Statistics NEW ZEALAND Te Tari Tatau

Prototype Outputs Using Linked Employer-Employee Data

Neil Kelly

30 September 2003

Contents

Summary

The Linked Employer-Employee Data (LEED) project is a data integration project that uses record linkage to establish relationships between various entities such as individuals in employment, their employers, and equivalent statistical collection units. Future stages of this project will bring together, or link, datasets held by the Inland Revenue Department (IRD) as administrative tax records, with data held by Statistics New Zealand. As such, it will provide a rich source of information on both labour market and employment flows, as well as firm performance and dynamics.

This paper presents a selection of some of the first prototype outputs derived from Stage 1 of the LEED project. It is proposed that these measures, when fully developed, will comprise the first official statistical series to be produced from this data. This stage of the project is intended to demonstrate the ability of the data to produce new, and enhance existing, labour market and business demographic measures, as well as to highlight future work needed. As such, the investigations that underlie the results in this paper serve two main purposes: firstly, to test the ability of the administrative data used in LEED to produce ongoing employment and business demographic indicators, and secondly, to demonstrate the types of new indicators that can be produced using this data.

In addition to the outputs shown here, findings suggest that the LEED data infrastructure, when fully cleaned and integrated with official data sources, would be able to support the production of a range of additional new statistics; both at industry and regional level.

The primary intention of this paper then is to illustrate the potential of LEED to produce the kinds of measures presented, not to produce these outputs as official statistics.

(1) New and Enhanced Outputs

The LEED infrastructure would support the enhancement of existing outputs related to employment and firm dynamics, as well as the production of new statistics that are unable to be produced using existing data sources. Some example of the outputs that could be produced from LEED include:

- Job flows: the contribution to net employment change between two points in time made by those firms that create jobs, either by expansion or start-up, and by those firms that destroy jobs, either by contraction or closure.
- Worker flows: the number of employees who either commence a new job, or leave an existing position between two points in time.
- Employment tenure: the length of time an individual is employed by the same employer.
- Multiple jobholding: the numbers of individuals who hold multiple jobs and their earnings in each job.
- Business demographic outputs such as longitudinal firm or employment dynamics. These measure the performance of cohorts of firms over time in terms of their employment and survival.

These measures are currently able to be produced on either an annual or quarterly basis and to be broken down by firm characteristics such as firm size or growth rates, as well as employee characteristics such as age, sex or income distribution. Future stages of LEED will add other firm characteristics from SNZ's Business Frame, such as industry or region.

(2) The Way Forward

While the summary outputs in this paper show the potential of the data to enhance the understanding of the New Zealand labour market and firm performance, further work is required before these outputs can be produced as official statistics. Broadly speaking, attention needs to be given to finalising development in the following three areas:

- cleaning and transforming the input data into a form that will support statistical research
- integrating administrative data sourced from IRD with data held by Statistics New Zealand, and
- further analysis of many of the aggregate prototype outputs produced to determine the factors that underlie the broad patterns that can be observed.

(3) Summary of Selected Results

- Apparent rates of job creation and destruction and their distribution, derived from LEED, are broadly consistent with earlier New Zealand findings. See Carroll et al (2002a, 2002b).
- The New Zealand labour market is characterised by turbulence. Initial findings suggest that up to one in every seven jobs in existence will have been destroyed, either by firm contraction or firm death, 12 months later. A similar proportion of new jobs will have been created.
- Worker turnover is similarly high. Up to 40 percent of all jobs in existence will not be filled by the same worker in 12 months time.
- Small firms are less stable than larger firms. Twenty-eight percent of jobs where the employer has between one and five employees will have been destroyed 12 months later. In the case of employers with more than 100 staff, the proportion is 7.5 percent.
- Older employees (those over 45 years of age) have more stable employment patterns than 'prime age' employees (those aged between 25 and 44 years). Prime age employees in turn have more stable employment patterns than those under 24 years of age.
- There is not a large difference in the stability of employment patterns between male and female employees. The proportion of jobs filled by female employees, which will not be filled by the same worker in 12 months time, is approximately 1 percent higher than the equivalent proportion for male employees.

Note

Outputs shown in this paper have been compiled using input data that is largely in its 'raw' form. Limitations exist in the ability of the raw data to support statistical outputs, some of which are outlined in this paper. Investigations suggest that the effects of these limitations may be reduced using appropriate methods that are being developed. However, these methods have not been incorporated into the outputs in this document. As such, results shown are to be regarded as illustrative of potential only. This document is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

Tables in this paper contain information about groups of people or firms so that the confidentiality of individuals is protected. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person or firm. The results are based in part on tax data supplied by the Inland Revenue Department to Statistics New Zealand under the Tax Administration Act. This tax data must be used only for statistical purposes and no individual information is provided back to IRD for administrative or regulatory purposes. Careful consideration has been given to the privacy, security and confidentiality issues associated with using tax data in this project. A full discussion can be found in the paper *Linked Employer-Employee Data Project: Privacy Impact Assessment.*

The IRD collects this data to support the efficient operation of the New Zealand taxation system, and its use as a base for the production of statistics places new and quite different demands on the data. Any discussion of data limitations or weaknesses is in the context of this latter use, and is not related to the ability of the data to support IRD's core operational requirements.

1. Background

1.1 Introduction to LEED

The Linked Employer-Employee Data project (LEED) is a record linkage project that establishes and verifies relationships between various entities. LEED brings together, or links, datasets held by the Inland Revenue Department (IRD) as administrative tax records with data held by Statistics New Zealand. A number of links exist within and between the various data sources. Specifically, employees are linked to their employer, and employers will in future be linked to equivalent statistical collection units on Statistics New Zealand's Business Frame (BF). In addition, as all entities within LEED are linked longitudinally, it is possible to track individuals' job histories over time, and to link this to longitudinal firm dynamics. For more detail on the links contained in LEED, see Bycroft (2003).

As both employers and employees are explicitly identified within LEED, it is possible to produce outputs that are classified by characteristics of the employer (for example firm size), the employee (such as age or sex) as well as derived variables such as firm growth rates, or individuals' income quantiles. In future, additional firm characteristics drawn from the BF will be incorporated, such as industry and institutional sector. The potential of LEED to produce a range of new regional outputs is also being investigated. As such, LEED provides a rich and powerful source of information on labour market performance, and business and employment dynamics.

This paper presents a selection of some of the outputs that are able to be produced using LEED. Some of these outputs are equivalent to either official SNZ statistics, or to the results of similar labour market studies, most notably Carroll et al (2002a, 2002b). These are produced both as experimental outputs in their own right and also to test the 'fitness for use' of LEED by comparison with existing statistics. Other outputs are new and are unable to be produced from existing data held by SNZ.

2. Description of the Source Data

2.1 Source Data

Outputs contained within this paper have been produced using a single administrative data source, as well as two derived tables that contain the birth date (used to derive age) and sex of each employee.

The administrative data table, termed *return_line_items* is sourced from the Inland Revenue Department's (IRD) Data Warehouse, and is derived from the Employer Monthly Schedule (EMS) payroll return. As the name suggests, the EMS is filed monthly, and covers all payers and recipients of income taxed at source, other than interest and dividends. Each payer and recipient is explicitly identified by a unique IRD identifier, termed the *employer_ird_number* and *employee_ird_number*, respectively. The dataset used in Stage 1 of LEED covers the periods April 1999 to March 2002.

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. Broadly speaking, those individuals who have PAYE deducted are employees, while those who pay withholding tax are a subset of the self-employed. However, not all self-employed individuals will appear on the EMS, as the majority do not pay their tax at source, and evidence suggests that a small number of self-employed individuals may earn income with PAYE tax deducted.

The recipients of most social welfare benefits are also covered by the EMS; these can be readily identified, as the IRD's unique identifiers of the payers is known.

The *return_line_items* data is used in its 'raw' form – no editing or imputation has been incorporated into outputs shown, other than the deletion of records with invalid IRD identifiers, which comprised approximately 2 percent of all records.

For additional detail on the editing/imputation of date of birth and sex of individuals refer to Graham (2003).

Variables used are:

return_line_items
employer_ird_number - the IRD number of the employer
employee_ird_number - the IRD number of the employee
return_period_date - the final day of the month covered by the return
gross_earnings_amount - the income earned by the employee before tax is deducted

date_of_birth
ird_number - the IRD number of the taxpayer (may be an employee or employer if an individual)
date_of_birth - the date of birth of the taxpayer

employee_sex
ird_number - the IRD number of the taxpayer (may be an employee or employer if an individual)
sex - the sex of the taxpayer

2.2 Population

Included:

All employers and employees, subject to output definitions and exclusions below.

It should be noted definition of 'employee' used will include any individuals who pay their income tax at source and with PAYE deducted. To the extent that working proprietors, partners or other self-employed choose to pay income tax in this manner, they have not been separated from 'true' employees.

Excluded:

- All self-employed who are able to be identified by virtue of:
  - Paying tax at source with a withholding tax type of 'W'
  - Not paying tax at source, and so not appearing on the input dataset at all.
- All persons who are working without pay (such as unpaid workers in a family business), or who are working 'under the table' and so not paying tax.
- All employers and employees with missing or invalid IRD numbers.
- All employees who were under 15 at the start of the reference month.
- All payment records in which the employee received zero gross earnings in the reference month.
- All payment records relating to the receipt of social assistance benefits.

2.3 Collection Unit

The collection unit used in the analysis presented in this paper is the legal entity that files the EMS return. The unique identifier is the *employer_ird_number.*

2.4 Strengths and Weaknesses

Without doubt, the strength of the LEED data in general and of the EMS in particular is in its almost universal coverage. Every taxpayer who pays tax at source on income related to labour input is captured. Being effectively a census, outputs can potentially be produced for very small groups of interest without introducing sample error. Each employer and employee carries a unique identifier, and the two are linked. Furthermore, the identifiers are generally stable, which means that the data is longitudinal; that is, employers and employees can be tracked over time.

However, there are some weaknesses in the data:

- Currently information on mergers, acquisitions and splits is not available, which means that some employers cannot be readily tracked longitudinally. As a result, some firms that are continuing enterprises may be incorrectly classified as firm births or deaths.[1] This will upwardly bias both job and worker flow statistics, and will also adversely affect the ability to produce statistics on the dynamics of new firms, as opposed to continuing enterprises. While we are confident that methods can be developed to deal with this issue (see Kelly, 2003), they have not been incorporated into the results presented in this paper.
- As stated above, the EMS is filed at the legal entity level, which generally corresponds to an enterprise on the BF. While in many cases this will also correspond to a single geographic unit, data is not available for multi-geographic-unit enterprises at the finest unit level. Evidence also suggests that in some cases a single EMS return may in fact cover several legal entities. For more details, see 4. Matters Outstanding, at the end of this paper, and Statistics New Zealand (2003).
- Identification of persons in self-employment is difficult, with only partial coverage on the EMS. For this reason, analysis in this paper has been confined to employees only.
- The EMS captures all persons engaged in paid employment in the month covered by the return, regardless of the timing, or the length of that employment. In contrast, many official labour market statistics are based on employment at a specific point in time. To the extent that employees change jobs during the EMS reference period (one month), they will be counted as having two jobs during that period, which will upwardly bias both the number of jobs and the number of multiple job holders. Furthermore, the count of persons in employment will be higher than point in time outputs as a result of individuals leaving their job during the return month, but before the reference point in time, and **not commencing** another job. The EMS return does contain start and finish dates for each job, which in theory will facilitate the derivation of point in time employment and number of filled jobs. Some data quality issues exist with these fields, and while editing and imputation methods have been developed, they have not been incorporated into this paper.

In addition, approximately 2 percent of records have invalid employee IRD identifiers and an as yet unknown proportion may have a valid but incorrect *employee_ird_number*. It is not expected that these latter issues will have a sizeable effect on any of the stocks presented at the level of aggregation used in this paper such as counts of employees or firms, although problems may arise at the finer levels of aggregation that may be used in customised statistics, or policy-

1. Firm birth and death is proxied in Stage 1 of LEED by the change in firm status from non-employing to employing, and vice versa; refer to Statistics New Zealand (2003a).

relevant research. However, the effect on flow statistics such as job creation of destruction is significant, especially where the rate of those flows is small. The ability to produce job tenure outputs is most affected, as this depends on the employee identifier being correct for not just one but all periods.

3. Results

All tables and graphs that follow can be produced on either an annual or quarterly basis. The focus of this paper is to demonstrate the kinds of outputs that can be produced and the feasibility of the data to support new outputs, not to produce these as official statistics at this stage.

The measures in this paper have been produced at a high level of aggregation. Understanding the patterns and determinants of these new measures is a complex task requiring detailed analysis. This work is beyond the scope of this paper, the focus of which is to introduce the potential outputs and concepts to future end users.

Furthermore, there is an insufficiently long series of input data to produce meaningful time series on an annual basis; only three years of data are available, meaning that only two years of flow statistics can be produced. Therefore, the majority of outputs, unless otherwise stated, are shown as summaries of annual snapshots as at February 2000, 2001 and 2002. These periods were chosen for comparability with earlier studies, which were based on the February business demography snapshots.

As far as possible, output definitions used in this report have been aligned with those used in existing official statistics produced by SNZ and/or accepted international practice. In some cases however, the conceptual 'ideal' has been deviated from due to practical considerations.

In brief, the main definitions which underlie the prototype outputs presented in this paper are:

Employee: any individual who appears on the EMS with non-zero gross earnings in the reference month, who was over 15 years old at the start of that month, and who does not have withholding tax deducted. The employee is identified by the *employee_ird_number.*

Employer: any taxpayer entity (for example a company, partnership, trust or individual) who employs and pays a non-zero number of employees, as defined above, in the reference period. In this paper, the payers of known source-deducted social assistance benefits are excluded. The employer is identified by the *employer_ird_number.*

In line with international practice (see Davis and Haltiwanger, 1992) the denominator in all flow statistics, whether by firm or at aggregate level, is the mean of employment in the current and previous reference periods, ie of periods *t* and *t*-1.

For more detail on the definitions used in Stage 1 of LEED, see Statistics New Zealand (2003).

3.1 Comparison with Existing Official Statistics

In order that the characteristics of the LEED source data be understood in the context of comparison with data sources used to produce existing official statistics, some brief comparisons were carried out with corresponding official outputs.

3.1.1 Business Demography: Count of Firms and Persons Engaged, by Firm Size

To provide a context for the aggregate results presented in this paper, Table 1 presents a cross-sectional summary of some simple business demographic statistics derived from the EMS. The table is a simple snapshot of a single month's EMS return, with the time period chosen for comparability with the latest official business demographic data available. The first column in the table shows firm size as at February 2001. The second column shows the distribution of EMS filing units (the statistical unit used in the LEED outputs in this paper), while the third column shows the distribution of filled jobs.

The table may be contrasted with Table 2, which summarises counts of persons engaged from SNZ's official Business Demography tables.

Table 1

Cross-sectional Business Demographic Snapshot from LEED

Summary Business Demography Employment Statistics Snapshot as at Feb 2001		
sizegroup	count_employer	jobs
b) 1-5	93700	208100
c) 6-9	16800	121100
d) 10-19	12700	169100
e) 20-49	6500	195100
f) 50-99	1900	126300
g) 100+	1600	778600
Total	133200	1598300

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

Table 2

Official Cross-sectional Business Demographic Snapshot[2]

As at February 2001

Sizegroup	Enterprises	Persons Engaged
0-5	235796	394490
6-9	17724	147480
10-49	17700	382320
50-99	1443	113400
100+	1298	631990
Total	273961	1699700

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

It may be seen that, subject to definitions and populations constraints already outlined, in February 2001 there were 133,200 legal entities that filed an EMS covering paid employees in approximately 1.6 million jobs. It should be noted that for comparison with the official table which follows, the definition of firm size used to aggregate firms and jobs is based on employer size as at February 2001. This is different from other tables in this paper, which cover dynamics over more than one point in time, and hence aggregate counts and flows based on mean firm size over all periods.

2. Source: Statistics New Zealand website http://www.stats.govt.nz/domino/external/pasfull/pasfull.nsf/7cf46ae26dcb6800cc256a62000a2248/4c2567ef00247c6acc256c63006b0274?OpenDocument).

Table 1, and the official table, which follows, highlight the fact that New Zealand's economy is characterised by firms with a relatively small number of employees. Approximately two-thirds of firms have five or fewer employees, covering around 13 percent of employment. Many international studies that focus on analysis of employment dynamics restrict the population of interest and exclude small firms; this option is not tenable in the New Zealand context.

There are some conceptual differences between the two tables: most importantly, table 2 excludes most agricultural industries as well as private households employing staff, but includes working proprietors. However, it may be seen that the most salient differences between the two tables are that:

1. The official table (table 2) has many more firms and persons engaged in the smallest firm size group.
2. The LEED-derived table has more firms and employees in the largest firm size group.

These differences are in the main due to three points:

1. The LEED population of interest in this paper has been confined to employees only. Those firms that have zero employees will not file an EMS return. The official Business Demographic table encompasses approximately 50,000 such enterprises. In addition, any sole traders or partnerships that do not engage paid employees, and whose proprietors do not pay their tax at source, will also not appear on the EMS. Such firms may be expected to be small in size.
2. LEED data is compiled at a higher level of aggregation than shown in table 2, above. Specifically, evidence suggests that in some cases a single representative of a consolidated group of companies may file an EMS return covering employees of other firms within the same group. To the extent that this occurs firm size will be overstated, which may upwardly bias the number of firms and employees in the larger size groups.
3. The EMS covers all employees engaged in a given return period, whether permanent or casual. Investigation of selected unit records suggests that in some cases, employers may engage large numbers of temporary staff, who may not be regarded as employees for the purpose of response to SNZ questionnaires. Detailed analysis has not been performed, but many of these firms appear to be engaged in the agriculture and horticulture industries.

Previous analysis of job flows in the New Zealand context has used the geographic unit as a base: see Carroll et at (2002a, 2002b). The geographic unit is defined as "a separate operating unit... engaged in predominately one kind of economic activity, from a single physical location and base". This is consistent with the analysis undertaken by Davis and Haltiwanger, which was performed at plant level. It may be expected that the higher level of aggregation inherent in LEED datasets compared with the geographic unit may cause downward bias in job and worker flows, to the extent that workers move between physical locations in multi-unit firms, and between firms in multi-firm groups, as well as downward bias in counts of firms. Employee counts should not be adversely affected by this issue.

3.1.2 Household Labour Force Survey: Count of Paid Employees

The pattern of counts of persons in employment was contrasted with the equivalent Household Labour Force Survey (HLFS) series.[3] The EMS series is a simple count of all unique employee

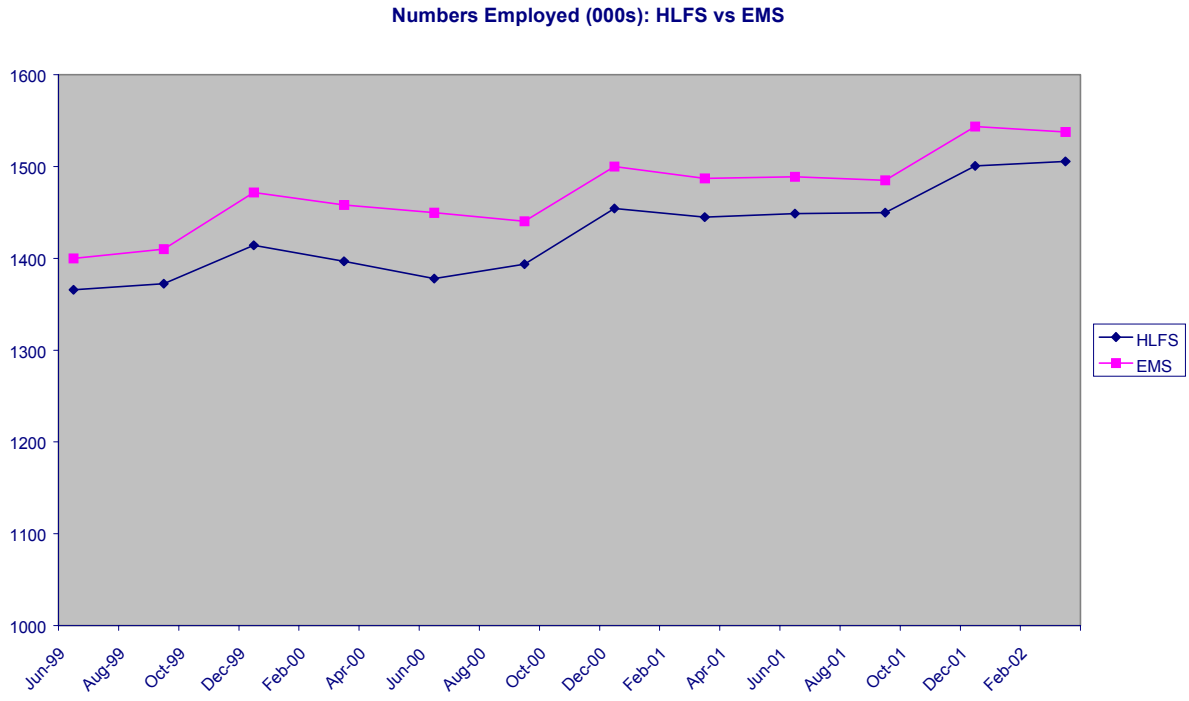
3. Household Labour Force Series HLFQ.SLA3HA: Wage or Salary Earners. For more information on the Household Labour Force Survey refer to the SNZ website: http://www.stats.govt.nz/domino/external/web/prod_serv.nsf/htmldocs/Household+Labour+Force+Survey+-+Information+Releases).

identifiers (the *employee_ird_number*) captured on any EMS return in the reference quarter. The HLFS series represents a count of all individuals in the working age population engaged as wage or salary earners in the reference quarter, and should, therefore, have essentially the same conceptual coverage as the EMS data.

Figure 1

Numbers Employed (000)

HLFS vs EMS



It may be seen that quarterly patterns in the two series are very similar. There is a level difference, with the EMS being generally around 3 percent higher than HLFS. While the reason for this discrepancy has yet to be explored in detail, it may nevertheless be seen that, at the aggregate level, the use of EMS sourced data does not give a radically different picture of changes in net employment from that observed in official statistics.

3.2 Summary of Job Creation, Destruction, and Net Employment Change

This paper first presents some statistics covering job creation and destruction.

Conceptually, a job is created when a firm, either new or existing, creates a new position for an employee and is destroyed when an existing position with a firm is no longer required, either due to firm contraction or closure.

These statistics represent the first 'level' of flows that underlie the net employment change presented in official statistics. Typically, studies in New Zealand and internationally have found that relatively small changes in net employment may be underpinned by much higher values of job creation and destruction; some firms grow while others shrink. For example, an increase in net employment of 1,000 jobs may simply mean that 1,000 new jobs have been created; it may also represent creation of 10,000 new jobs and the loss of 9,000 existing positions.

3.2.1 Annual Summary

The tables below show the mean annual values of job creation and destruction, averaged between the three annual snapshots taken as at February of each year.

Table 3

Mean of Annual Job Creation and Destruction

Job creation, destruction and net change			
jobs	create_ mvt	destruct_mvt	net_mvt
1625800	15.8%	(13.3%)	2.5%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals

Table 4

Annual Summary by Firm Size

Job creation, destruction and net change- mean over all periods by mean size group of firm				
sizegroup	jobs	create_mvt	destruct_mvt	net_mvt
b) 1-5	207800	32.2%	(28.4%)	3.8%
c) 6-9	105300	25.8%	(22.0%)	3.9%
d) 10-19	161900	21.5%	(17.3%)	4.2%
e) 20-49	193400	17.0%	(14.6%)	2.4%
f) 50-99	127200	14.8%	(12.7%)	2.1%
g) 100+	830100	9.3%	(7.5%)	1.8%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals

It may be seen that, on average over the three-year period covered by the EMS data, the annual rate of job creation in firms was 15.8 percent while 13.3 percent of existing jobs were destroyed; the average net rate of job creation across all firms was 2.5 percent. It may therefore be seen that relatively small changes in the number of jobs at the total economy level disguise comparatively large rates of job creation and destruction.

Turning attention to job creation and destruction by firm size, it may be seen that smaller firms are characterised by higher rates of job creation and destruction than larger firms, and furthermore that job flow rates decrease monotonically as firm size increases. While this finding is broadly consistent with the results of Carroll et al (2002a, 2002b), the LEED data shows a greater decrease in job flows as firm size increases. While detailed unit record investigation between the data used in this paper and that used by Carroll et al has not been undertaken, it may be surmised that the differences may be due to the following factors:

- The EMS collects data at the legal entity level, which generally corresponds to an enterprise. In the case of single unit firms, the legal entity also corresponds to a geographic unit. In the case of multi-unit firms, one EMS may cover several geographic units. To the extent that job flows exist between geographic units in multi-unit firms, LEED will understate these job flows. Almost by definition, multi-unit firms tend to be larger in size than single unit firms. Hence flows in the larger size groups may be understated.
- The practice of groups of firms filing a consolidated return on behalf of more than one enterprise will exacerbate this effect.
- Offsetting this effect is that fact that all paid employees, whether permanent or casual, will be captured by the EMS return if they have PAYE deducted. However, respondents may not regard casual staff as employees for the purposes of responding to the questionnaires used to maintain SNZ's Business Frame. Brief investigation of selected unit record data shows, for example, that firms in the agriculture and horticulture industries, which are seen to employ

large numbers of apparently casual staff from time to time, may record very few employees on the BF. Movements in the count of jobs filled by casual staff may lead to higher job flows than reported in Carroll et al.

3.2.2 Quarterly Summary

Job flows were also produced on a quarterly basis, shown below:

Table 5

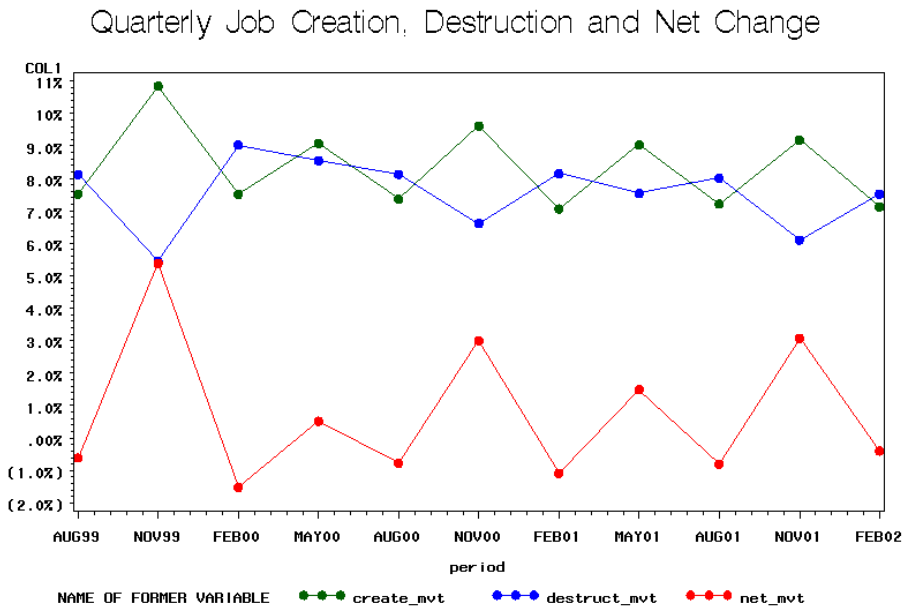
Quarterly Job Creation and Destruction

Job creation, destruction and net change by quarter				
return_ period	jobs	net_mvt	create_mvt	destruct_mvt
AUG99	1511900	(.6%)	7.5%	(8.1%)
NOV99	1595500	5.4%	10.8%	(5.5%)
FEB00	1571500	(1.5%)	7.5%	(9.0%)
MAY00	1579800	.5%	9.1%	(8.6%)
AUG00	1567800	(.8%)	7.4%	(8.1%)
NOV00	1615600	3.0%	9.6%	(6.6%)
FEB01	1598300	(1.1%)	7.1%	(8.2%)
MAY01	1622300	1.5%	9.0%	(7.6%)
AUG01	1609400	(.8%)	7.2%	(8.0%)
NOV01	1659800	3.1%	9.2%	(6.1%)
FEB02	1653300	(.4%)	7.1%	(7.4%)

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

Figure 2

Quarterly Job Creation and Destruction



The length of time series is insufficient to determine seasonal patterns with certainty. Nevertheless, it may be seen that, as well as the known seasonality in net employment, there is apparent seasonality in both the patterns of job creation and job destruction. In the November month, net employment growth is achieved by both increased job creation, and decreased destruction, relative to the preceding period. However, in the February, May and August snapshots, job destruction is relatively stable, with seasonality in net employment being driven by

movements in job creation. Future work may decompose these aggregate movements by industry to determine what is driving these patterns.

3.3 Decomposing Job Creation and Destruction Rates

The tables and graphs below show how job creation and destruction are spread across different magnitudes of job flows. All values are averages of movements between the annual February snapshots.

Table 6 decomposes job creation and destruction into three aggregate magnitudes of job flow – those firms that:

- grow/shrink between 0 and 10 percent
- grow/shrink by more than 10 percent, and
- are birthed/ceased.

The histogram in Graph 6 breaks these aggregate categories down into finer groups based on magnitude of job flow: those firms that have job flows greater than 0 and less than 10 percent (labelled +/- 5, respectively), those with job flows greater than or equal to 10 percent and less than 20 percent (labelled +/- 15) and so on.

Table 7 further decomposes the aggregate groupings of job flow based on firm size.

Table 6

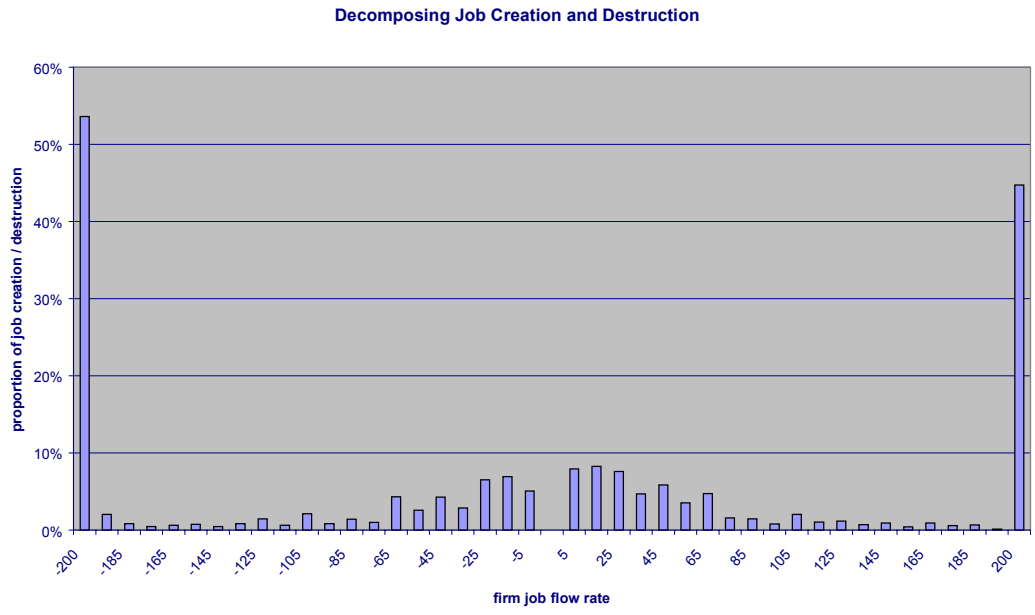
Contribution to Job Creation and Destruction

Decomposing job creation and destruction rates all periods							
death	lt -10 0 %	between 0 and -10%	total destruction	between 0 and +10%	ge +10%	birth	total creation
(7.1%)	(5.5%)	(.7%)	(13.3%)	1.3%	7.5%	7.1%	15.8%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

Figure 3

Contribution to Job Creation and Destruction



Note that the histogram bars sum to 100 percent of *each of job creation and destruction* respectively.

Table 7

Contribution to Job Creation and Destruction

By firm size

Decomposing job creation and destruction rates – all periods by sizegroup								
sizegroup	death	lt -10%	between 0 and -10%	total destruction	between 0 and +10%	ge +10%	birth	total creation
b) 1-5	(17.9%)	(10.6%)	.	(28.4%)	.	13.1%	19.0%	32.2%
c) 6-9	(12.1%)	(9.7%)	(.0%)	(22.0%)	.0%	11.9%	14.0%	25.8%
d) 10-19	(9.0%)	(7.5%)	(.8%)	(17.3%)	.9%	10.4%	10.2%	21.5%
e) 20-49	(7.5%)	(6.2%)	(.9%)	(14.6%)	1.1%	8.8%	7.1%	17.0%
f) 50-99	(6.4%)	(5.3%)	(1.0%)	(12.7%)	1.3%	8.1%	5.4%	14.8%
g)100+	(3.6%)	(3.1%)	(.8%)	(7.5%)	1.8%	4.7%	2.8%	9.3%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

The above tables and graph imply that approximately half of all job flows are the result of firm birth and death. However, there is an important data caveat which bears reiteration and further expansion at this point: enterprises may change their structure or legal form for a number of reasons. Often enterprises that appear as either new births or deaths in administrative systems may in fact relate to continuing enterprises. Examples include mergers, acquisitions and splits, as well as firms changing their legal character; for example, from a partnership to a company. These are administratively driven events and do not necessarily represent true births and deaths.

Some trial methods have been developed to facilitate the longitudinal linking of firms in such cases. The results of these investigations suggest that approximately 37 percent of all flows attributable to firm birth or death may not represent true flows. For more detail, see Kelly (2003).

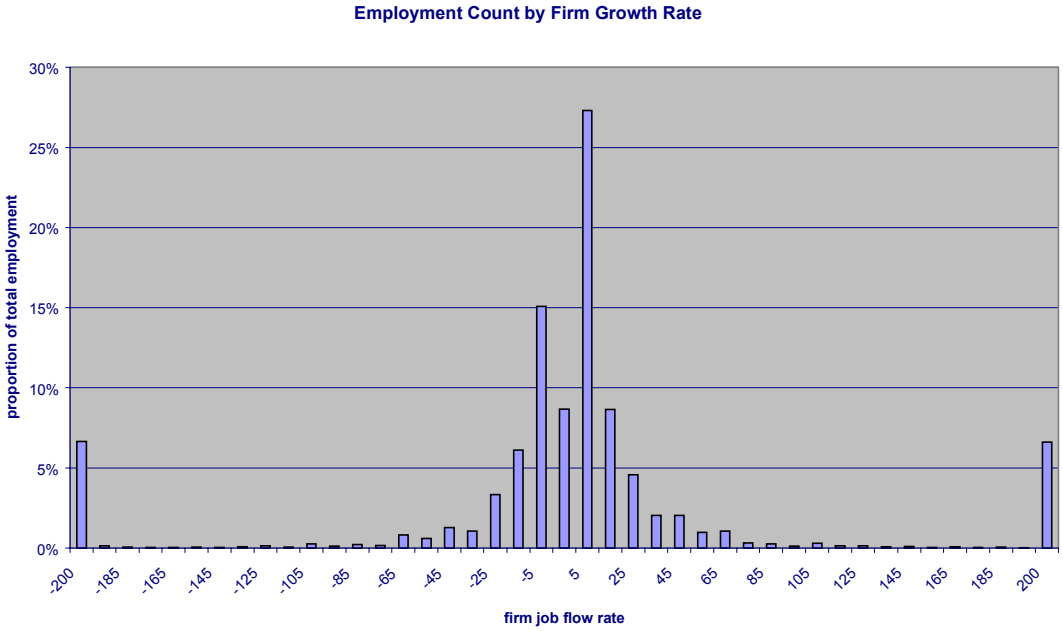
Due to time constraints, the methods developed have not been incorporated into the trial outputs presented in this paper. However, it may be assumed that, at the total level, the 'true' value of job (and worker) flows resulting from firm birth and death is in the order of 3 to 4 percent of total employment, or around 30 to 40 percent of all job creation and destruction. Distribution by firm size is not known.

The above data shows that, even after correcting for the effects of spurious birth and death, most of the job flows in the New Zealand economy occur in firms that have comparatively large changes in employment. This effect may be seen across all sizes of firm; however, the relative contributions of firm birth and death are higher in small firms than in larger ones.

3.4 Distribution of Employment by Firm Growth Rate

Figure 4

Distribution of Employment by Firm Growth

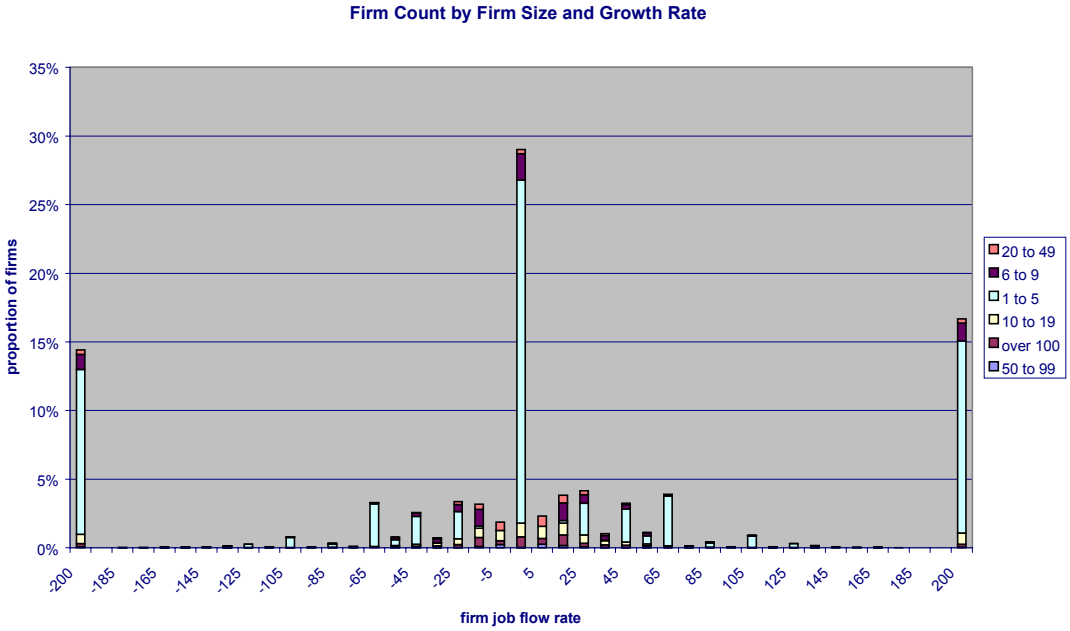


The above graph shows that, while most of the contribution to job creation and destruction occurs in firms which have comparatively large job flows, most *employment* is concentrated in firms with relatively small flows; almost 50 percent of jobs are in firms that have absolute job flows of less than 10 percent, including those firms that neither create nor destroy jobs at all. Employees in approximately 7 percent of jobs were engaged in firms which were either new births, with a further 7 percent in firms that ceased to employ staff (before adjustment for spurious birth or death).

3.5 *Distribution of Employers by Firm Growth Rate and Size*

Figure 5

Distribution of Employers by Firm Growth Rate and Size



The graph above shows the composition of job creation and destruction rates by firm size. This shows that 28 percent of firms exhibit no job growth or destruction at all, the majority of these (25 percent) being firms which have five or fewer employees. Recalling that approximately two-thirds of firms have five or fewer employees, it may be seen that smaller firms are disproportionately represented in both births and deaths, but also in the count of firms that have zero job growth or destruction. The latter effect may be due to the fact that small firms cannot achieve small values of either job creation or destruction; a firm that has two employees can only lose or gain half of its workforce. Therefore, smaller firms may instead achieve small adjustments in their output without changing the size of their workforce. The former effect in relation to firm death may represent the increased vulnerability of small firms to shocks or adverse economic conditions.

3.6 Summary of Worker Flows, Job Flows and Net Employment Change

Values and distributions of job flows outlined in the above tables and graphs may be further broken down into flows of employees between firms. That is, to decompose net job change at the firm level, whether creation or destruction, into the movements of individuals which lead to those job flows. International studies have shown that, while relatively small changes in net employment over time may be underpinned by comparatively large job flows, those job flows themselves may be further derived from still larger movements of individuals between firms.

The ability of LEED to produce worker flow statistics is based on the unique identification in the input data of both employers and employees, as well as the employer-employee link. It has not been previously possible to produce these statistics in New Zealand.

3.6.1 Annual Summary

All tables below show the mean annual values of job accession and separation rates (those employees who either commenced a new job, or left an existing position respectively, expressed as a proportion of total employment), averaged between the three annual snapshots. Once again,

it must be noted that, as with job flows, worker flows will be upwardly biased due to spurious firm birth and death caused by administrative changes.

Table 8

Worker Flows, Job Flows and Net Employment Change

Job and worker flows All period					
jobs	access_ mvt	separate_ mvt	create_ mvt	destruct_mvt	net_mvt
1625800	39.6%	(37.0%)	15.8%	(13.3%)	2.5%

Table 9

Worker Flows, Job Flows and Net Employment Change by Firm Size

Job and worker flows All period sizegroup						
sizegroup	jobs	access_mvt	separate_mvt	create_mvt	destruct_mvt	net_mvt
b) 1-5	207800	49.4%	(45.7%)	32.2%	(28.4%)	3.8%
c) 6-9	105300	48.4%	(44.5%)	25.8%	(22.0%)	3.9%
d) 10-19	161900	46.9%	(42.7%)	21.5%	(17.3%)	4.2%
e) 20-49	193400	44.6%	(42.2%)	17.0%	(14.6%)	2.4%
f) 50-99	127200	43.1%	(41.0%)	14.8%	(12.7%)	2.1%
g) 100+	830100	32.9%	(31.1%)	9.3%	(7.5%)	1.8%

It may be seen that, while average net employment change over the period covered by the EMS data was 2.5 percent, achieved by job creation of 15.8 percent and job destruction of 13.3 percent, the proportion of employees commencing a new job or leaving an existing position is higher still. Over the period of study, 39.6 percent of jobs will be filled by a 'new' employee after 12 months, while employees who occupy 37.0 percent of existing jobs will have resigned.

As with job flows, worker turnover rates are higher among smaller firms, and decreases monotonically as firm size increases.

3.7 Decomposing Worker Flows

Attention is now turned to further determining which types of firm are contributing to flows of employees between jobs.

Figure 6 below shows the count of employee accessions and separations, as well as the job creation or destruction that represents the net of these worker flows, by magnitude of firm job flow. Values shown represent the mean of annual snapshots.

Figure 6

Count of Worker Flows and Job Flows

By rate of job flow

The horizontal axis in the above graph represents the rate of job flows by firm whether job creation (above zero) or destruction (below zero).

Contrasting this graph with figure 3 shows that, while most of the contribution to job flows results from firms which have relatively large flow rates, in particular those firms which are births or deaths, most of the contribution to worker flows occurs in firms that are comparatively stable. There is also a bias in the distribution of worker flows toward those firms that exhibit net job growth. It may also be seen that job flows also occur in firms that remain completely stable over time; over a 12-month period on average 39,000 employees will leave, as well as commence jobs in firms that remain identical in size.

We know then that most worker flows occur in relatively stable firms. Tables 10 and 11 show the contribution to total accessions and separations, respectively, by size of firm in the current period, and size of firm 12 months later.

Table 10

Contribution to Total Accessions by Firm Size in Consecutive Periods

Proportion of total accessions by firm size in period 1 vs firm size in period 2											
Firm size in period t	Zero employees in t+1	1 employee in t +1	2 employees in t +1	3 employees in t +1	4 employees in t +1	5 employees in t +1	6 to 10 employees in t +1	11 to 20 employees in t +1	21 to 100 employees in t +1	101+ employees in t +1	total
a) 0	.	1.9%	1.5%	1.1%	1.0%	.8%	2.7%	2.4%	3.1%	3.4%	17.9%
b) 1	.0%	**.6%**	.9%	.4%	.2%	.1%	.2%	.1%	.0%	.0%	2.6%
c) 2	.0%	.1%	**.6%**	.8%	.4%	.2%	.3%	.1%	.0%	.0%	2.7%
d) 3	.0%	.0%	.2%	**.6%**	.6%	.4%	.5%	.1%	.0%	.	2.5%
e) 4	.0%	.0%	.1%	.2%	**.5%**	.5%	.7%	.1%	.1%	.	2.3%
f) 5	.0%	.0%	.0%	.1%	.2%	**.4%**	1.1%	.2%	.0%	.	2.3%
g) 6-10	.0%	.0%	.0%	.1%	.2%	.4%	**4.5%**	2.2%	.3%	.0%	7.8%
h) 11-20	.0%	.0%	.0%	.0%	.0%	.0%	.7%	**5.8%**	2.2%	.1%	8.9%
i) 21-100	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.6%	**15.8%**	1.9%	18.4%
j) 101+	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.4%	**34.7%**	35.1%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

Table 11

Contribution to Total Separations by Firm Size in Consecutive Periods

Proportion of total separations by firm size in period 1 vs firm size in period 2											
Firm size in period t	Zero employees in t+1	1 employee in t +1	2 employees in t +1	3 employees in t +1	4 employees in t +1	5 employees in t +1	6 to 10 employees in t +1	11 to 20 employees in t +1	21 to 100 employees in t +1	101+ employees in t +1	total
a) 0	.	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%
b) 1	1.9%	**.6%**	.2%	.1%	.0%	.0%	.0%	.0%	.0%	.0%	2.8%
c) 2	1.4%	.8%	**.7%**	.3%	.1%	.0%	.0%	.0%	.0%	.0%	3.3%
d) 3	1.1%	.4%	.7%	**.6%**	.3%	.1%	.1%	.0%	.0%	.	3.3%
e) 4	.9%	.2%	.3%	.6%	**.5%**	.3%	.2%	.0%	.0%	.	3.1%
f) 5	.8%	.1%	.2%	.3%	.5%	**.4%**	.5%	.0%	.0%	.	2.8%
g) 6-10	2.6%	.2%	.3%	.4%	.6%	.9%	**4.7%**	1.0%	.0%	.0%	10.8%
h) 11-20	2.3%	.1%	.1%	.1%	.1%	.1%	1.8%	**5.9%**	1.0%	.0%	11.5%
i) 21-100	3.7%	.1%	.1%	.1%	.0%	.1%	.3%	1.5%	**15.5%**	.7%	22.1%
j) 101+	4.6%	.3%	.0%	.0%	.3%	.2%	.1%	.1%	1.4%	**33.4%**	40.4%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

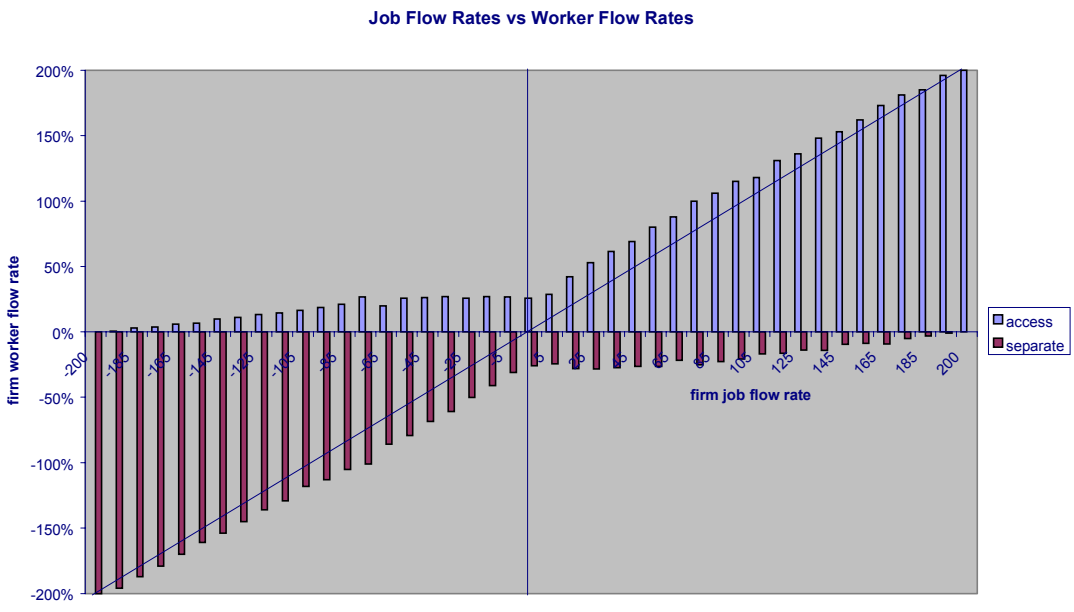
Bold cells in the above tables represent those firms that stay either exactly the same size, or in the same size group, between consecutive periods. For example, it may be seen that 34.7 percent of all accessions occur in firms that have more than 100 employees in both current and successive periods.

The tables confirm that most worker flows take place in firms that are relatively stable, and furthermore demonstrate that more than half of both accessions and separations occur in firms that have 20 or more employees, with more than one-third occurring in firms with more than 100 employees. Although larger firms tend to have lower worker flow rates than smaller firms, the high proportion of total employment in larger firms results in a high contribution to total worker flows.

Figure 7, below, summarises firm accession and separation rates by rate of firm job flow.

Figure 7

Rates of Accession and Separation by Rate of Job Flow



Both the horizontal axis, as well as the diagonal line, represent the rate of job flows by firm, whether creation (above zero) or destruction (below zero). This diagonal also represents the difference between the rates of accession and separation shown.

It may again be seen that even completely stable firms exhibit employee accessions and separations, equal to approximately 25 percent of their workforce. This appears to be inherent in firms until they begin to either rapidly grow or shrink. That is, it may be seen that firms continue to hire new staff at a rate of approximately 25 percent of their workforce, until their rate of job destruction exceeds 100 percent, and conversely tend to lose staff at a rate of approximately 25 percent until their rate of job creation exceeds 100 percent.

3.8 Annual Summary by Sex of Employee and Firm Size

Tables 12 and 13, below, show the annual rates of worker flows broken down by sex, first over all firms and second by firm size.

Table 12

Worker Flows, Job Flows and Net Employment Change by Sex

Job and worker flows All period sex						
sex	jobs	access_ mvt	separate_ mvt	create_ mvt	destruct_mvt	net_mvt
M	820600	38.8%	(36.4%)	17.7%	(15.4%)	2.4%
F	805100	40.3%	(37.6%)	16.0%	(13.3%)	2.7%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

Table 13

Worker Flows, Job Flows and Net Employment Change by Sex and Firm Size

Job and worker flows All period sex sizegroup							
sex	sizegroup	jobs	access_mvt	separate_mvt	create_mvt	destruct_mvt	net_mvt
M	b)1-5	111000	50.9%	(46.8%)	35.7%	(31.6%)	4.2%
M	c)6-9	54400	47.6%	(43.8%)	28.2%	(24.3%)	3.9%
M	d)10-19	85000	45.0%	(41.2%)	22.8%	(18.9%)	3.9%
M	e)20-49	103400	42.3%	(40.1%)	18.1%	(15.8%)	2.3%
M	f)50-99	67600	40.9%	(38.6%)	15.5%	(13.2%)	2.3%
M	g)100+	399200	31.8%	(30.4%)	10.6%	(9.2%)	1.4%
F	b)1-5	96800	47.7%	(44.4%)	33.1%	(29.8%)	3.3%
F	c)6-9	50900	49.2%	(45.3%)	27.9%	(24.0%)	3.9%
F	d)10-19	76900	49.0%	(44.4%)	24.0%	(19.4%)	4.6%
F	e)20-49	90000	47.1%	(44.6%)	18.8%	(16.2%)	2.6%
F	f)50-99	59600	45.6%	(43.7%)	15.9%	(14.0%)	1.9%
F	g)100+	430900	33.9%	(31.7%)	8.8%	(6.6%)	2.2%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals

The above tables highlight a number of points:

- Males occupy more filled jobs than females in all firm sizes except for those firms that have more than 100 staff.
- At the total level, there is a small difference between the sexes in the proportions of employees who commence a new job or leave an existing position, with worker flows being slightly higher for female employees.
- However, while worker flows are higher, the rates of job creation and destruction in those jobs that are filled by females are slightly lower than for jobs filled by males.
- Job creation and destruction decreases monotonically as firm size increases for both sexes. However, while accessions to new jobs and separations from existing jobs occupied by males decreases monotonically as firm size increases, this is not the case for jobs occupied by females. Female staff engaged in jobs in firms which have between six and nine employees have higher rates of accessions and separations than in both larger, and smaller firms.
- In addition, while the rates of accession and separation are higher for female staff than male staff at the total firm level, this is not true for all sizes of firm. Specifically, females are slightly less likely to commence employment in firms of five or fewer employees, as well as slightly less likely to leave an existing job in a firm of that size.

These points may indicate that the types of jobs filled by females differ from those filled by males. This may be due to a number of factors; for example, the industrial compositions of those firms at which females tend to be employed. Future work may decompose this table by various firm characteristics in order to better understand these patterns.

3.9 Annual Summary by Age of Employee

Tables 14 and 15, below, show the annual rates of worker flows broken down by age of employee, first over all firms and second by firm size. Note that job flows and net employment change are not shown in these tables, as there will be changes in the counts of jobs filled by

persons in each age group that are due not to flows of individuals between jobs, but simply a result of workers getting older.

Table 14

Worker Flows by Age of Employee

Job and worker flows All period age_grp			
age_grp	jobs	access_ mvt	separate_ mvt
15-24	334100	63.4%	(56.3%)
25-44	802600	38.2%	(36.3%)
over 45	489100	25.2%	(24.8%)

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

Table 15

Worker Flows by Age of Employee and Size of Firm

Job and worker flows All period age_grp sizegroup				
age_grp	sizegroup	jobs	access_mvt	separate_mvt
15-24	b)1-5	47900	67.1%	(60.9%)
15-24	c)6-9	25300	67.2%	(60.4%)
15-24	d)10-19	40400	67.7%	(59.2%)
15-24	e)20-49	45200	66.5%	(60.1%)
15-24	f)50-99	29200	65.7%	(59.4%)
15-24	g)100+	146200	58.9%	(51.6%)
25-44	b)1-5	98200	49.0%	(45.4%)
25-44	c)6-9	51100	47.0%	(43.5%)
25-44	d)10-19	77700	44.6%	(41.7%)
25-44	e)20-49	95000	42.5%	(41.1%)
25-44	f)50-99	61700	41.1%	(40.0%)
25-44	g)100+	418900	32.1%	(30.7%)
over 45	b)1-5	61700	36.2%	(34.0%)
over 45	c)6-9	28900	34.1%	(32.1%)
over 45	d)10-19	43800	31.7%	(29.1%)
over 45	e)20-49	53300	29.4%	(28.5%)
over 45	f)50-99	36300	28.1%	(27.7%)
over 45	g)100+	265100	19.4%	(20.0%)

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

The above tables show that older employees (those over 45 years of age) have more stable employment patterns than 'prime age' employees (aged between 25 and 44 years), who in turn have more stable employment patterns than those under 24 years of age. This pattern is repeated across all sizes of firm, and worker flow rates decrease monotonically as firm size increases with one exception. Employees aged between 15 and 24 are more likely to accede to, and less likely to separate from, firms that have between 10 and 19 employees than to either smaller or larger firms.

In addition, worker flows by firm size of workers in the 15 to 24 age group are less variant, depending on firm size than other age groups, especially those engaged in firms of fewer than 100 employees.

The above patterns may be due to the characteristics of those firms that tend to employ younger staff; for example, the industry in which the firm is engaged. Alternatively, the patterns may be indicative of the employment preferences of younger workers. These issues have not been investigated in this paper.

It should be noted that the rate of accessions in the younger age group will also comprise employees who enter the labour force for the first time, and that separations in the older age group will encompass employees who leave the labour force permanently.

3.10 Counts and Earnings of Multiple Job Holders

The tables below present some summary statistics on numbers of employees and gross earnings by count of jobs held. Tables are presented at the all economy level only; future work will decompose results by age, sex, industry or other derived variables.

All values presented in the tables relating to multiple job holding are means of three annuals snapshots taken in February of each year.

The tables should be interpreted in the light of the data weaknesses presented at the start of this paper; specifically, that the count of multiple job holders will be upwardly biased by the extent of persons who leave their job and commence a new job in a given reference month. Preliminary investigations suggest that the value of this bias will be in the order of 3 to 4 percent of all jobs.

It may also be expected that there will be some downward bias in mean monthly earnings measures to the extent that earnings by individuals in the reference month represent less than a full month of employment. For example, when an individual either commences a new job or leaves an existing job in a given month, the earnings in that period will almost always represent a partial month. Future work may restrict the population of employees for the purposes of this table to those individuals who are employed throughout a given reference period.

Table 16 summarises the counts of individuals by number of jobs held and the proportion of each group to the total number of individuals in employment.

Table 16

Counts of Multiple Job Holders

Multiple Job Holding Mean of Annual Snapshots		
Number of employers that employee is engaged at	Number of persons	Proportion of total
1	1365900	92.4%
2	101600	6.9%
3	9300	.6%
>=4	2100	.1%

It may be seen that the majority of individuals are engaged in a single job, with most of the remainder holding two jobs.

Table 17

Earnings of Multiple Job Holders Count of Jobs Held by Job

Mean earnings per job by count of jobs held					
Number of Jobs Held	First Job	Second Job	Third Job	Fourth and Subsequent Jobs	Total Earnings
One Job	2720	.	.	.	2720
Two Jobs	1770	500	.	.	2280
Three Jobs	1390	530	220	.	2140
Four or More Jobs	1160	530	320	190	2190

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

The above table shows that in general total individual earnings fall as the count of jobs held increases, and that earnings in each job held also falls as the count of jobs increases.

The exception is that those persons who hold four or more jobs have higher earnings than those who hold three jobs, and that earnings in the third job are higher than in the case of individuals who hold three jobs. However, earnings of persons holding four or more jobs should be treated with caution as:

1. The count of persons holding four or more jobs is low.
2. There are some individuals who hold a very high (more than 10) count of jobs, and whose income is significantly higher than average.
3. Investigation of selected unit records suggests that a high proportion of payments to these individuals may in reality reflect the payment of either directors' fees or shareholder salaries, and not payments related purely to the provision of labour.
4. Other individuals who appear to hold a high number of jobs appear to be casual agricultural employees who are paid, with PAYE deducted, from a number of different farms in the reference month. In some cases, their individual job histories suggest that these employees may also from time to time receive end-of-season or similar 'settle ups'. A number of these may be received in a given month, which appear as multiple jobs.

3.11 Earnings of Accessions versus Earnings of Continuing Employees

The table below shows the mean gross earnings of full-month accessions versus mean gross earnings of full-month continuing employees.

The population of interest is restricted to full-month employees only: those employees who were engaged throughout the entire mid quarter reference month. This is done to ensure that earnings are not downwardly biased by persons who work less than the full month, as this will disproportionately affect new accessions whose first earnings payment will almost always represent a partial month.

Full-month employment is proxied by the condition that all employees, whether accessions or continuing employees, engaged by a given employer in the reference month, must also have been engaged by that same employer in both the previous and successive mid-quarter reference months. Implicit in this is the assumption that there will be no intra-month worker flows. As a result, the accessions column in the table below refers not to those employees who commenced a new job in the current reference period, but to those who were accessions in the previous mid-quarter reference month.

Table 18

Earnings of New Accessions versus Earnings of Continuing Employees

By period

Mean gross earnings of full month employees who were accessions in previous quarter versus full month continuing employees			
return_ period_date	accessions	continuing	ratio
NOV99	2050	2810	73%
FEB00	1940	2870	68%
MAY00	2060	3030	68%
AUG00	2130	3130	68%
NOV00	2240	3110	72%
FEB01	2110	2850	74%
MAY01	2250	3260	69%
AUG01	2110	3230	66%
NOV01	2190	3010	73%

The above table shows that, on average, persons who commence a new job will be paid approximately two-thirds of the earnings of persons in continuing employment; this is broadly consistent with the findings of the U.S. Census Bureau Longitudinal Employer-Household Dynamics programme. This ratio appears to be relatively consistent over time without immediately apparent seasonal effects, although the time series available is insufficient to be certain.

Many different factors may be contributing to this difference; for example, while it is likely that new entrants to any given job will in reality be paid less than continuing staff in a similar role, the earnings of accessions in the above table may also be affected by a high proportion of temporary labour such as pieceworkers, by casual part-time employment, or by prevailing earnings rates in certain industries. Future work may attempt to quantify some of these factors by decomposing this table by industry, by sex, and by tenure of new full-month accessions.

3.12 Job Tenure of New Accessions

The table below shows a count of new jobs commenced by duration of tenure and start period. While the table is presented for three start periods one year apart, the value of accessions in each period is derived from quarterly, not annual, worker flows. That is, the value of accessions in February 2001 represents all filled jobs in that reference month that were not filled by the same employee in the mid-month of the previous quarter – November 2000. Quarterly worker flows were used so that a longer time series (though still short) could be seen.

More than any other table, job tenure statistics are likely to be adversely affected by missing or incorrect employee IRD identifiers. This is because, for an entire job history to be correct, the relevant IRD identifier must be correct not only in the current period, but in all periods in which the job is in existence. Therefore, the table is to be taken as indicative only of the kind of output that is able to be produced.

As with other outputs, the table will in future be able to be broken down by characteristics of both the employer, such as industry or region, as well as of the employee such as age, sex, or income distribution. Statistics such as these on tenure, as well as data on transitions from receipt of social assistance benefit to paid employment (refer Carroll and Wood 2003) will provide new data sources for analysing individuals tenuously attached to the labour market.

Table 19

Job Tenure of New Accessions

February Reference Months

Duration of job by quarter of accession			
duration	FEB00	FEB01	FEB02
1 qtr	70600	66300	68600
2 quarters	77000	75500	.
3 quarters	32900	31300	.
4 quarters	17100	17700	.
5 quarters	16300	15500	.
6 quarters	10000	.	.
7 quarters	7100	.	.
8 quarters	5200	.	.
9 quarters	5200	.	.
Not finished	36900	68900	215800

The table shows, for example, that 70,600 of employees who acceded to a new job in February 2000 separated from that job within one quarter. At the other end of the scale, 36,900 were still in that job more than nine quarters later – job duration is unknown due to the length of time series available.

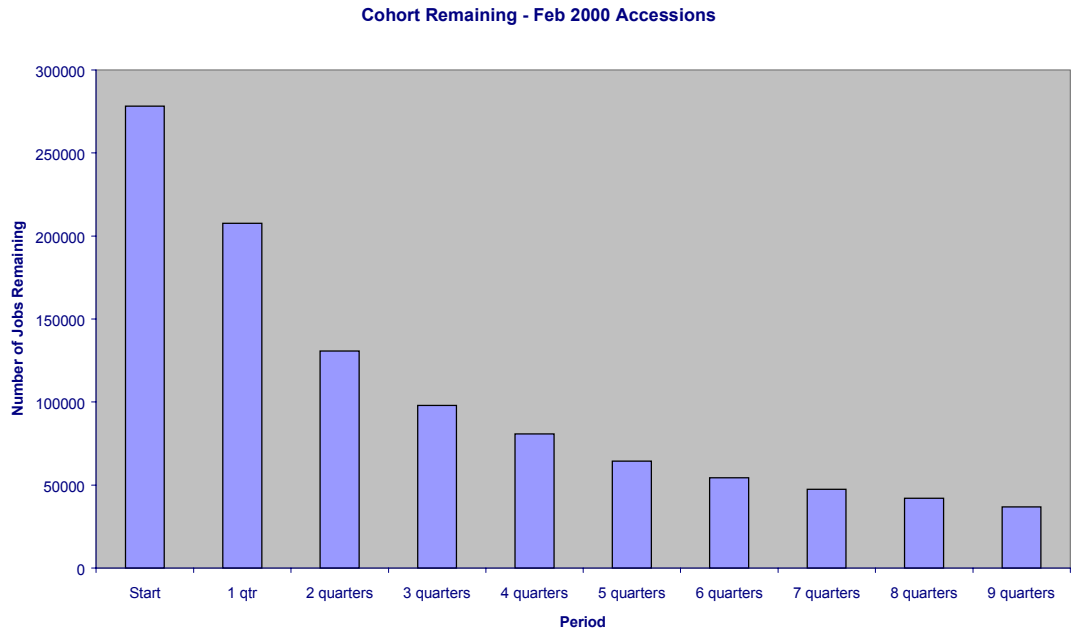
Interpretation of the table is risky due to the factors mentioned above and should be treated with caution. Nevertheless, a pattern may be seen, which is highlighted in the following graph and table.

The above table is re-expressed to show the number of jobs from the February 2000 accession cohort which remain filled by the same employer, by quarter.

Figure 8

February 2000 Accessions Cohort:
Number of Jobs Filled by Same Employee

By quarter



It may be seen that a substantial proportion of the February 2000 accessions cohort separate from their new jobs in a relatively short period of time, between one and three-quarters after accession. Over time, however, the remaining cohort becomes more stable with proportionally fewer separations in each successive quarter. This may be seen in the table below which re-expresses Table 20 to show the number of separations occurring in each quarter as a proportion of the original accessions cohort remaining as at the end of the previous quarter; this may be interpreted as the probability of an employee leaving their job.

Table 20

February 2000 Accessions Cohort: Probability of Separation

By quarter

Quarters After Accession	Probability of Separation
Start	0%
1 qtr	25%
2 quarters	37%
3 quarters	25%
4 quarters	18%
5 quarters	20%
6 quarters	15%
7 quarters	13%
8 quarters	11%
9 quarters	12%

From this table it may be seen that the probability of an employee separating from their job decreases as the length of tenure in that job increases.

Recall that approximately 4 in every 10 filled jobs will not be occupied by the same employee after 12 months. This implies, ignoring intra-year flows, that employees will separate from around 10 percent of jobs per quarter. However, it appears that there is considerable variation in the propensity of individuals to separate from their current employer, with some individuals having a very tenuous attachment to the labour force, characterised by short job tenure and a consequently high probability of separation, with others having more stable employment patterns. Future work may compare the employment patterns of new accessions with those of individuals in longer term continuing employment, as well as seasonal and other effects. Ideally this work would use a longer time series than currently available.

3.13 Firm Size Transition Rates

The following tables examine the dynamics of firms and their employment over the three annual snapshots of February 2000, February 2001 and February 2002. These summary results provide some insight into patterns of firm growth and decline over this short time period.

The tables cover all enterprises in existence and employing staff in February 2000. Other studies (Carroll et al, 2002a, 2002b) have contrasted the dynamics of continuing firms with those of new births. This comparison has not been undertaken in Stage 1 of LEED as evidence suggests that a significant amount of employment in firms which appear to be new births using Stage 1 definitions may in fact represent employment in continuing enterprises. Refer Statistics New Zealand (2003) and Kelly (2003).

The table below presents the counts of firm by firm size in February 2000 and February 2002, expressed as proportions of start period firm composition and rates of transition to closing period firm size.

Table 21

Firm Size Transition Rates February 2000– February 2002

Firm Size Transition Rates 29 Feb2000- 28 Feb 2002 29 Feb 2000 stock cohort								
employer size group at distribution Feb2000	Feb2000 distribution of units	1 to 5 employees	6 to 9 employees	10 to 19 employees	20 to 49 employees	50 to 99 employees	over 100 employees	dead
b) 1-5	70%	**62%**	5.3%	.8%	.1%	.0%	.0%	32%
c) 6-9	13%	21%	**42%**	15%	.9%	.0%	.0%	20%
d) 10-19	9.5%	4.6%	13%	**53%**	11%	.2%	.1%	17%
e) 20-49	4.9%	1.7%	1.4%	11%	**63%**	7.3%	.3%	15%
f) 50-99	1.4%	.5%	.3%	1.0%	13%	**61%**	11%	13%
g) 100+	1.2%	.9%	.2%	.4%	1.3%	5.8%	**81%**	11%
x) At End	100%	47%	10%	8.2%	4.5%	1.3%	1.1%	28%

Column 1 shows the cross-section distribution of employers in February 2000, showing, for example, that 70 percent of firms in that period had five or fewer employees. The rows in the table represent firm size transition rates between the opening and closing periods. For example, row 1 shows that of all the firms that had between one and five employees in February 2000, 62 percent also had between one and five employees two years later, 5.3 percent had between six and nine employees, and 32 percent had ceased to employ staff at all. (Note that other investigations imply that at least 10 and possibly over 20 percent of firms in the final 'dead' column may in fact represent continuing employers.) The final row shows the distribution by size in February 2002 of those firms that were in existence in February 2000.

The bold figures in the matrix represent the proportions of firms that remain in the same size group over the two-year period. They show that, in general, firms are more likely to remain in the same size group than to either expand or contract. The exception is those firms that had between six and nine employees in February 2000; these are more likely to contract or to 'die'. The table also shows that while in general existing firms tend to remain in the same size group, they are nevertheless more likely to contract or to die than they are to expand. As job creation exceeded job destruction over the period in question, the implication is that the tendency toward contraction or firm death is 'propped up' by new firm births.

3.14 Longitudinal Employment Dynamics

The table below shows the annual employment dynamics of this cohort of firms.

Table 22

Longitudinal Employment Dynamics

February 2000 Cohort

Longitudinal Employment Dynamics 2000 stock cohort Employment relative to start period employment				
Employer size group at start period	Start period employment	29FEB00	28FEB01	28FEB02
b) 1-5	206200	100%	90.7%	86.2%
c) 6-9	119500	100%	87.9%	83.0%
d) 10-19	166200	100%	90.6%	86.2%
e) 20-49	192600	100%	92.6%	88.9%
f) 50-99	126500	100%	97.0%	94.2%
g) 100+	760500	100%	97.4%	95.9%
x) All	1571500	100%	94.5%	91.6%

Note: Figures have been rounded and discrepancies may occur between sums of components and totals.

The first column shows the distribution of employment in firms in February 2000. This shows that in the LEED data, while the proportion of firms that engage over 100 employees is small, they nevertheless account for a large proportion of employment in New Zealand.

Subsequent columns show the numbers employed in those same firms in two subsequent years, expressed as a proportion of base period employment. The overall pattern is that employment in all sizes of firm fell each year for this cohort of firms. In general, small firms tend to destroy jobs (to contract or to die) at a higher rate than larger firms. Once again, however, it must be borne in mind that a high proportion of the loss in jobs due to firm death (approximately 37 percent) is likely to be spurious and to in fact represent continuing employment, and that the distribution of spurious flows resulting from this issue is not known by firm size.

The findings above are consistent with job flows presented earlier in this paper, in that smaller firms tend to have a higher rate of job destruction and firm death than larger firms.

4. Matters Outstanding

As stated in the introduction to this paper, the outputs presented herein have been produced in order to demonstrate the kinds of outputs that are able to be derived from the LEED source data, and the feasibility of doing so. While enabling some useful insights into the New Zealand labour market, none of the prototype outputs is produced to the standard required of official statistics. Some of the outstanding data issues that would need to be addressed before official outputs could be produced, as well as some likely directions of future work, are summarised below:

- The source data contains no information on mergers, acquisitions and splits; all events where an apparent firm birth and/or death may occur, but which in reality represent administrative changes to a continuing entity. While trial methods have been developed to deal with this issue, they have not been incorporated into this paper.
- The definition of employment that underlies all outputs in this paper rests on the receipt of non-zero gross earnings by employees at any time in the return period chosen. This is conceptually different from the point-in-time measures that underlie many official statistics, and

may be expected overstate both the number of jobs, and the number of multiple jobholders. Again, methods have been developed to deal with the issue, but not incorporated into the paper.

- Outputs are, in theory, able to be aggregated by any characteristics related to either individuals or to their employers. Employer characteristics of interest include industry, region, and institutional sector. The most appropriate source of many of these variables is the SNZ Business Frame. Therefore, production of many desired outputs requires the linking of employer and employee information derived from the EMS with equivalent statistical units on the BF. This linking mechanism has not yet been finalised.
- Approximately 2 percent of EMS records have an invalid or missing *employee_ird_number* and as yet unknown proportion may have an identifier that is valid, but in reality belongs to another individual. It may be expected that such errors will upwardly bias many flow statistics, and will downwardly bias measures of job tenure. Solving problems with employee identifiers is likely to involve the use of specialised record linking and repair software, and has not been implemented in the source data used in this report.
- The collection unit that underlies all outputs in this paper is the legal entity that files the EMS tax return, identified by the *employer_ird_number*. In most cases, this will correspond to a single enterprise, which is in turn composed of a single geographic unit on the Business Frame. This unit represents the lowest level of firm employment information available on the BF, and is conceptually consistent with the plant level analysis undertaken by Davis and Haltiwanger. In some cases, however, a single EMS return may represent a legal entity that has many geographic units, or in some cases more than one legal entity. To the extent that workers move between physical locations in multi-unit firms, and between firms in multi-firm groups, job and worker flows will be understated. Furthermore, to the extent that the industrial classification and region of the filing firm(s) differs from that of component geographic units, distortions will result in summary statistics by industry and region. In addition, the ability to produce business demography by firm size measures will be compromised. It is intended that future work will develop methods to disaggregate multi-unit and multi-firm returns.
- The definition of an employee used in this paper rests on, among other things, the receipt of non-zero gross earnings in a given EMS return period. This differs from most official statistical outputs, which rely on a criterion of employment for a specific length of time. Future work will look at ways of applying definitions that are conceptually closer to those used in official outputs.
- As this paper forms part of a series of feasibility studies, and is intended to show the possibility of LEED to produce new outputs and enhance existing statistics, a balance was struck between the level of detail and associated analysis produced, and the number of outputs presented. As such, many outputs are presented at the total economy level only, and analysis has been confined to the broad patterns that are readily discernible. Depending on user needs and feedback, many of these tables may be produced at a finer level of aggregation in future.
- The EMS return covers all persons engaged as paid employees, with partial coverage of persons who are self-employed. The majority of self-employed individuals do not pay their income tax at source, but instead file an end of year return known as the IR3. However, as this return is quite broad in scope and also covers the recipients of non-source-deducted income other than that related to self-employment, integration with the EMS is not straightforward. Stage 2 of LEED will investigate methods to combine these two data sources.

5. References

Bycroft C (2003). *Record Linkage in LEED: An Overview*, unpublished paper, Analytical Support Section, Statistics New Zealand, Christchurch.

Carroll N and Wood J (2003). *Preliminary Research into Sustainable Employment Measures using the Linked Employer-Employee Data,* unpublished paper, Statistics New Zealand, Wellington.

Carroll N, Hyslop D, Maré D, Timmins J and Wood J (2002a). "The turbulent labour market," paper prepared for the New Zealand Association of Economists' Conference, Wellington, 25-27 June 2002, http://nzae.org.nz/files/%2316(2)-CARROLL-TEXT.PDF [22 August 2003].

Carroll N, Hyslop D, Maré D, Timmins J and Wood J (2002b). "An analysis of New Zealand's business demographic database" *New Zealand Economic Papers,* 36(1), 59–62.

Davis S and Haltiwanger J (1992). "Gross job creation, gross job destruction and employment reallocation" *Quarterly Journal of Economics*, 107:3, 819-863.

Graham P (2003) *Specifications for Editing and Imputation of Dates of Birth and Sex,* unpublished paper, Statistics New Zealand, Christchurch.

Kelly N (2003). *Repairing EMS Employer Longitudinal Links*, unpublished paper, Statistics New Zealand, Christchurch.

Statistics New Zealand (2003a). *LEED Stage One Feasibility Report,* unpublished paper, Statistics New Zealand, Christchurch.

Statistics New Zealand (2003b). *LEED Project: Base Concepts and Definitions*, unpublished paper, Statistics New Zealand, Christchurch.

US Bureau of the Census, Longitudinal Employer-Household Dynamics programme, http://lehd.dsd.census.gov/ [8 September 2003].