasons of confidentiality.

Business size measures on the LBF

The LBF records the employee count of each physical site of a business (geographic unit). For all months prior to August 2002, this is the sum of four variables: the number of male and female employees, each split by full-time or part-time employment. From August 2002 onwards the employee count is a single headcount of wage and salary earners that does not distinguish between gender and type of employment. Both measures are derived from the BF. For more information on the change from full-time equivalent persons engaged (FTE) to EC measures on the BF, see Appendix A.2.

The LBF holds an additional indicator for business size, which is derived from the LEED database. LEED measures total filled jobs for each enterprise and breaks the number down to each physical site of that enterprise. In addition, working proprietor data is expected to be added to LEED and the LBF in the middle of 2006.

The experimental business demography series from the LBF uses the LEED employee count to measure business size. The continuity of the measure from 1999

to date is seen as its major advantage. An option for further development is to introduce the rolling mean employee (RME) count derived from the LEED EC. The RME is the average EC over the previous 12 months for each monthly snapshot and therefore is less affected by seasonality effects than the EC of a single month.

3.6 Updates to business demography data

Current business demography methodology

As mentioned above, business demography data is extracted from the BF as at February of each reference period. However, the BF is not designed specifically to support statistical output with a single reference point (or period) in time. Rather, it is set up as a population source for economic and financial surveys, using the most up-to-date information at a given point in time. It is therefore a dynamic data source, which is continually updated with information obtained from survey feedback, administrative data, media reports, and so on.

Changes to attributes on the BF are not lost over time, but are recorded in BF history tables. Indeed, as explained earlier, the LBF uses these history tables to unwind the BF attribute variables in order to set up a back series to April 1999. However, the current business demography system does not have the functionality to update previously published statistics that may have changed since the last business demography population extraction from the BF. Updates are only published if significant changes or errors have been identified.

LBF development

The LBF is updated on a monthly basis, and data can be extracted from the LBF at a monthly frequency going back to April 1999. Each time the LBF is updated, changes to attributes on the BF are incorporated onto the LBF for all months that are affected, back to April 1999. This is possible because changes on the BF are recorded using a 'real world' date. Therefore, it is likely that there will be changes to previously published business demography data if the most recent information is extracted from the LBF for the next snapshot.

Monthly updates of business demography variables on the LBF affect around 20,000 to 30,000 geographic units and/or enterprises. Seventy percent of the changes relate to recent real world dates of the last five months. Extracting a snapshot for February of the current year with a lag of at least six months (in August) should therefore limit the updates to the series in the next year.

Consultation with stakeholders in June 2005 over the issue of whether updates should be made available as revisions to previously published data met with consensus that all revisions should be put through, no matter how small.

The experimental series was produced using a copy of the LBF database as of March 2006. By the time the official series based on the LBF goes into production, the LBF will have changed, so that the official statistics may differ from the experimental series due to updates that are processed in the meantime.

4. Presentation of Experimental Statistics

In the following section, references to LBF denote the experimental series for business demography from Statistics New Zealand's Longitudinal Business Frame, while BD denotes the official Business Demography series as published in November 2005.¹⁴ In the graphs, agriculture is abbreviated as 'ag'.

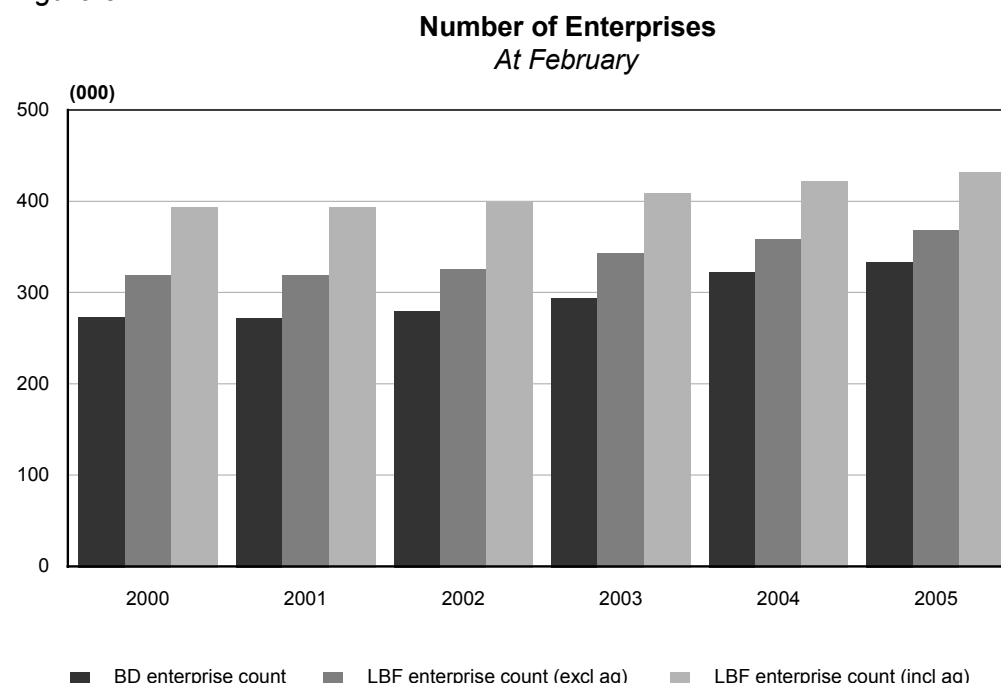
4.1 Stock of active enterprises and geographic units

Total number of enterprises and employees

When comparing the full coverage LBF population with the current BD population it is important to note that the former includes all economically significant enterprise units. To ensure comparability with previous periods, the BD continued to exclude agriculture (A01) units even after they became available on the Business Frame (BF), as well as units that would have entered the population due solely to BF maintenance changes (ie Business Surveys Tax Expanded Population Strategy (STEPS) changes on the BF – see the discussion in Appendix A.1 for details). Primarily, the historical data available on the LBF allows the construction of a series from 2000 onwards that is less susceptible to change as a result of BF maintenance. It also allows the consistent inclusion of the agriculture industry.

Around 25 to 30 percent of the LBF population are not included in the BD population due to the current BD conditions for population selection (see section 3.1). Although the BD series was designed to be comparable over time by excluding units as set out above, the BD growth rates vary more widely over the years than the comparative rates of the LBF (see Figures 7 and 8 below).

Figure 5



¹⁴ For release information see Statistics New Zealand's website at: <http://www.stats.govt.nz/products-and-services/info-releases/bus-demo-stats-info-releases.htm>

In comparison to the Linked Employer-Employee Data (LEED) employee count on the LBF, about 2 to 6 percent of the employee count (EC) has been underestimated by the BD methodology. Compared with the BD, the LBF annual growth for total employment count appears smoother over time. Additionally, the LBF total employee growth trend in general moves together with the growth trend of its total number of enterprises. The BD series by comparison shows more volatility, and often the BD's total enterprise number and total employee count move against each other (see Figures 7 and 8 below).

Figure 6

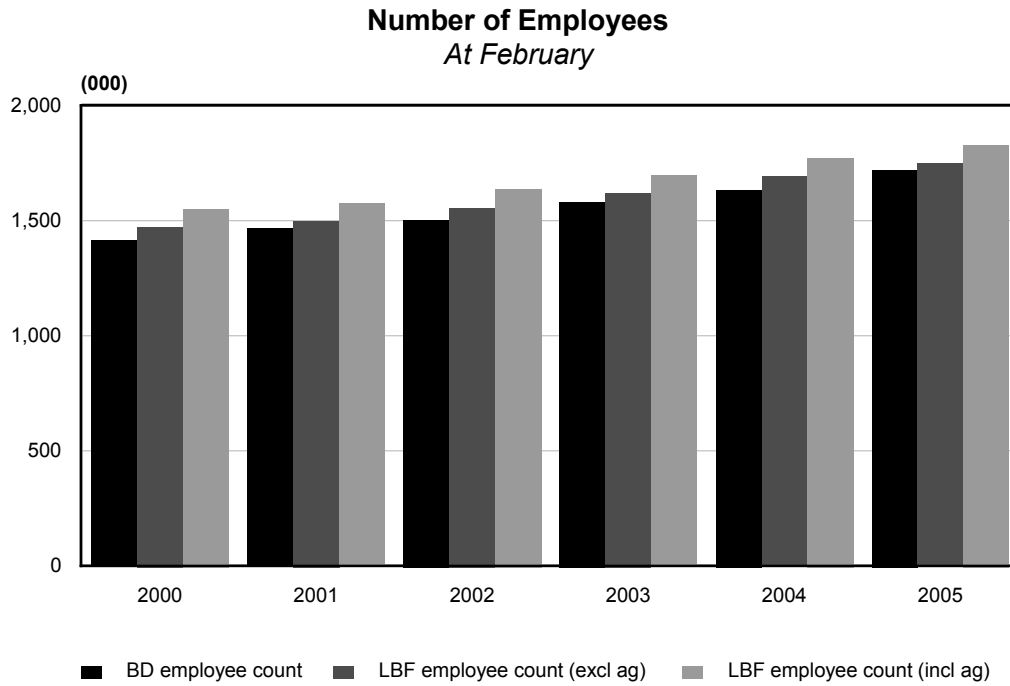


Figure 7

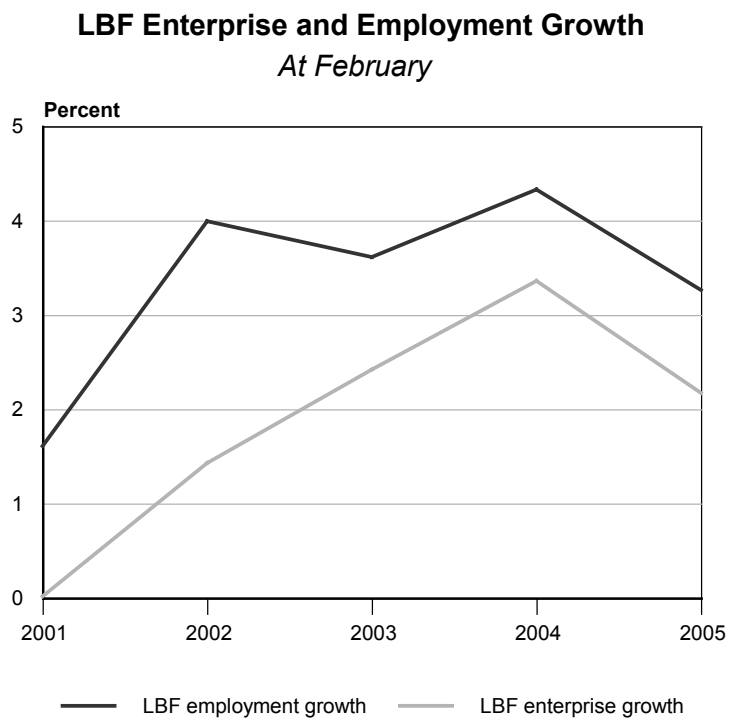
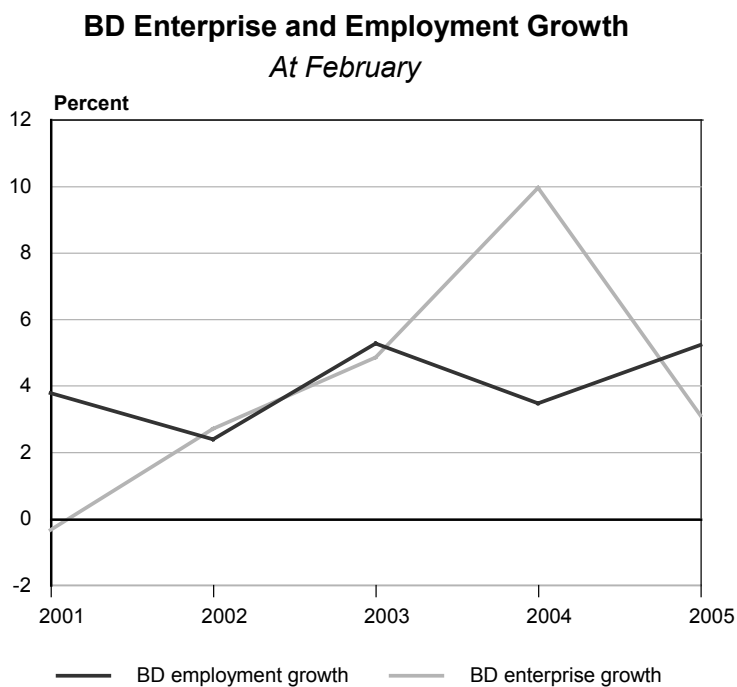


Figure 8



Annual percentage changes for both the LBF number of enterprises and number of employees by EC size groups show positive growth. In 2005, the strongest growth came from businesses with size groups of 10 to 19 employees, or more than 100 employees.

Table 2

Annual Percentage Change in Enterprise Numbers from LBF
Including agriculture
By EC size group

Business EC Size Group	Year				
	2001	2002	2003	2004	2005
0	-0.5%	0.5%	2.0%	3.5%	2.0%
1-5	1.0%	2.6%	2.9%	2.7%	2.4%
6-9	1.6%	5.5%	3.7%	2.9%	1.3%
10-19	1.2%	3.8%	5.4%	3.7%	4.6%
20-49	1.4%	5.6%	3.0%	5.4%	0.2%
50-99	1.3%	3.8%	2.7%	6.2%	3.6%
100+	2.6%	3.8%	2.3%	3.5%	6.0%
Total	0.0%	1.4%	2.4%	3.4%	2.2%

Table 3

Annual Percentage Change in Employment Count from LBF
Including agriculture
By EC size group

Business EC Size Group	Year				
	2001	2002	2003	2004	2005
0					
1-5	1.4%	3.2%	2.7%	2.8%	2.6%
6-9	1.3%	5.7%	3.7%	2.9%	1.5%
10-19	1.4%	3.7%	5.5%	3.5%	4.9%
20-49	1.6%	4.8%	2.6%	5.0%	0.6%
50-99	0.6%	4.3%	3.4%	6.5%	3.6%
100+	2.0%	3.6%	3.7%	4.6%	4.1%
Total	1.6%	4.0%	3.6%	4.3%	3.3%

The improvement of BF maintenance from administrative sources (STEPS – see section 3.1 and Appendix A.1) has seen the incorporation of enterprise units exempt from Goods and Services Tax (GST) that are reporting annual IR10 returns exceeding \$40,000, reactivation of units that had previously been ceased, and the inclusion of all employing units. These additions are included in the LBF population, but not accounted for in the BD.

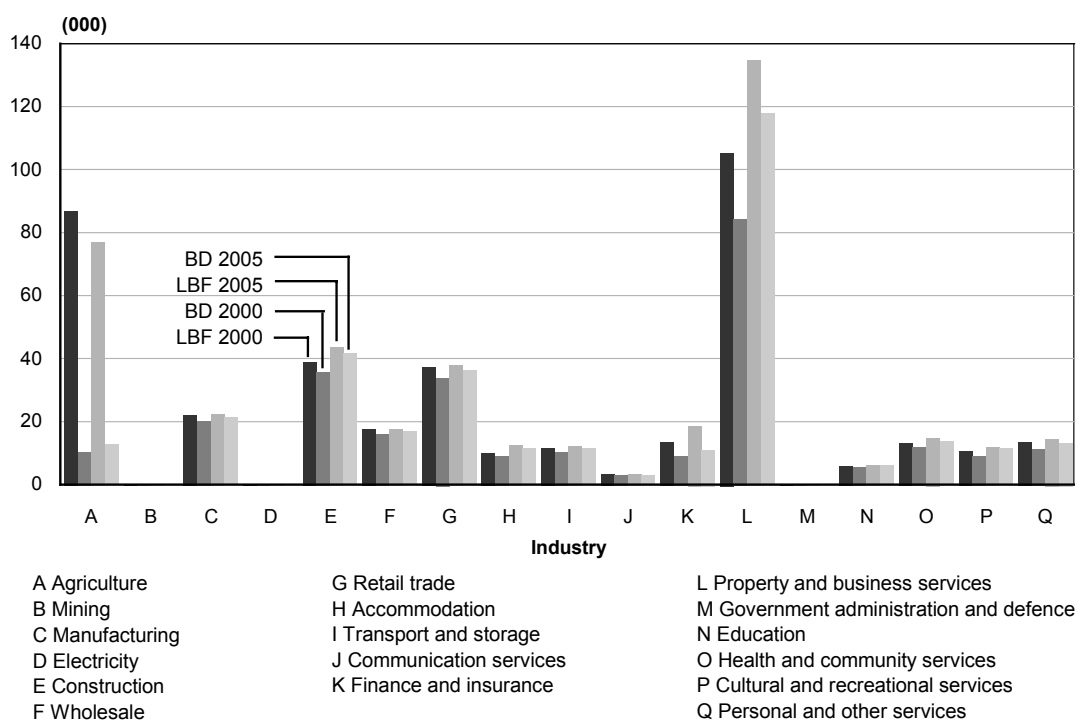
Industry

Figure 9 shows the key advantage of full industrial coverage on the LBF by comparing the LBF and BD snapshots by industry at February 2000 and 2005. After property and business services, the agriculture industry is the second most dominant industry on the LBF, accounting for around 20 percent of all economically significant enterprises in New Zealand for both periods. In contrast, the non-farming series of the BD only includes agricultural enterprises other than farming (eg services to agriculture, forestry), which contributed only about 4 percent of enterprises in each period.

Since 2003, additional information from administrative sources is also being used to administer the ongoing improvement of industrial classification on the BF to better reflect the real-world industrial activities of enterprises and geographic units. On the LBF, the most up-to-date information on industry is available, whereas updates of previously published BD statistics are not feasible (see section 3.6).

Figure 9

Number of Enterprises by Industry
At February 2005 compared with February 2000



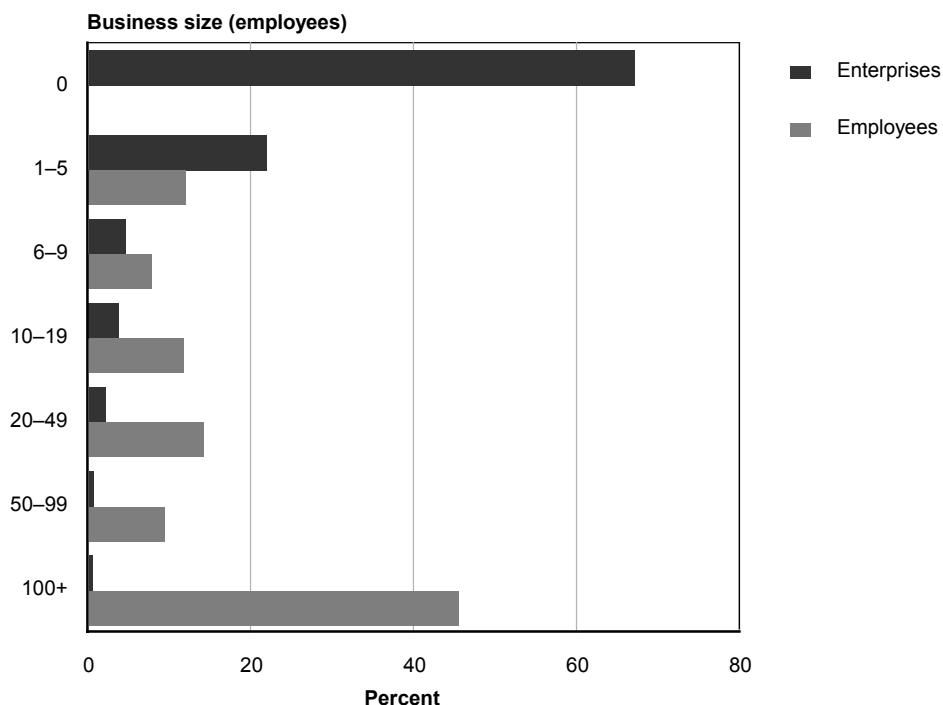
Business size

As shown by current BD statistics, most economically significant enterprises (97 percent) on the LBF in 2005 have fewer than 19 employees. However, these enterprises account for only 31 percent of all employees. Conversely, enterprises with more than 100 employees make up only 0.5 percent of the total number of enterprises in New Zealand, but employ 45 percent of the total number of employees.

From 2000 to 2005, non-employing enterprises consistently contributed over 60 percent to the total number of economically significant enterprises in New Zealand. The LBF has yet to include working proprietors – they are only included in the employment count on the LBF if they pay themselves a wage or salary. Therefore, care should be taken when interpreting results for non-employing enterprises.

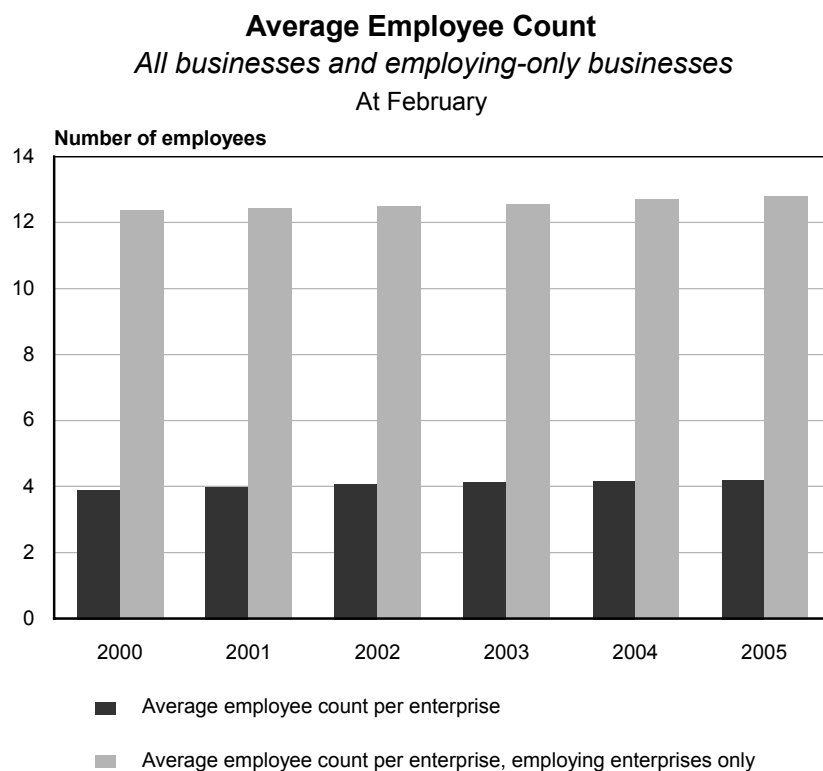
Figure 10

Proportion of Enterprises and Employees by Business Size Groups
At February 2005



The average business size for all economically significant enterprises in New Zealand is about four employees. When non-employing enterprises are excluded, the average business size is about 12 employees, approaching 13 towards the end of the series.

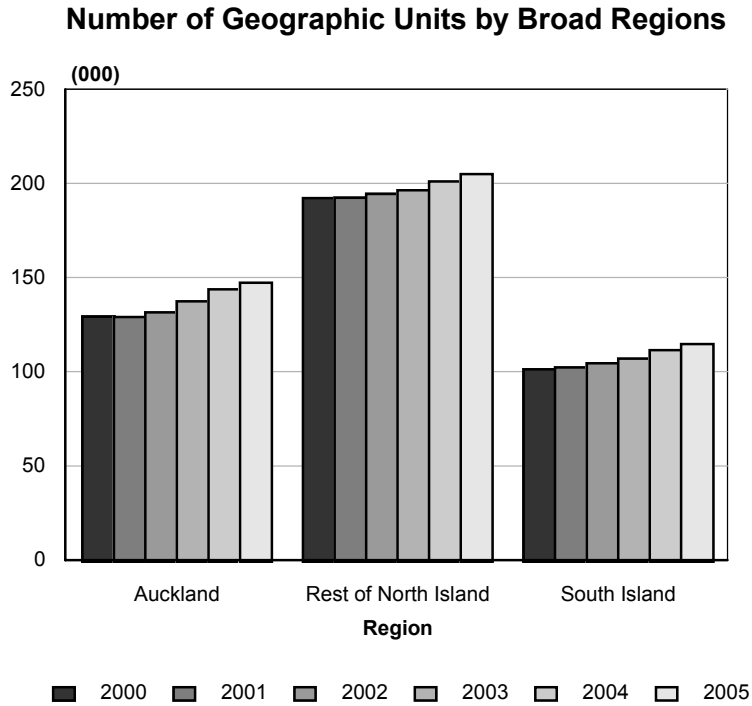
Figure 11



Region

As shown by current BD statistics, about one-third of all LBF geographic units (business locations) are located in Auckland. In fact, three-quarters of all geographic units are located in the North Island. The South Island hosts one-quarter of all geographic units.

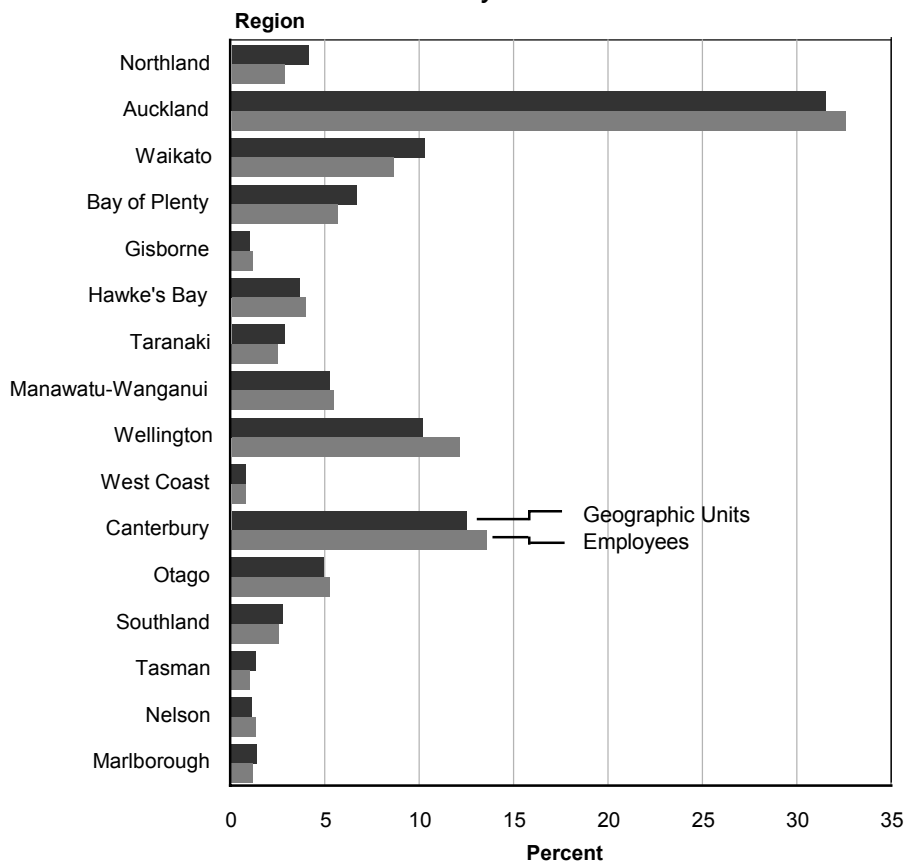
Figure 12



The distribution of businesses and the number of employees across regional councils as seen on the LBF reflects that for the BD. Businesses located in Auckland engage one-third of all employees in New Zealand. Auckland is followed by Canterbury (12 percent of enterprises, 14 percent of employees, respectively), Wellington (10 percent, 12 percent, respectively) and Waikato (10 percent, 9 percent, respectively).

Figure 13

Proportion of Geographic Units and Employees by Regional Council
At February 2005



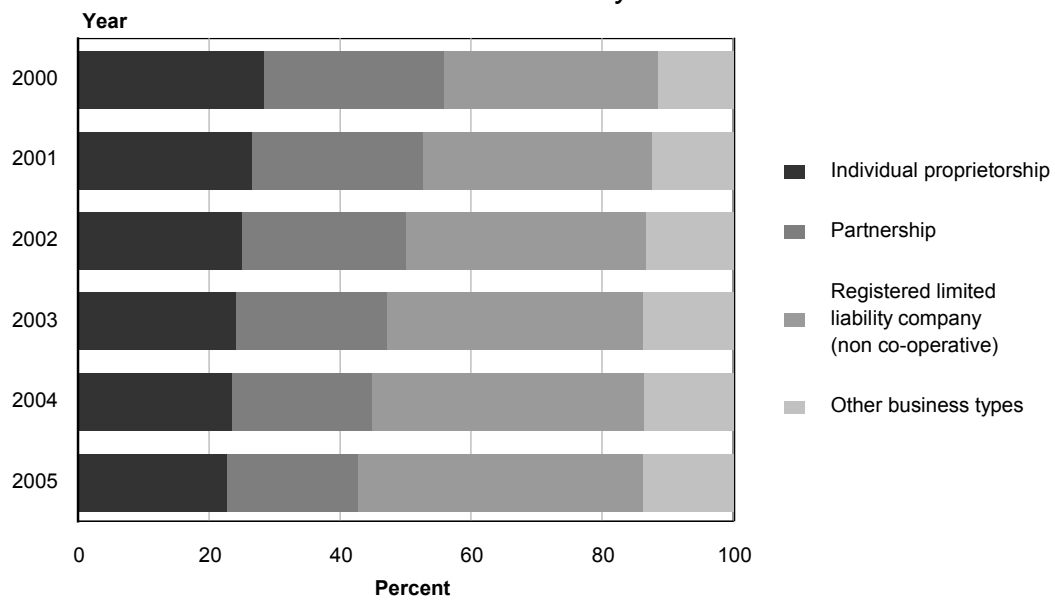
Business type and institutional sector

Of all economically significant enterprises in New Zealand in 2005, 40 percent are registered limited liability companies, followed by individual proprietors or sole traders, and partnership business types (both at 20 percent). Registered limited liability companies employ over 60 percent of the total number of employees in New Zealand, when not counting working proprietors (proportions of total employees are not shown in Figure 14).

From 2000 to 2005, the share of registered limited liability companies in the enterprise population has risen continually, while the shares of both individual proprietors and partnerships have declined.

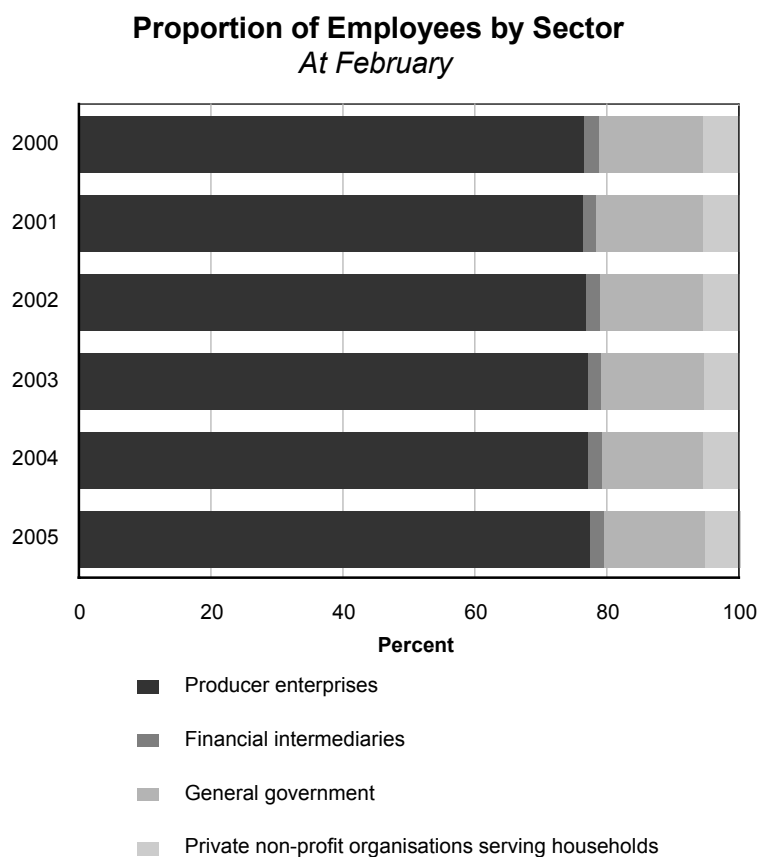
Figure 14

Proportion of Total Enterprises by Business Type
At February



Ninety-five percent of all economically significant enterprises are in the producer enterprises sector, which includes private producer enterprises, producer boards and government enterprises. The sector employs over three-quarters of the total number of employees in New Zealand. Sixteen percent are employed in the general government sector, although the sector contributes only 1 percent to the total number of enterprises (proportions of enterprises are not shown in Figure 15).

Figure 15



4.2 Enterprise births

With the new birth methodology, the annual birth rate calculated from identified births on the LBF lies between 12 and 13 percent for the period February 2000 to February 2005. This is a marked drop from the BD entry rate of 17 to 24 percent. The annual death rate, between 10 and 12 percent on the LBF, is also lower than the BD exit rate of 14 to 18 percent. The resulting business turnover rate¹⁵ of entering and exiting businesses throughout the economy, at February of each year, decreased from between 32 and 40 percent on the BD to 22–23 percent on the LBF. This means that, according to the LBF, not more than 23 percent of economically significant enterprises in the New Zealand economy are affected by business dynamics (births and deaths). This leaves at least 77 percent of enterprises on the LBF that are stable in the economy.

Table 4

Overview of BD Population, Entries and Exits

	Year				
	2001	2002	2003	2004	2005
BD total population	273,961	281,338	294,954	324,293	334,340
BD entries	47,324	50,522	53,972	67,559	58,144
BD exits	48,280	43,145	40,356	38,220	48,097
BD entry rate	17%	18%	18%	21%	17%
BD exit rate	18%	15%	14%	12%	14%
BD turnover rate	35%	33%	32%	33%	32%

Table 5

Overview of LBF Population, Births and Deaths

	Year				
	2001	2002	2003	2004	2005
LBF total population	394,766	400,409	410,099	423,877	433,089
LBF births	46,896	46,272	52,053	56,482	51,934
LBF deaths	45,473	40,662	42,462	42,396	42,283
LBF birth rate	12%	12%	13%	13%	12%
LBF death rate	12%	10%	10%	10%	10%
LBF turnover rate	23%	22%	23%	23%	22%

Table 6

LBF Births by Subcategories

	Year				
	2001	2002	2003	2004	2005
LBF births	46,896	46,272	52,053	56,482	51,934
LBF surviving births	38,735	38,765	43,376	47,765	n/a
LBF short-lived births	8,161	7,507	8,677	8,717	n/a
LBF pure births	40,450	39,383	43,688	47,072	43,338
LBF surviving pure births	33,698	33,122	36,539	40,317	n/a
LBF short-lived pure births	6,752	6,261	7,149	6,755	n/a
LBF pure birth rate	10%	10%	11%	11%	10%

¹⁵ The turnover rate is defined as the sum of the entry/birth rate and the exit/death rate.

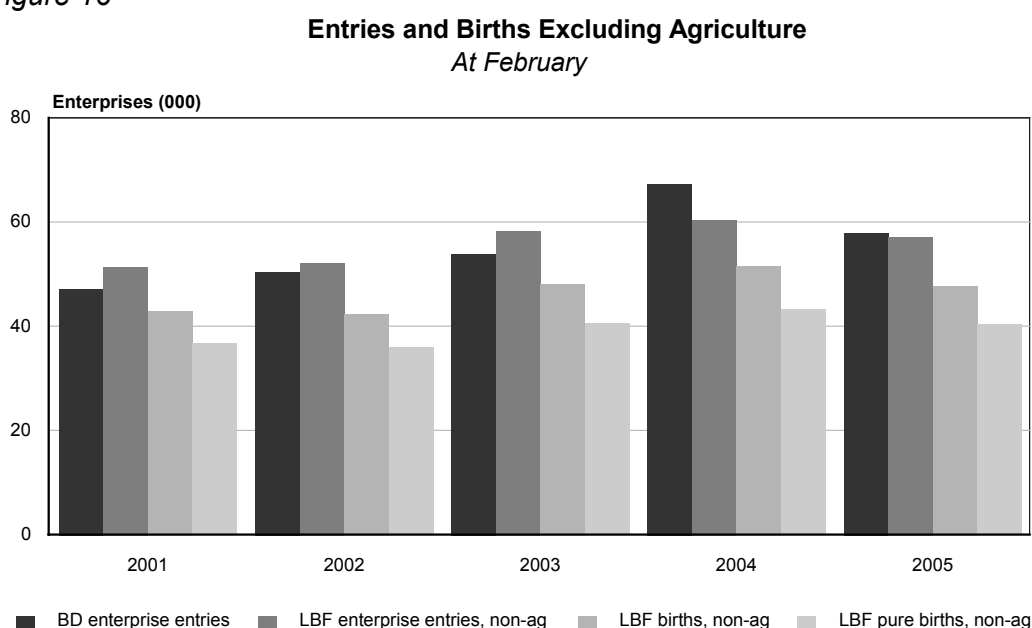
Number of births and birth rates

Figures 16 to 18 show the numbers of enterprises birthed on the LBF – excluding and including agriculture – compared with enterprise entries in the BD. LBF births are consistently lower than BD entries, even when agriculture is included.

LBF entries are the number of new enterprise identifiers appearing in a snapshot when compared with the previous one – that is, defined the same way as BD entries. Although the BD population is smaller than that from the LBF, the BD entry rates exceed LBF entry rates in 2004 and 2005. The increase of entries coincides with the adjustments made to the BF in the course of the STEPS maintenance change, which involved the addition of units to the BF from July 2003 to March 2006. Despite the methodology applied to exclude these units from the BD series to make it more comparable with previous periods, BD is a ‘live’ extraction from the BF. The timing of the extraction has the potential to impact the BD. The example also shows that it can be misleading to rely on administrative identifiers for the identification of business dynamics, especially at times when the administrative data source undergoes change.

The lower number of births on the LBF is reflected in the comparison of BD entry rates with LBF birth rates. The LBF birth rates align with international results from Europe, the United Kingdom, Canada, the United States and Australia, where birth rates lie between 9 and 14 percent, dependent on the population and methods applied (Vale, 2006, pp 7 and 31-35). Moreover, they are consistent with rates from independent research conducted on firm entry and exit in New Zealand.¹⁶

Figure 16



16 A paper by The Treasury on entry and exit rates for 1995 to 2003, based on data from the BF, constructed an average national entry rate of 12 percent, and an average turnover rate of 20 percent. See Law D and McLellan N (2005). “The Contributions from Firm Entry, Exit and Continuation to Labour Productivity growth in New Zealand”, New Zealand Treasury Working Paper 05/01, Wellington, pp 2–5.

Figure 17

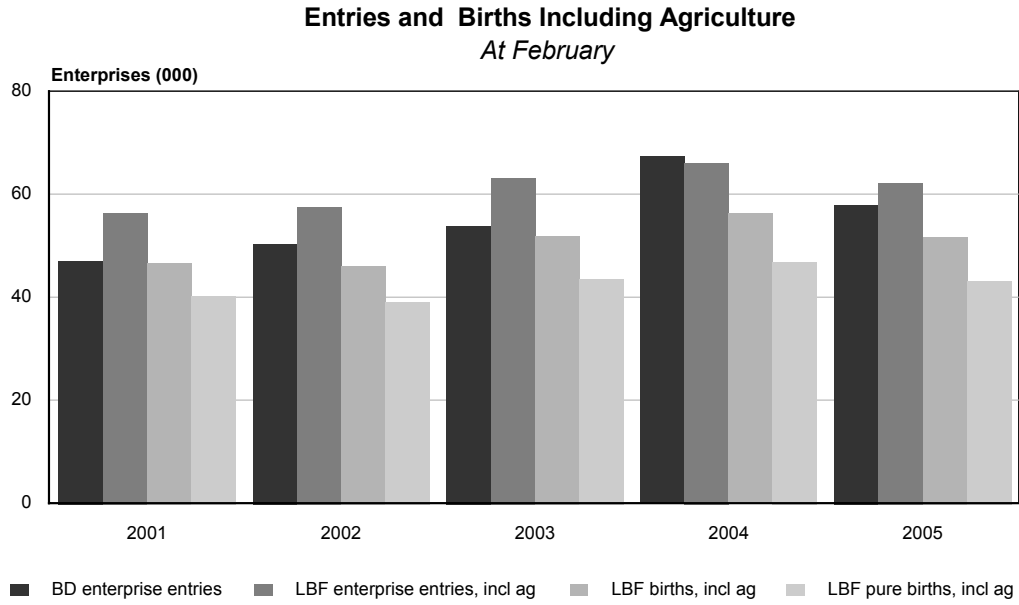
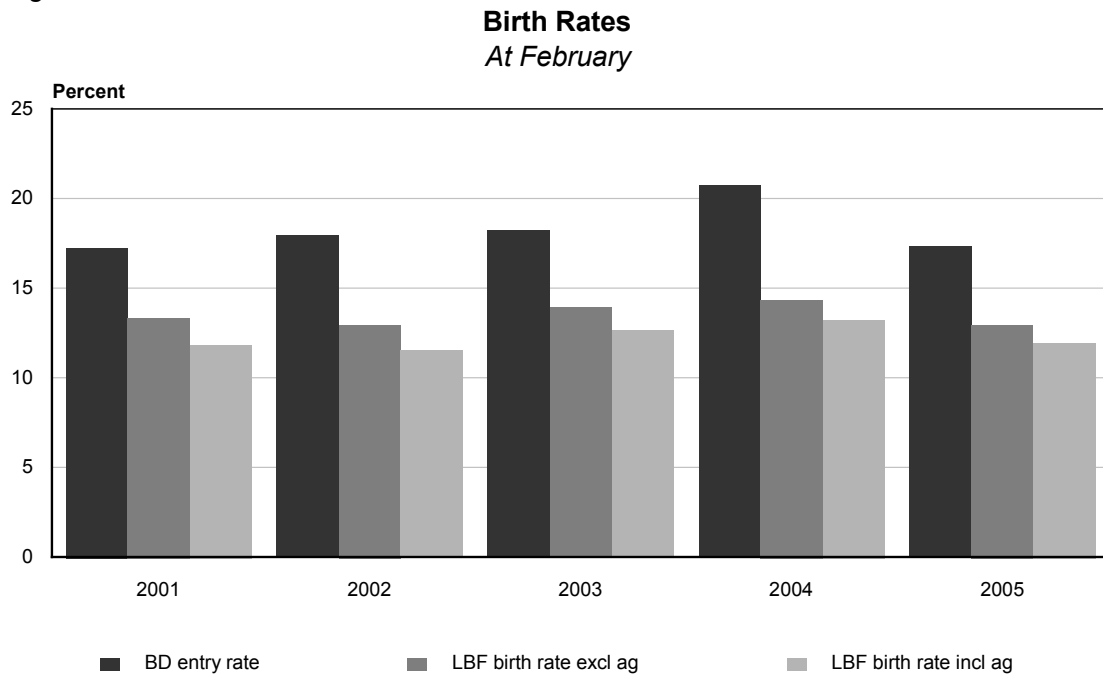


Figure 18



When calculating and comparing birth rates, it is essential to be clear about the underlying definitions, as the notion of what constitutes an enterprise birth varies from country to country. There is ongoing discussion on whether to count reactivations after more than two years as births at all, and whether to exclude from the count short-lived births that don't survive for at least one year. In the LBF series:

- Pure births are identified births that have a recent birth date. They make up 83 to 86 percent of total identified births.
- The majority of births not accounted for under pure births ('other births') are likely to be reactivations – further analysis is necessary to confirm this.
- Short-lived births, which disappear again in the period after their birth due to death or dormancy, make up 15 to 17 percent of births, both as a percentage of all births, and of pure births only.

An assessment of the survival of births is not possible for the last available snapshot.

Figure 19

Identified LBF Births by Subcategory
At February

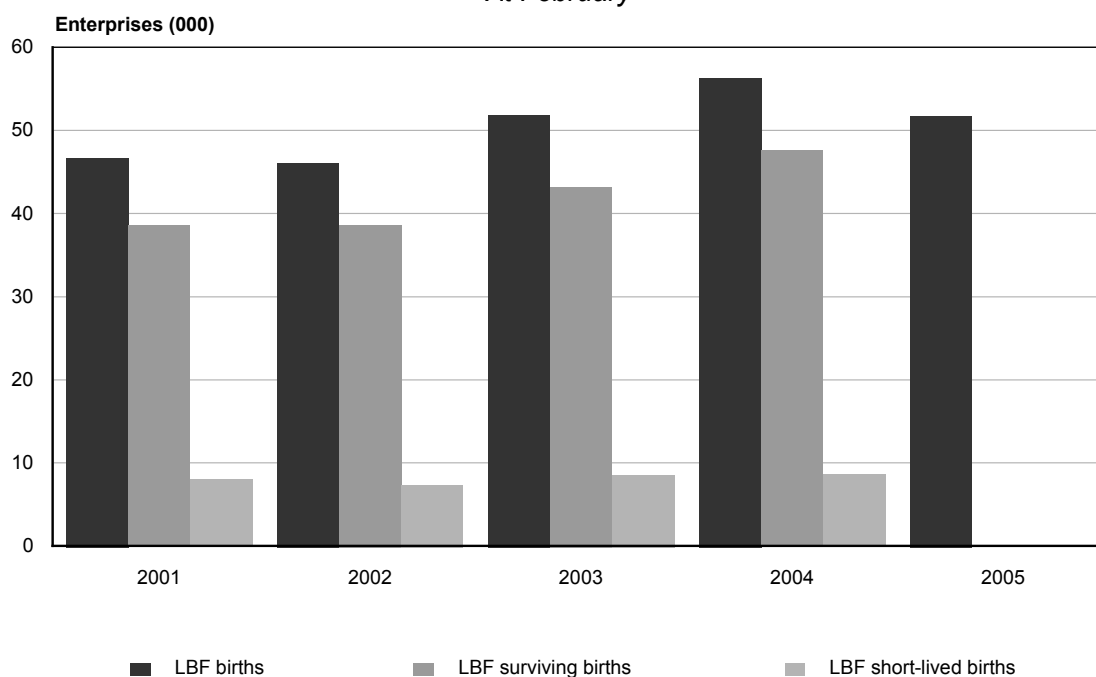
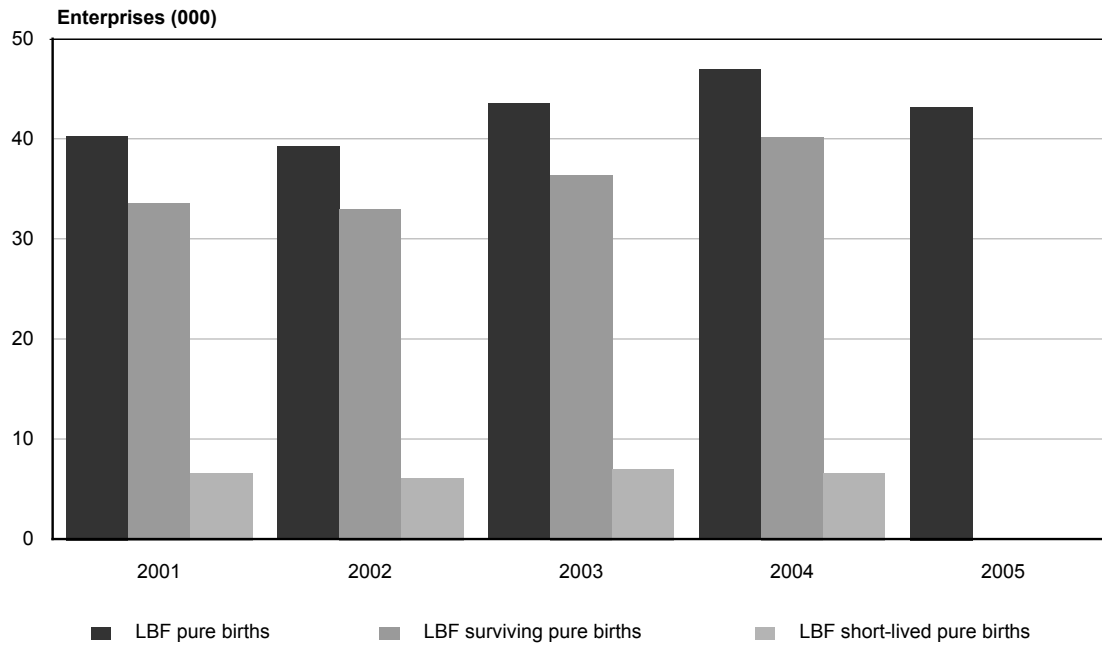


Figure 20

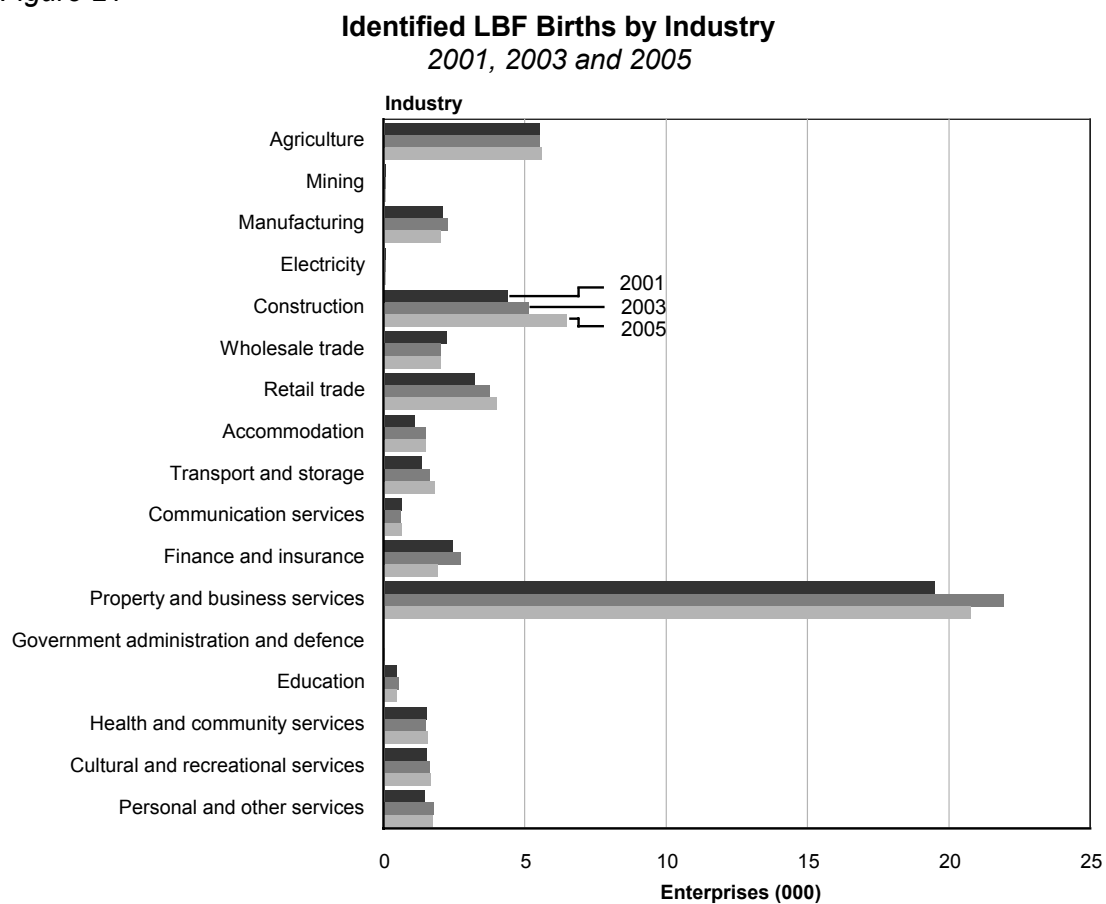
Identified LBF Pure Births by Subcategory
At February



Industry

Birthed enterprises are most likely to be active in the property and business services industry (40 percent), followed by the agriculture and construction industries, both with approximately 10 percent of all births. The government administration and defence sector is the least likely industry for businesses to be birthed into.

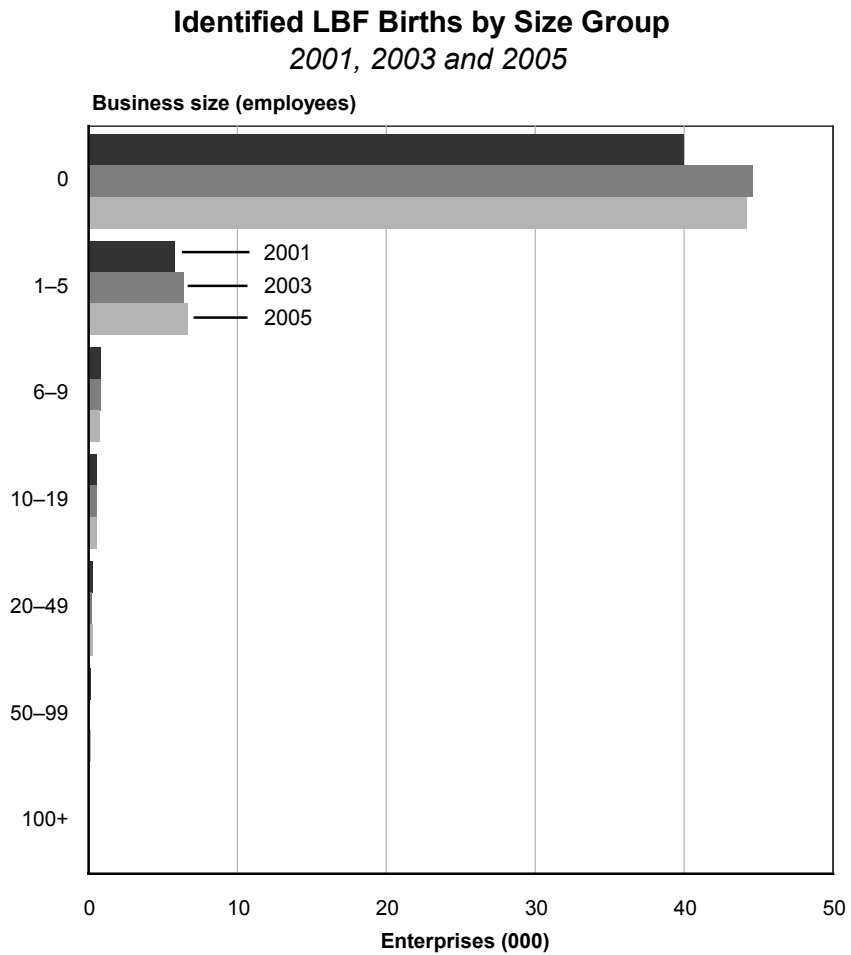
Figure 21



Business size

On the LBF, the majority of enterprise births are non-employing businesses (85 percent), followed by businesses that employ one to six employees (12 percent).¹⁷ Most births fall into the category of small and medium enterprises, which employ fewer than 19 employees (99.6 percent). Conversely, enterprise births are very unlikely to be businesses that employ more than 20 employees (less than one-half of a percent).

Figure 22



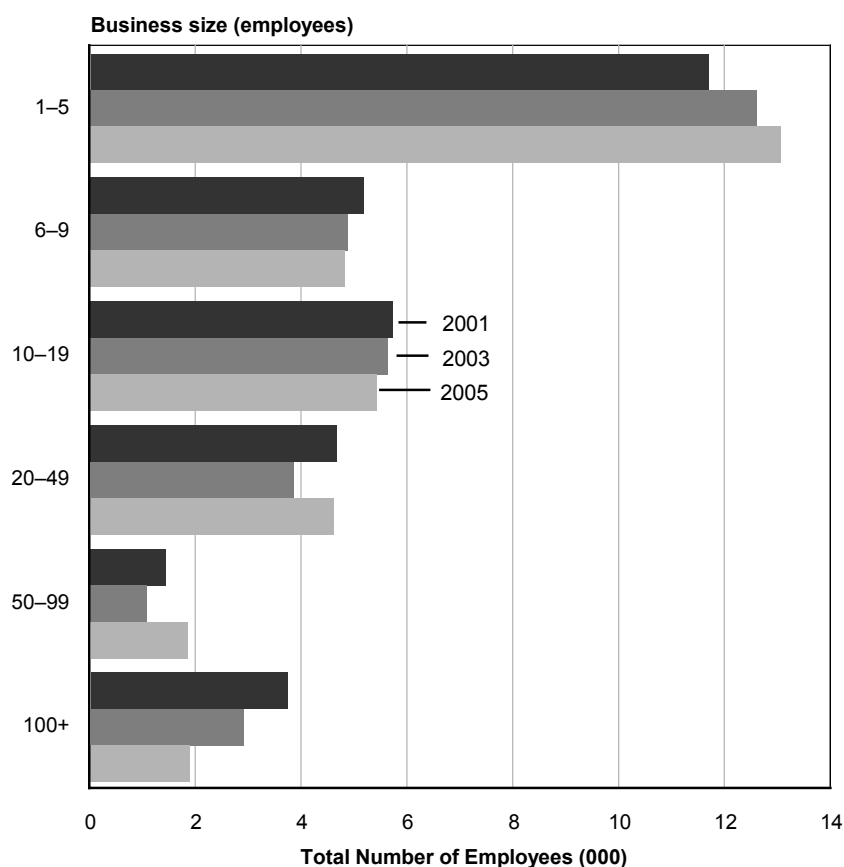
¹⁷ For statistics by business size, the size of births is the employee count at the time of the first snapshot that includes them.

Employment in enterprise births is most likely to come from businesses that employ fewer than 19 employees (70 percent). Forty percent of the employment count of new births comes from businesses with one to five employees. No implications, however, should be drawn about the further growth of births in these size groups. The size changes of new enterprises over the years after their birth merit a separate analysis, which is outside the scope of this report.

Figure 23

Employees of LBF Births by Size Group

2001, 2003 and 2005



For the average size of enterprise births and deaths, see the combined graph of Figure 30 below.

Statistics for births or deaths by region have not been produced for this report. While enterprises with only one geographical unit have a clearly defined region, this attribute is not as easily assigned to enterprises with multiple locations. It is an issue for discussion how to define a 'predominant region' for an enterprise.

4.3 Enterprise deaths

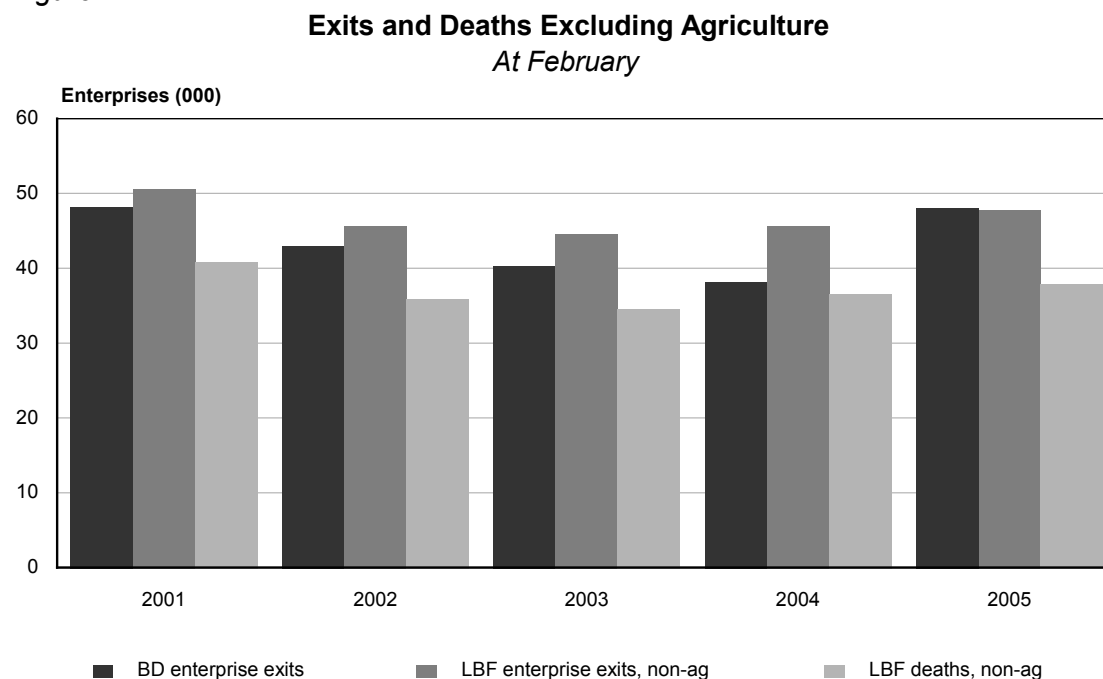
Number of deaths and death rates

With the new methodology to identify deaths, the number of enterprise deaths on the LBF (both excluding and including agriculture) is less than the number of BD exits. Exits for both the BD and LBF are derived on the basis of the current BD exit methodology – that is, by comparing enterprise identifiers from one snapshot to the other. Exits include enterprises that do not actually die, but change their administrative identifiers. Exits on the LBF (excluding agriculture) are higher or equal to those of the BD, reflecting the larger population. The use of repaired longitudinal links on the LBF in conjunction with the new methodology to identify deaths eliminates false deaths of about 15 percent of BD exits and 20 percent of LBF exits.

When studying the time series for deaths, it should be kept in mind that 2005 results for enterprise deaths are preliminary until the snapshot for 2006 is available.

From 2000 to 2005, the death rates on the LBF ‘including agriculture’ were in a range of 9.8 to 11.5 percent¹⁸ and significantly lower than BD exit rates.

Figure 24



¹⁸ In the working paper by The Treasury (see footnote 16), the average national exit rate for 1995–2003 was 8.2 percent (Law and McLellan, 2005, p 5).

Figure 25

Exits and Deaths Including Agriculture
At February

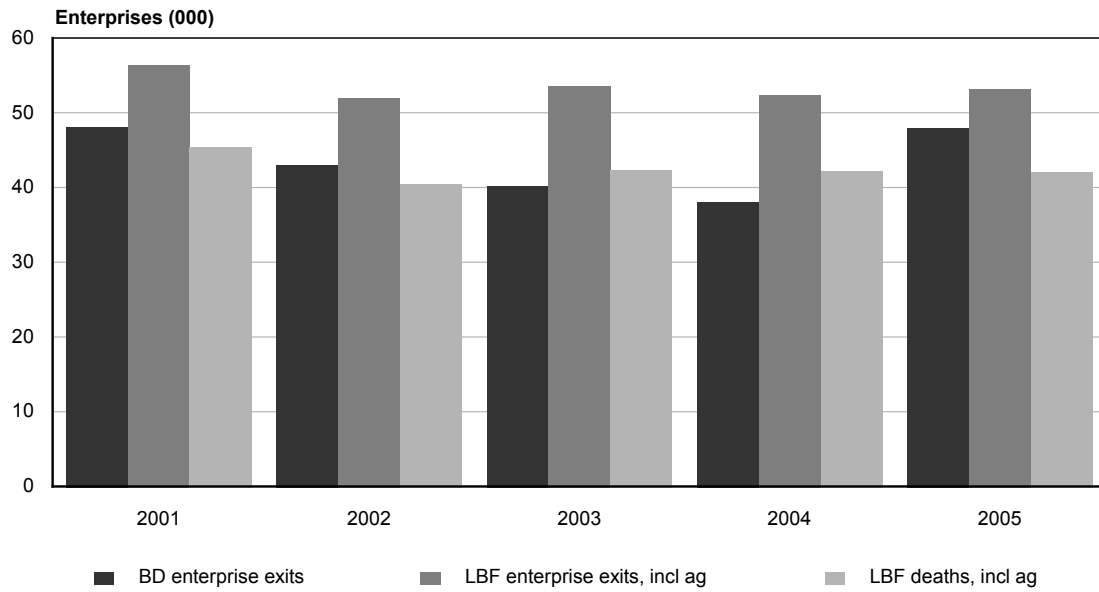
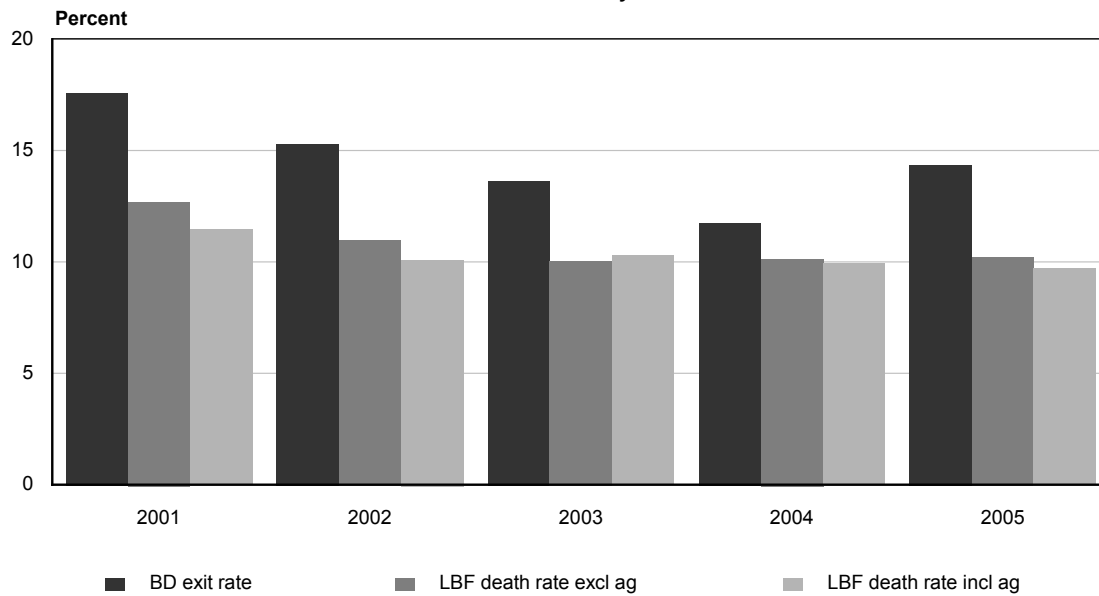


Figure 26

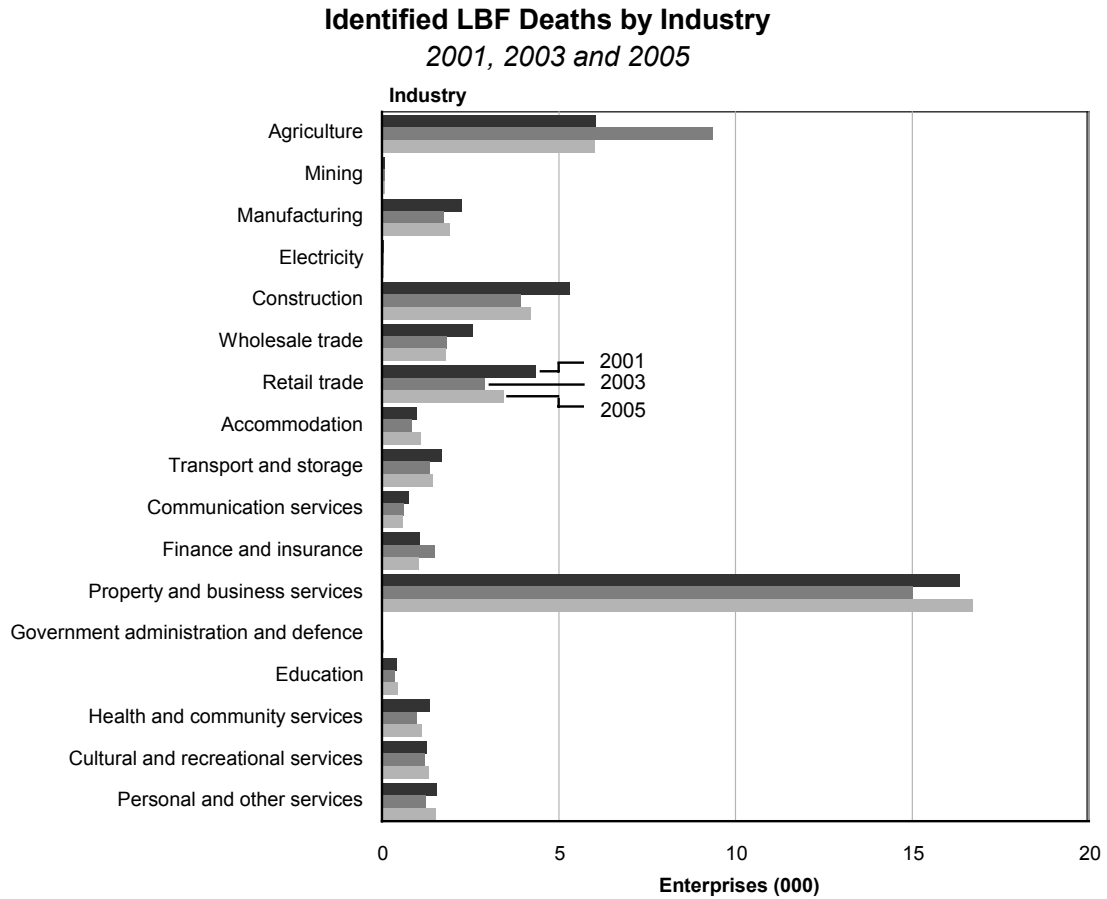
Death Rates
At February



Industry

Enterprise deaths on the LBF are most likely to originate from the business and property services industry (35–39 percent), followed by agriculture (13–22 percent), construction (9–12 percent) and retail (7–9 percent).

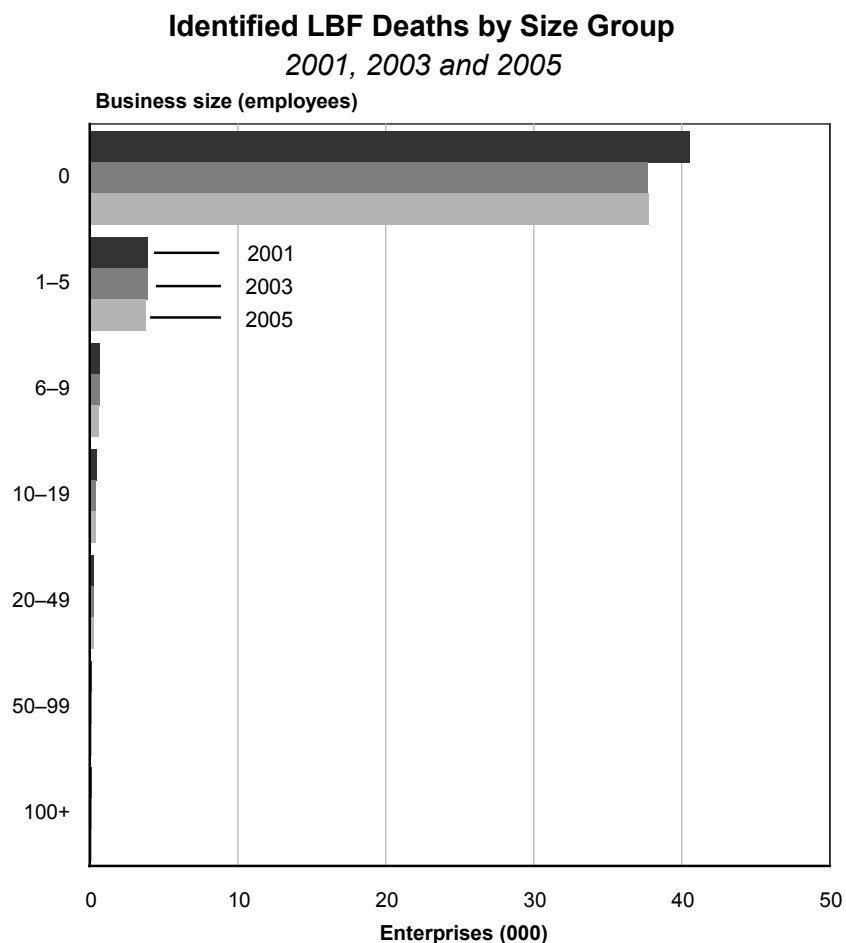
Figure 27



Business size

On the LBF, non-employing businesses are the most likely businesses to die (88 percent). Most deaths fall into the category of small and medium enterprises, which employ fewer than 19 employees (99.6 percent).¹⁹ Conversely, enterprise deaths of businesses that employ more than 20 employees are uncommon (less than one-half of a percent). It should be kept in mind that enterprises may reduce their staff numbers over a period before their actual death, and therefore may appear as deaths in a lower EC size group than the group they belonged to during their productive period.

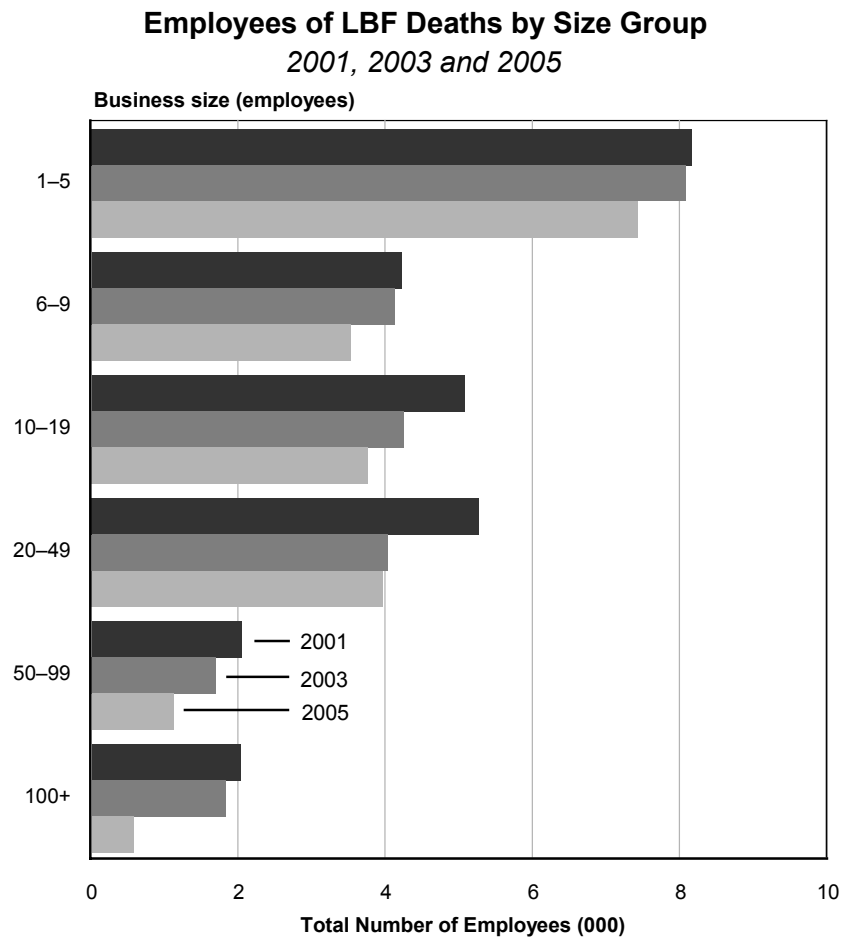
Figure 28



¹⁹ For statistics by business size, the size of deaths is measured at the time of the last snapshot that includes them.

About two-thirds of the loss of employment as a result of enterprise deaths comes from businesses with fewer than 19 employees, with one-third coming from businesses with fewer than five employees.

Figure 29

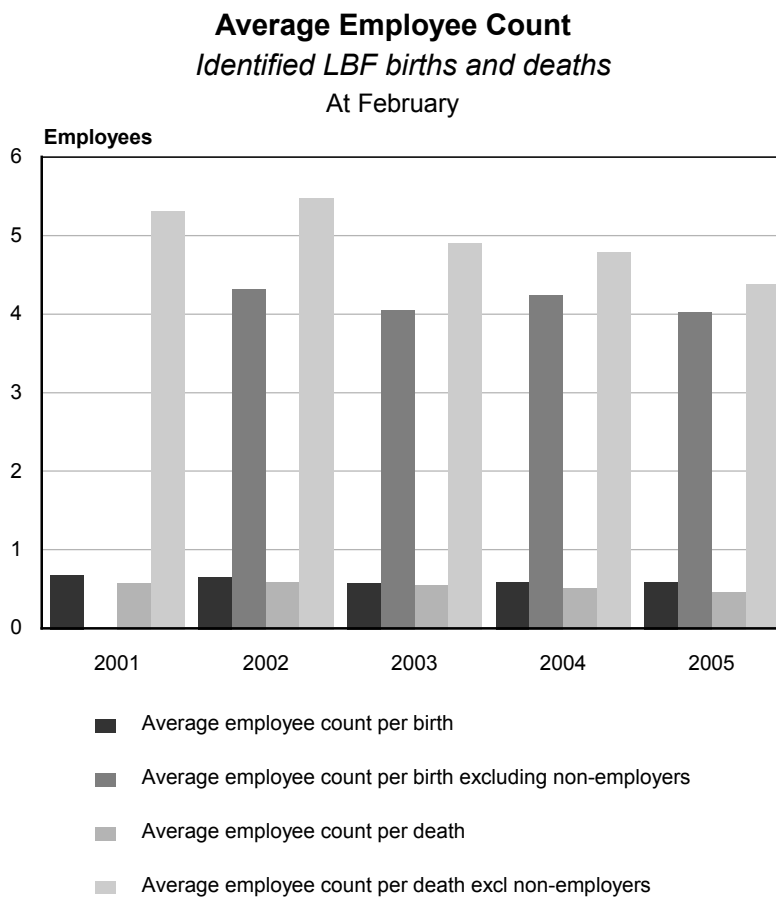


When non-employed enterprises are excluded, the average number of employees engaged by *birthed* enterprises lies at around four employees per enterprise birthed. In contrast, the average number of employees per enterprise death is slightly higher at 4.5 to 5.5 employees.

From 2001 to 2005, the overall average net gain in employees per net additions to enterprises (births minus deaths) lies between 0.7 and 3.9 employees²⁰ (not shown in graph).

²⁰ Calculated as: (employees gained from births – employees lost from deaths) / (enterprises gained from births – enterprises lost from deaths).

Figure 30



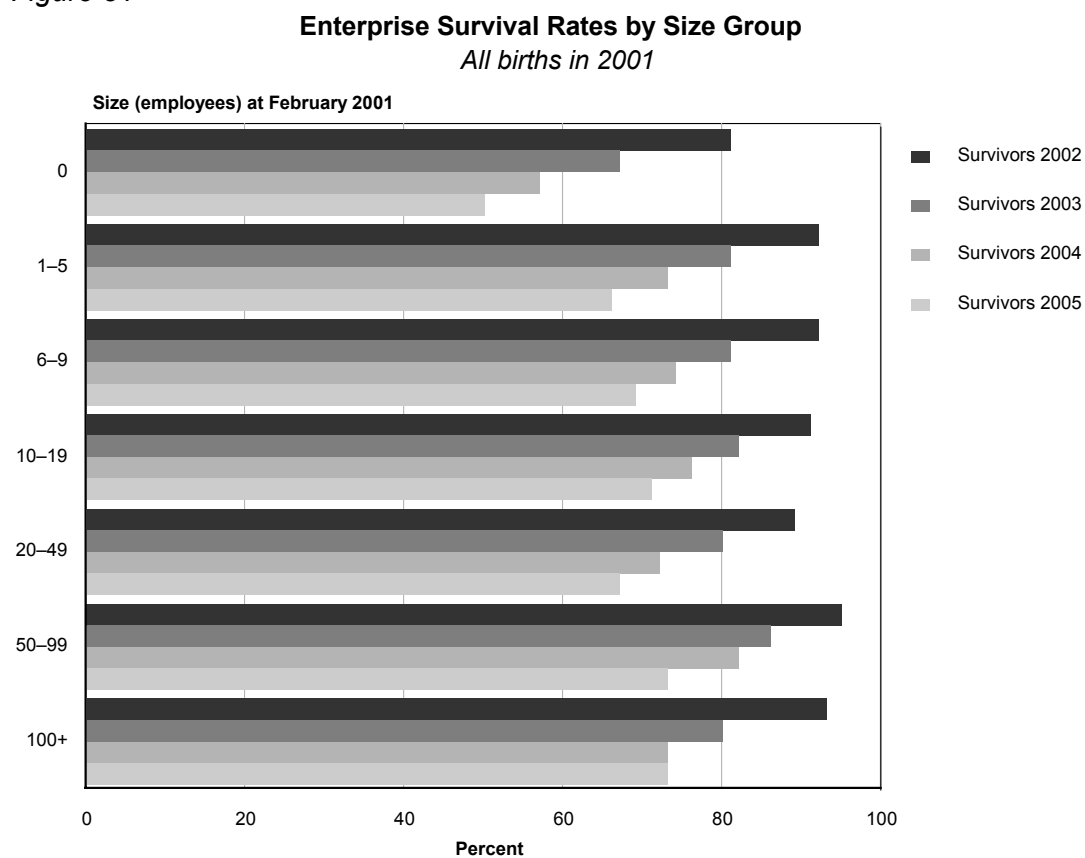
4.4 Survival and transitions

Survival of enterprise births

Survival rates measure the probability that new enterprises will live beyond a given time. For enterprises birthed in 2001 with a business size lower than 19 employees, the chance to survive up to February 2005 improved with growing business size, and was highest for births with 10 to 19 employees (71 percent survived from 2001 to 2005). Non-employing enterprise births had the lowest chance of surviving for four years (less than 50 percent), while those with more than 50 employees had a 73 percent chance of surviving until 2005.

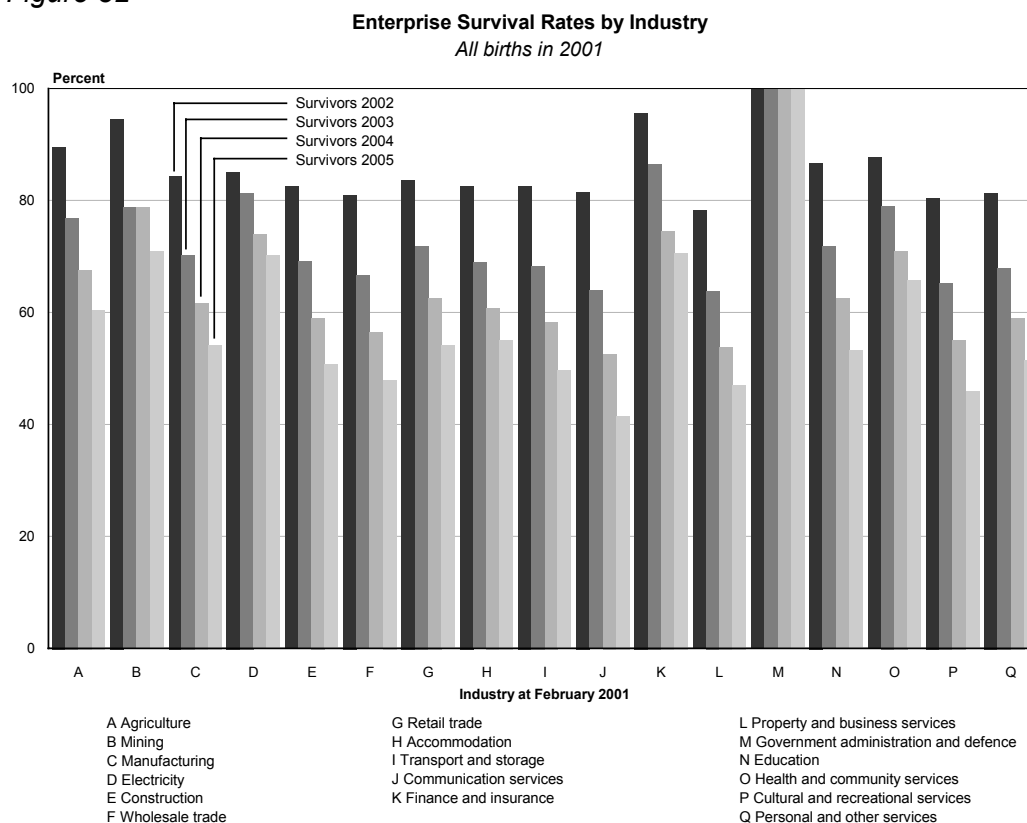
Aggregated survival rates (for all size groups) for enterprises birthed in 2001 are: 76 percent of births survived the first year, 65 percent the second, 57 percent the third, and 50 percent the fourth (not shown in Figure 31).

Figure 31



Enterprise births entering the government administration and defence industry in 2001 had a 100 percent chance of surviving into February 2005. This was followed by businesses birthed into the mining industry, the finance and insurance industry, and electricity, gas and water supply industry. All three industries recorded a chance of survival for four years of around 70 percent. Enterprises birthed in 2001 into the communication services industry had the least chance of survival until 2005 (42 percent).

Figure 32



Continuation rates for continuing enterprises

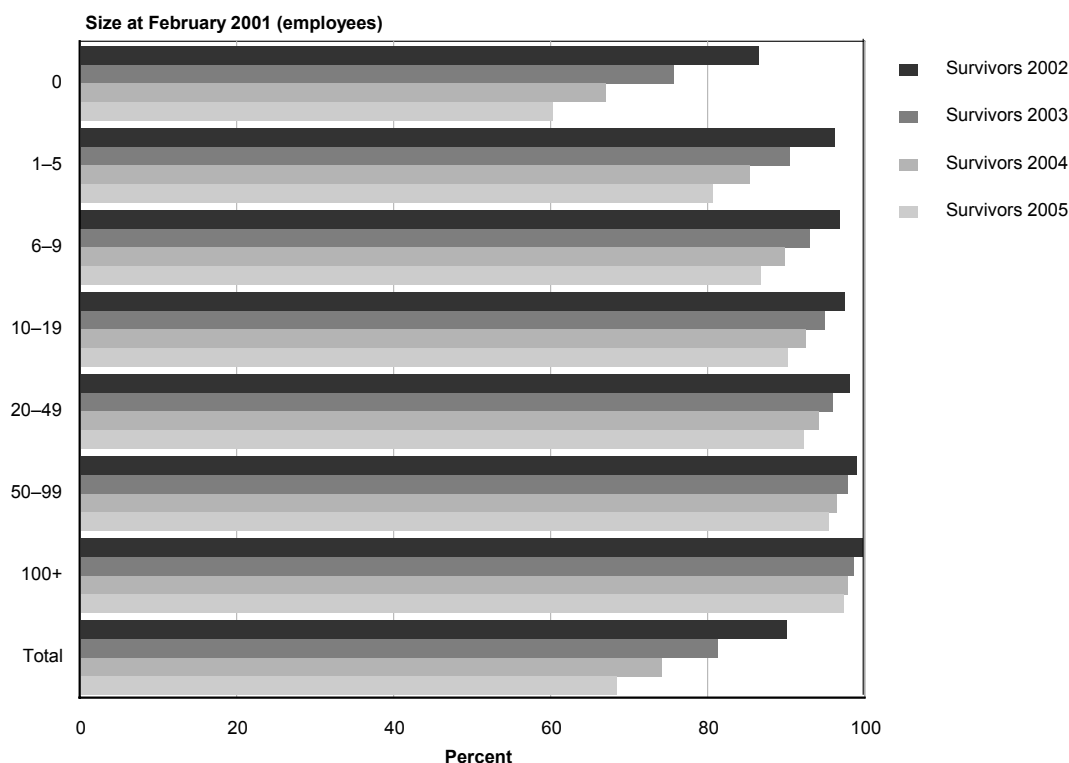
Given the natural churn of businesses entering and exiting the population, the factors influencing the survival of incumbent enterprises is of interest. For the purpose of this report, an enterprise is called incumbent if it was present in at least one snapshot before the current one. An incumbent of a period is said to survive or continue for further periods if it appears continually in the following snapshots without any breaks. The continuation rates thus measure the probability of the incumbent enterprise surviving beyond a given timeframe. Figure 33 shows continuation rates of incumbent enterprises in 2001 over the following four years.

As for the survival of enterprise births, of the total number of incumbent businesses with fewer than 19 employees on the LBF in 2001, the businesses most likely to continue until 2005 were those with 10 to 19 employees (90 percent continue). The chance of continuation strictly increases with business size. Enterprises with a size of more than 100 employees in 2001 had the highest chance to continue until 2005 (97 percent). Overall, 68 percent of incumbent businesses in 2001 continued until 2005.

Figure 33

Enterprise Continuation Rates by Size Group

Incumbent enterprises in 2001



Of the total number of incumbent businesses in any year from 2000 to 2003, almost 90 percent continued for at least one more year (over all EC size groups). The continuation rates for longer periods are similarly consistent and do not seem to depend on the year of observation. Roughly 81 percent of incumbent businesses survive for at least two years after observation, 74 percent for three years and 68 percent for four years. With further years coming up in the time series, it remains to be seen whether this pattern will hold over longer periods of survival (five years and more). Although the definition of incumbent enterprises allows for the inclusion of births of the preceding period, the aggregated continuation rates for incumbent enterprises are significantly higher than the aggregated survival rates for births.

Table 7

Continuation Rates of Incumbent LBF Enterprises

February 2000 to February 2003

Incumbent Enterprises in	Percentage of incumbent enterprises continuing				
	for at least one year	for at least two years	for at least three years	for at least four years	for at least five years
2000	89.4	81.3	73.8	67.7	62.6
2001	89.8	81.1	74.0	68.2	
2002	89.5	81.2	74.5		
2003	89.9	81.9			

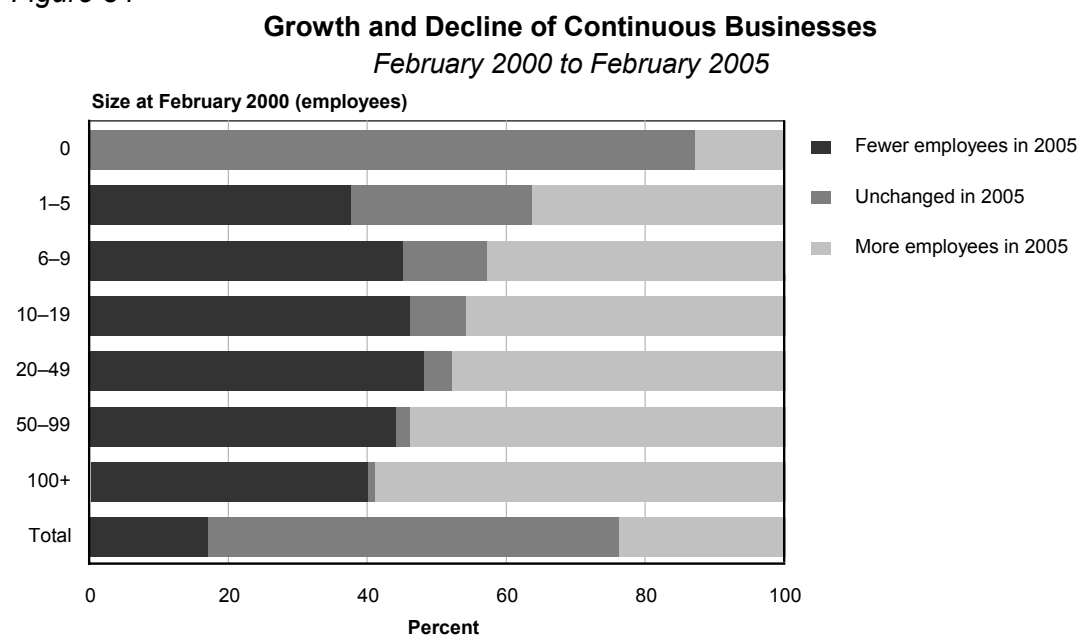
Growth and loss of employment 2000–2005

For enterprises that existed continuously from February 2000 to February 2005 – that is, that were present in all snapshots over six years – the employee counts of enterprises at February 2000 were compared with the employee counts for those enterprises at February 2005. Based on their February 2000 EC size group, the continuous businesses were grouped into three bands: those with fewer employees (if EC of 2005 was less than in 2000), those that remained unchanged, and those with more employees (if EC in 2005 was greater than in 2000).

The results show that of all continuing enterprises that started out in February 2000 as non-employing businesses, 87 percent remained non-employing until February 2005, while only 13 percent became employing businesses. As would be expected, the higher the EC size group of an enterprise was in 2000, the lower was the probability that its employee count would remain unchanged over six years.

Overall, 59 percent of continuous businesses remained unchanged, 24 percent increased their employee count, and 17 percent reduced staff over the period from 2000 to 2005.

Figure 34



5. Future Development and Call for User Feedback

The results in this report show that basing business demography statistics on the Longitudinal Business Frame (LBF) is feasible and that the differences between the experimental and official series are explicable. Considerable scope remains for future development and analysis.

To bring the experimental series into production, a new system has to be developed and integrated into Statistics New Zealand's overall information technology (IT) environment (see section 5.1). Before changing the official series to the new data source (the LBF), Statistics New Zealand is seeking to agree with stakeholders on a set of business demography outputs that best meets user needs. Section 5.2 invites your comments on this report and the future business demography series.

During the preparation of this report, the following issues were identified, which merit further analysis:

- How can business demography results best be tied in with results from the Linked Employer-Employee Data project (LEED) and other Statistics New Zealand surveys, particularly those measuring business performance?
- How can the data on self-employed individuals/working proprietors, which is expected to be added to the LBF later this year, best be used in business demography statistics? How would the addition of working proprietors into the employee count affect business size measures?
- Can and should the employee count at a point in time be replaced by a rolling mean employee (RME) count calculated over the whole period (one year)?
- What is the impact on statistics of businesses changing their industry classification between snapshots?
- How can (predominant) region be defined for enterprises?
- How do aggregate statistics compare if they are expressed separately for the private and public sector?
- How many very short-lived births are not included in the statistics, because their lifespan does not include the month of February?
- Can death statistics be expressed by age of businesses?
- Should the size of births/deaths be measured at other times than the first/last snapshot in which they appear? For example, is the number of jobs lost through enterprise deaths underestimated because the number of staff may be reduced over a period of time before the actual death?

The business demography team plans to incrementally address these questions during stage 2 of the Business Population Statistics Review Project. Issues of data integration using the LBF are currently being addressed by the Improved Business Understanding via Longitudinal Database Development (IBULDD) feasibility study at Statistics New Zealand, which aims at linking annual enterprise data to the LBF to measure enterprise productivity and performance.

5.1. Plan and timeframe for system development

The Business Frame (BF) is a large relational database that consists of over 150 tables. The existing Business Demographic Statistical System is a specialised system tailored to extract an annual snapshot as at February of each year from the BF's complex history tables. The snapshot is taken about six months after the month of February, so that the state of the BF as at February has to be reconstructed. As the BF undergoes constant change, such a reconstruction of data at a point in time is complex and prone to error.

As an alternative source for business demography statistics, the LBF simplifies the large BF database into a comprehensive and fit-for-use relational database. All historical data is readily available on the LBF on a monthly basis. The monthly data that is relevant to business demography can be seen and analysed from a single table.

To bring business demography statistics based on the LBF into production, a new system within the IT environment of Statistics New Zealand has to be created that will perform the following tasks:

- select data from the LBF and store it in an input data environment
- calculate additional BD attributes such as predominant industry, enterprise employment count, EC size group
- perform analysis of births and deaths, and mark identified births and deaths in the dataset
- aggregate data into statistical series by attributes and calculate totals
- store results in an output data environment.

The new system will be aligned with Statistics New Zealand's Business model Transformation Strategy, which is an overall framework for processes and systems within Statistics New Zealand.

It is expected that system specification will be completed by the end of 2006, and system development carried out in the first half of 2007. The first official publication of business demography statistics based on the LBF is planned for the end of 2007.

5.2 Opportunity for user feedback

Statistics New Zealand welcomes user comments on this report. In particular, we are interested in consulting on the following issues:

- Are users comfortable with the Longitudinal Business Frame as the data source for business demography statistics?
- Does the population selection meet user requirements?
- Does the observation of the population at a point in time (snapshot), as opposed to over a whole period of time, meet user requirements?
- Are users comfortable with the proposed methods of identifying enterprise births and deaths?
- Are users comfortable with the employee count sourced from the LEED project as the business size measure?
- Which of the presented experimental statistics are of most value to users and should therefore be included in annual official publications (Hot Off The Press etc)?
- Which of the presented experimental statistics may need more background information to make them more accessible and useful?
- Which other statistics published by Statistics New Zealand would users like to see integrated in with business demography statistics?

The Business Demography team can be contacted at:

Business Demography
Statistics New Zealand
Private Bag 4741
Christchurch
New Zealand

Telephone: +64 (03) 964 8700

Fax: +64 (03) 964 8759

Email: info@statistics.govt.nz

A. Appendix

A.1 Change in Business Frame maintenance strategy

In 2003, there was a significant change in the strategy used to maintain the Business Frame (BF), from which business demography statistics are sourced. This strategy involves greater use of administrative data to maintain the BF. A summary of the changes that have resulted from the change in strategy include:

- changed employment measure on the BF from a full-time equivalent persons engaged (FTE) measure to employee count (EC, see Appendix A.2)
- increased coverage of the BF to include all employing businesses – with the exception of individuals that are employers but are not registered for Goods and Services Tax (GST) – and reactivation of previously ceased businesses that are showing GST activity
- reduced compliance costs by decreasing the reliance on survey-sourced information to maintain the BF
- improved coverage of GST-exempt industries by making greater use of tax data (sourced from the Employer Monthly Schedule (EMS) and IR10 tax returns)
- faster processing of the births and deaths of enterprises to more accurately reflect real-world changes
- inclusion of farming businesses in the maintenance strategy (previously excluded)
- definition of boundaries for maintenance of enterprises on the BF on the basis of business size – larger enterprises continue to be primarily updated using annual maintenance survey data and tax data, while smaller enterprises are principally maintained using tax data.

The effect of these changes on business demography statistics has been examined. The outcomes of this work can be summarised as follows:

- Births and reactivations of enterprises identified solely as a result of changes in the maintenance strategy have been excluded to ensure comparability of statistics with previous years.
- The faster processing of births and deaths had a minimal impact on business demography statistics, so no adjustments were made.
- The farming industry has been excluded in the release of business demography statistics to ensure comparability of results with recent years. Statistics on the farming industry are available to users on request.

The outcome of changing the business size measure in business demography statistics from FTE to EC is discussed in Appendix A.2.

A.2 Changes in employment data

An important change from the 2004 release onwards has been that the indicator of business size (employment levels) now used is the employee count (EC). This replaces the previously available FTE measures (including full-time and part-time employees and working proprietors). To enable trends to be studied, the EC measure has been cast back to the year 2000.

The EC is sourced primarily from Inland Revenue's IR348 form – the EMS. This form is required to be completed on a monthly basis by employers, and allows for the number of salary and wage earners to be derived. The EC used for the business demography statistics is for the February month. There are a small number of enterprises whose employee count is collected by Statistics New Zealand surveys.

The change to the EC measure has the following benefits:

- reduced compliance load for small and medium-sized businesses
- improved coverage – information on businesses involved in farming is now maintained on the BF after being excluded in recent years. Business demography statistics on the farming industry are available to users on request
- improved accuracy – the attributes of businesses on the BF will now be updated more regularly, improving accuracy as a result.

The following table summarises the main differences between EC and FTE:

	EC	FTE
Source	mainly sourced from Inland Revenue's EMS (there is a small number of enterprises whose employee count is collected by Statistics New Zealand surveys)	was updated on the BF using survey feedback from respondents
Business Frame maintenance	updated monthly on the BF	was updated annually, as at February on the BF
Measure	EC is a head count of all salary and wage earners. It is mostly a count of employees, but can include a small number of working proprietors who pay themselves a taxable salary or wage	FTE is the total number of employees and working proprietors working full-time, plus half the number of employees and working proprietors working part-time
Full-time/Part-time split	not available	available
Gender breakdown	not available	available
Availability	available from 2000	discontinued in 2003

A.3 Industry coverage

Below is a summary of the industry coverage available in the business demography statistics, dating back to 1994.

The coverage of the business demography statistics has changed in recent years, as more industries have been included in the population. Historically, most of these industries were excluded because they contained a large proportion of enterprises that were not registered for GST, or a large proportion of enterprises that fell below the threshold of economic significance.

Since 1997, the selection criteria and standard published industry categories for the business demography statistics have been based on the Australian and New Zealand Standard Industrial Classification (ANZSIC). In 1996, the statistics were published using ANZSIC, but the selection criteria were based on the New Zealand Standard Industrial Classification (NZSIC).

2004–to date

All industrial activity covered.

1999–2003

Excludes agriculture production (ANZSIC subdivision A01).

1998

All industrial activity covered.

1997

Excludes agriculture production (ANZSIC subdivision A01).

1996

Excludes:

- agriculture production (NZSIC major group 111)
- residential property leasing and rental (NZSIC subgroup 83121)
- religious organisations (NZSIC subgroup 93910)
- social and related community services (NZSIC 93990)
- sporting and recreational clubs (NZSIC subgroup 94402).

1994–1995

Excludes:

- agriculture production (NZSIC major group 111)
- residential property leasing and rental (NZSIC subgroup 83121)
- commercial property leasing and rental (NZSIC subgroup 83123)
- day care centres and crèches (NZSIC subgroup 93402)
- other welfare institutions (NZSIC subgroup 93403)
- business, professional and labour associations (NZSIC subgroup 93500)
- religious organisations (NZSIC subgroup 93910)
- social and related community services (NZSIC 93990)
- sporting and recreational clubs (NZSIC subgroup 94402).

Statistics can be produced using two forms of industrial coverage:

- 1997 industrial coverage: this coverage excludes the agriculture production industry (ANZSIC subdivision A01). This data can be produced from 2000–2005, using the EC employment measure, to allow for time series comparison.
- 2004 industrial coverage: this includes all industrial activity. This data is only available from 2004.

Due to the differences in industry coverage and methodology applied, data produced using 1997 and 2004 industry coverage are not directly comparable.

A.4 Glossary of terms

Vale (2006, pp 53–54) provides a list of working definitions that have been adapted for use in this report.

Births – A birth is the creation of a combination of production factors, with the restriction that no other national businesses are involved in the event. Births do not include entries into the population due to reactivations, mergers, break-ups, split-offs or other restructuring of a group of businesses linked by ownership or control. Births also exclude entries into a population resulting from changes to characteristics of existing businesses. (Note: this is largely based on, and fully consistent with, the Eurostat definition of enterprise births.)

Deaths – A death is the dissolution of a combination of production factors, with the restriction that no other domestic businesses are involved in the event. Deaths do not include exits from the population due to temporary inactivity, mergers, takeovers, break-ups or other restructuring of a group of businesses linked by ownership or control. Deaths also exclude exits from a population resulting from changes to characteristics of businesses which remain active. (Note: this is largely based on, and fully consistent with, the Eurostat definition of enterprise deaths).

Entries – Businesses that are present in the population at the end of the period, but were not present at the start of the period (Vale, 2006, calls these businesses ‘joiners’).

Exits – Businesses that are present in the population at the start of the period, but are not present at the end of the period (Vale, 2006, calls these businesses ‘leavers’).

Other entries – All entries that are not births.

Other exits – All exits that are not deaths.

Other births – Identified births whose birth date is earlier than a defined cut-off date (mostly reactivations).

Population – All businesses that meet certain pre-defined criteria at a specific temporal reference point.

Pure births – Identified births with a recent birth date later than a defined cut-off date.

Short-lived births – Identified births that disappear in the period immediately after the period in which they emerged (due to death or dormancy).

Surviving births – Identified births that survive for at least one period after the period in which they emerged.

Survivors – All businesses that are in the population at both the start and the end of the period.

Bibliography

Eurostat (2004). "Business Demography Recommendations Manual, Status August 2004", unpublished paper, Eurostat, Luxembourg.

Green W (1993). *Econometric Analysis*, 2nd edition, Prentice Hall, Upper Saddle River, New Jersey.

Kelly N (2003). "Repairing EMS Employer Longitudinal Links", Statistics New Zealand, Wellington. Available at: <http://www.stats.govt.nz/NR/rdonlyres/4479EFF7-B28C-4611-9C75-8333B5FCC2C4/0/RepairingEMSEmployerLongitudinalLinks.pdf>

Law D and McLellan N (2005). "The Contributions from Firm Entry, Exit and Continuation to Labour Productivity Growth in New Zealand", New Zealand Treasury Working Paper 05/01, Wellington.

Organisation for Economic Co-operation and Development (2002). "OECD Small and Medium Enterprise Outlook, Enterprise Demography: Examining Business Dynamism in OECD Countries" (DSTI/IND/PME(2002)6/PART2), OECD, Geneva.

Seyb A (2005). "Statistics New Zealand's Longitudinal Business Frame", in *Maximising Data Value. Data use and Reuse*, Statistics New Zealand, Wellington.

Statistics New Zealand (2005). Business Demography Release Information. Available at: <http://www.stats.govt.nz/products-and-services/info-releases/bus-demo-stats-info-releases.htm> [21 November 2005].

Statistics New Zealand (2006). Information on the Improved Business Understanding via Longitudinal Database Development (IBULDD) Feasibility Study. Available at: <http://www.stats.govt.nz/developments/longitudinal-research-of-business-dynamics-project.htm>

Statistics New Zealand (2006). Linked Employer-Employee Data release. Available at: <http://www.stats.govt.nz/leed/default.htm> [13 February 2006].

Vale S (2006). *The International Comparability of Business Start-up Rates, Final Report*, OECD Statistics Directorate/Office for National Statistics, London, United Kingdom.