# Toward a Model of Firm Productivity Dynamics

David Law, Bob Buckle and Dean Hyslop

New Zealand Treasury
Working Paper 06/11

November 2006

THE TREASURY
Kaitohutohu Kaupapa Rawa

**NZ TREASURY WORKING PAPER 06/11**

# Toward a Model of Firm Productivity Dynamics

**MONTH/YEAR**

November 2006

**AUTHORS**

David Law
The Treasury
PO Box 3724
Wellington
New Zealand

| | |
|---|---|
| Email | David.Law@treasury.govt.nz |
| Telephone | 64-4-9176233 |
| Fax | 64-4-4731151 |

Bob Buckle
The Treasury
PO Box 3724
Wellington
New Zealand

| | |
|---|---|
| Email | Bob.Buckle@treasury.govt.nz |
| Telephone | 64-4-9176252 |
| Fax | 64-4-4990992 |

Dean Hyslop
The Treasury
PO Box 3724
Wellington
New Zealand

| | |
|---|---|
| Email | Dean.Hyslop@treasury.govt.nz |
| Telephone | 64-4-9176157 |
| Fax | 64-4-4731151 |

**ACKNOWLEDGEMENTS**

We thank Richard Fabling, Mario Di Maio, Dave Maré, seminar participants at The Treasury and participants at the 2006 Conference of the NZ Association of Economists for comments and discussions. We are grateful to Preeti for her capable assistance with data compilation. Access to the data used in this study was provided by Statistics New Zealand in a secure environment designed to give effect to the confidentiality provisions of the Statistics Act, 1975. The views, opinions, findings and conclusions or recommendations expressed in this document are strictly those of the authors, and do not necessarily represent and should not be reported as those of The Treasury or Statistics New Zealand.

**NZ TREASURY**

New Zealand Treasury
PO Box 3724
Wellington 6008
NEW ZEALAND

| | |
|---|---|
| Email | information@treasury.govt.nz |
| Telephone | 64-4-472 2733 |
| Website | www.treasury.govt.nz |

# Abstract

A common finding from international research on firm productivity dynamics is that within-firm productivity dynamics tend to dominate the effects of firm entry and exit on aggregate productivity. The aim of this paper is to explore the suitability of Statistics New Zealand's Business Demography (BD) and Goods and Services Tax (GST) data as a basis for modelling within-firm productivity dynamics. The paper first analyses and describes the cross-sectional and time-series properties of sales, purchases and a value-added measure of labour productivity. Cross-sectional results reveal a great deal of heterogeneity in average sales, purchases and labour productivity both across and within industries and cohorts. Univariate time-series properties of these variables are remarkably similar and sales and purchases are highly correlated contemporaneously. Transition probabilities are also calculated for movement of firms between quartiles of the labour productivity distribution over varying lengths of time. In order to understand the processes driving the data, a simple statistical model for sales, purchases and value-added per unit of employment is developed to calibrate to the stylised empirical facts. The model does a remarkably good job at mimicking the properties of the BD and GST data.

**JEL CLASSIFICATION** D21: Firm behaviour
L00: Industrial organisation – General
O12: Microeconomic analyses of economic development

**KEYWORDS** Firm Productivity; Labour Productivity; Firm Dynamics; New Zealand GST Data; New Zealand Business Demography; Firm Value-added.

# Table of Contents

# List of Tables



# List of Figures

# Toward a Model of Firm Productivity Dynamics

## 1 Introduction

Firm unit record data have proved to be a rich source of information for understanding how firm dynamics contribute to industry-level and economy-wide productivity growth. The survey by Bartelsman and Doms (2000) emphasised in particular the importance of the development of longitudinal micro-level data sets which have enabled documentation of micro-level productivity growth, examination of the factors associated with firm productivity and more rigorous exploration of causal relationships. The aims of this paper are to describe a longitudinal New Zealand firm unit record dataset compiled from various sources including goods and services taxation data, to explore the properties of these data, to draw out some inferences about New Zealand firm productivity, and to use the revealed statistical properties to guide the specification of a model of firm productivity dynamics.

Some of the inferences about New Zealand firm productivity performance that can be drawn from the firm unit record data analysed in this paper are the degree of heterogeneity of productivity performance across firms, the degree of persistence in firm productivity and the extent to which firms move over time within the distribution of firm productivity. The paper explores some relationships between firm entry, exit, size and productivity performance. It also calibrates a simple stylised statistical error components model that is (broadly) consistent with the stylised facts pertaining to the time series properties of the firm unit record data which will help inform the development of richer models of New Zealand firm productivity.

One approach to modelling firm productivity dynamics now commonly applied to longitudinal firm databases is to use techniques suggested by Griliches and Regev (1995) and by Foster, Haltiwanger and Krizan (1998). These techniques decompose total productivity growth over time into contributions from within-firm productivity growth, changes in market share, and contributions from exiting and entering firms. A common finding from this stream of research is that within-firm productivity growth tends to dominate the contributions of firm entry and exit to aggregate productivity growth. For example, Baily, Bartlesman and Haltiwanger (1996) conclude that the within-firm component accounted for almost half of the growth of total factor productivity for US manufacturing establishments during the 1980s while net entry accounted for about one quarter of that growth. The OECD (2004) found that for the manufacturing sectors of several OECD economies during the years 1987 to 1997, within firm productivity growth accounted for the bulk of overall labour productivity growth, the contribution from changes in market share varied across countries and time but was typically small, and that the net contribution from entry and exit accounted for between 20 and 40 percent of labour productivity growth. Broadly similar conclusions were drawn for New Zealand in a recent study by Law and McLellan (2005) which used exactly the same firm unit record database that will be compiled and analysed in this paper.

Understanding the dynamics of within-firm productivity growth is therefore a crucial step toward understanding industry and aggregate productivity growth. The aim of this paper is to develop an empirical model that sheds some light on within-firm productivity growth. Recognising the potential insights documented for example by Bartelsman and Doms (2000), the approach in this paper is to create a longitudinal firm unit record database to analyse within-firm productivity growth. This longitudinal database is developed from Statistics New Zealand's Business Demography (BD) and Goods and Services Tax (GST) data. The database is used to derive a measure of firm productivity. The statistical properties of the data are then analysed and applied to specify and calibrate a model of firm productivity that matches the stylised properties of the firm unit record data.

We focus particularly on the extent to which the GST-based measure of value-added we derive from these data is a useful proxy for labour productivity. As we don't observe a robust alternative firm-level productivity measure for comparison, we rely on inference from the time series statistical properties of the data. For example, assuming there is (some) persistent difference in true productivity across firms over time, in the extreme, if the year-to-year correlations in firm-level measured productivity are zero, this would suggest that it is dominated by "noise" and the measure is a very poor proxy for firm productivity. In the absence of an obvious metric to assess the absolute quality of the labour productivity measure, we instead use the calibrated model to interpret the variation in measured productivity.

The remainder of the paper is organised as follows. Section 2 describes the data and variable construction. Various summary measures for selected variables drawn from the dataset are presented in Section 3, while Section 4 provides some simple descriptive analysis including correlations between variables across time and transition probabilities. The revealed statistical properties of the data are then used to guide the calibration of a model of firm productivity dynamics. This model and its properties are described in Section 5 which also includes comparisons between the time-series properties of the BD and GST data and the calibrated model. The model of firm productivity is derived from an autoregressive model of firm sales and purchases that contains common and firm specific shocks. The model does a surprisingly good job of mimicking the properties of the BD and GST data. The final section summarises the key insights and provides some discussion of the implications for understanding firm productivity dynamics.

# 2 Background and data

This section provides a brief review of recent New Zealand literature applying firm unit record data to understand firm dynamics and productivity. We then describe the various data sources, and how they are combined and used in this study. Variable construction, procedures for dealing with missing data and weighting are discussed. The section concludes by outlining some of the potential problems with the data.

In New Zealand, several firm unit record datasets have been utilised to understand aspects of firm dynamics. Using Statistics New Zealand's Business Practices Survey (BPS), Fabling and Grimes (2003, 2004) examine the relationship between various firm practices, including innovation, and firm performance. Unit record data from the NZ Institute of Economic Research's Quarterly Survey of Business Opinion (QSBO) has been applied by Buckle and Carlson (2000, 1998) to evaluate the microeconomic foundations of business cycles, including the properties of firm pricing and output decisions. Maré and Timmins (2005) and Maré (2005) have applied Statistics New Zealand's Business Demography (BD) and Goods and Services Tax (GST) data to examine patterns of concentration, specialisation and agglomeration of firms and the association with firm productivity. BD and GST data has also been used by Mills and Timmins (2004) to evaluate the distribution of New Zealand firm size and turnover and by Dixon, Maré and Timmins (2005) to examine changes in the size distribution of firms and to assess some reasons for these changes.

These various sources of New Zealand firm unit record data are quite different, each with their own strengths and weaknesses. The BPS provides data on a wide range of variables but gives a snapshot of a firm only, has a relatively small sample size and is restricted to firms with 6 or more full-time employees. Data from the QSBO dates back to the 1960s and covers a wide range of variables. However, these data are primarily categorical. BD and GST data capture almost all firms in New Zealand over at least the last decade, but compared to the QSBO and BPS, contain significantly fewer variables.

Despite the modest number of variables contained within the BD and GST database, Law and McLellan (2005) have shown that it offers considerable potential to measure and enhance our understanding of firm productivity dynamics. This potential will be enhanced further when this database is combined with the Annual Enterprise Survey (AES). However, to date little is known about the properties of the BD and GST database or its reliability when used to form measures of labour productivity. Some potential difficulties arise for example from the fact that sales and purchases of capital goods are included in the variables that measure firms' sales and purchases. The typically infrequent and lumpy nature of capital goods transactions means that measures of value-added for individual firms could be quite volatile from period to period and could even include large negative values for some firms, with corresponding negative measures of labour productivity.

## 2.1 Description of source data

We derive a real value-added measure of firm-level labour productivity using firm-level sales, purchases and employment data supplemented by industry-level producer price indexes and hours worked data. Producer prices and hours worked data are not available at the firm level.

The primary firm-level data for sales, purchases and numbers employed come from two sources and cover the period 1994 to 2003:

1. Statistics New Zealand's Business Demography Statistics Database (BD)
2. Inland Revenue Department's Goods and Services Tax Database (GST)

The BD contains demographic and employment (both employees and working proprietors) information on enterprises (firms) from the New Zealand Business Frame that are deemed to be economically significant.[1] [2] These data are collected for mid February of each year as part of Statistics New Zealand's (SNZ) Annual Business Frame Update survey (ABFU).[3] From 1994 onwards an enterprise was deemed to be economically significant and therefore covered by the BD if it satisfied any one of the following criteria:

- The enterprise had annual GST expenses or sales greater than $30,000;
- The enterprise had more than two full time equivalent paid persons employed;
- The enterprise was in a GST exempt industry (except residential property leasing and rental);
- The enterprise was part of a group of enterprises;
- The enterprise was a new GST registration that was compulsory, special or forced (which normally means the enterprise was expected to have GST sales or expenses that exceed $30,000); or
- The enterprise was registered for GST and was involved in agriculture or forestry.

The BD industry coverage is not the same from year to year. Additional industries were included in the ABFU in some years while in other years industries were dropped. To maintain constant industry coverage over the period 1994 to 2003 it was necessary to drop enterprises in industries that were not included in the BD in every year between 1994 and 2003. This means that agriculture and livestock production, residential property leasing and rental, commercial property and leasing, child care services, residential and non-residential services, business professional and labour organisations, religious organisations, social and community groups, and sporting and recreational services industries have been excluded from the database of firms described in this study.

The Inland Revenue Department's GST data contain monthly GST sales and purchases information for all enterprises registered for GST. Enterprises can file GST returns at different frequencies over the year. Enterprises that are members of a group may file on each other's behalf and the responsibility for filing within a group can vary over time. We make no attempt to apportion GST returns across firms within a group.

In order to derive a real value-added measure of labour productivity per hour worked, the BD and GST data are then combined with the Producers' Price Indexes (PPI) and the Household Labour Force Survey (HLFS) compiled by Statistics New Zealand.

The PPI provide information on producer output and input prices at the two digit industry level. The output price indexes measure changes in the prices received by producers and input price indexes measure changes in the cost of inputs to production (excluding labour and capital costs). These two-digit industry level price indexes are used as deflators for

[1] In the final data set that we use in this paper there may be firms that do not necessarily meet any one of these criteria in a particular year but have some visible information, from GST returns for example, that enables us to retain them. So long as they meet one of these criteria at some stage during our sample period they will be included in the sample if at all possible.

[2] The BD also contains demographic information on geographical units (previously known as activity units) i.e., units engaged in one or predominately one economic activity from a single physical location or base.

[3] The ABFU survey is not sent to all firms in every year. Smaller firms may receive this survey infrequently, with some firms only having their information updated at birth. In these cases Statistics New Zealand applies a 'last known actual' rule.

firm-level sales and purchases. If the price of a firm's output is actually higher than the average industry price, then measured real gross output will be overstated. Similarly, if the price of a firm's inputs are actually higher than the average industry price, then measured real inputs purchased will be overstated.

A proxy for firm real value-added is then constructed by subtracting input price deflated purchases from output price deflated sales. Measures of value-added should ideally take account of changes in inventories of finished goods, unfinished goods and raw materials. However, data on firm level inventories are not available from the BD/GST data and therefore a measure of value-added based on firm sales and purchases is likely to be more volatile than one based on firm output. This is because inventories tend to act as a buffer and vary inversely with changes in demand (see Buckle and Meads, 1991). Just-in-time inventory methods and the growth in the relative size of the services sector may mean that this issue is becoming less significant over time.

All firm real value-added estimates are expressed in terms of per hour worked. The hours worked per firm are derived by combining employment data from the BD database with information on hours worked from Statistics New Zealand's Household labour Force Survey (HLFS).

The BD contains enterprise data on the number of full time working proprietors, part time working proprietors, full time employees, and part time employees for mid February of each year. These data could be used to construct a 'head count' measure of firm labour input by adding the number of full time working proprietors, part time working proprietors, full time employees, and part time employees. However, the total hours worked or paid for within an enterprise is the preferred measure of that enterprises labour input. This information is however not available from the BD.

The HLFS does however provide information on the average hours worked by full time and part time working proprietors and employees for three digit industries. The HLFS is a private household based survey and has greater industry coverage than the alternative Quarterly Employment Survey (QES),[4] which is a firm based survey that contains information on the average hours paid to full time and part time workers within an industry.[5] The HLFS classifies a person as a full time working proprietor or employee if they work 30 hours or more per week. A person is classified as a part time working proprietor or employee if they work less than 30 hours per week.

The four types of firm employment data from the BD are therefore combined with the three digit industry level average hours worked data from the HLFS to account for differences in the hours worked by different types of workers. Each worker type was assigned the average hours worked by the corresponding type of worker at the three digit industry level. This alternative measure of enterprise labour input assumes there is no variation in the average hours worked by different types of workers within an industry at the three digit level (although there will still be variation in labour inputs within a three digit industry because enterprises have different numbers and types of workers).

These firm data are combined with the aid of unique random firm identification numbers and industry and time classifications. The aim is for every firm in every year of its existence to have sales, purchases and employment information in order to generate

[4] The QES excludes the following industries: Agriculture and agricultural contracting, hunting and trapping, fishing, seagoing work, owning and leasing of real estate, armed forces (civilian staff are included), and domestic service in households.

[5] The HLFS was not designed to collect industry data. The QES therefore may provided a more stable hours measure. However, for consistency with Law and McLellan (2005) we continue to use the HLFS.

annual firm level measures of labour productivity. In practice however for a number of reasons firm records are often incomplete as will be discussed in the next section.

Because BD data on numbers of working proprietors and employees are recorded for mid February in each year, monthly GST sales and purchases data have been collapsed to an annual frequency for the year ending August. When forming annual GST sales and purchases for entering and exiting firms that had monthly sales and purchases data for less than a full year (which suggests these firms were operating for only part of the year), the aggregated monthly sales and purchases were annualised to ensure entering, exiting and continuing firms are analysed on a comparable basis. GST sales and purchases were annualised by multiplying the total recorded for each over any given year, centred around February, by 12/(12-n) where n is the number of months for which the firm had no GST information in that year.[6]

In principle these data sources provide scope to derive two proxies for firm output (real gross output and real value-added) and two measures of firm labour input (total number of workers and a proxy for the total number of hours worked). This means that it is possible to construct four different measures of labour productivity: gross output per worker; gross output per hour; value-added per worker; and value-added per hour. The analysis in this paper concentrates on our measure of firm-level real value-added per hour worked and the components that make up that variable, predominantly sales per hour worked and purchases per hour worked. The use of real value-added is consistent with Statistics New Zealand's (2006) estimates of aggregate labour productivity for New Zealand. Their measure of labour productivity is however based on estimates of hours paid whereas we use an estimate of hours worked. Industry coverage also differs somewhat.[7]

## 2.2 Dealing with missing data

Merging demographic and employment data with GST sales and purchases data highlighted several issues. First, there were enterprises that had GST sales information but no employment data for the entire period they existed, or conversely had employment data but no GST sales information. Because it was not possible to form a measure of firm labour productivity when either employment or GST sales data were missing for the entire period the firm existed, these firms were dropped.[8] Second, some enterprises had partial information on employment or GST sales for part of the period the firm was recorded as existing. When this occurred during the middle of a firm's existence the missing data observations were filled if at all possible using historical imputation. For example, a firm in existence between 1994 and 2003 with GST sales for the corresponding period but missing employment data in 1996 and 1997 would have the missing 1996 and 1997 employment data filled using employment in 1995. If historical imputation was not possible we then impute using data from subsequent years.

---

6 This approach was also applied to continuing firms that operated for only part of any given year.

7 Statistics New Zealand's aggregate New Zealand labour productivity estimates are for the "measured sector" which covers about 65 percent of total New Zealand industry GDP and about 69 percent of total paid hours. As Table 2 shows, the database used for this study covers close to 90% of total New Zealand GDP and of total hours worked.

8 Approximately 344,000 firms were first removed from the dataset due to missing employment information in all years they were observed. A further 77,000 firms were then removed due to missing GST sales in all years they were observed.

A partial explanation for situations similar to the example above is that some firms fail to respond to the Annual Business Frame Update (ABFU) questionnaire, despite the firm still operating (as indicated by the firm filing GST sales of $30,000 or greater). The non-response rate for the ABFU is estimated to be between 10% and 15%. The approach to imputing missing values is one that has been adopted by Statistics New Zealand in other contexts.

Cases occurred where a firm was in existence in the BD but there was no recorded information on total employment and GST sales or purchases at the beginning or end of the firm's life. In these situations the firm was deemed to be an entrant in the first period that either employment or GST was available and was 'ceased' in the period following the last observation for either employment or GST sales.[9]

Table 1 shows the numbers of firms deemed to be in operation for each year between 1994 and 2003 and the percent of missing observations on sales, purchases and total hours in each year that are subsequently filled as described above. Missing observations on total hours are more prevalent than those on sales or purchases. Across all years 22% of observations on total hours are missing[10] as compared to 11% for each of sales and purchases. The portion of firm-year observations with missing information on any one of sales, purchases or total hours is 29%. Fewer observations are missing in 1994 and 2003 than in other years, particularly on sales and purchases. This is likely to be due to the fact that potential missing observations for firms existing in previous (in the case of 1994) and later years (in the case of 2003) cannot be observed and therefore cannot be taken into account in the process used for imputing missing data.

***Table 1 Missing observations***

| Year | Firm count | Sales % | Purchases % | Hours % | Any data missing % |
|---|---|---|---|---|---|
| 1994 | 183,769 | 3 | 4 | 19 | 22 |
| 1995 | 213,846 | 9 | 9 | 20 | 27 |
| 1996 | 230,213 | 11 | 11 | 21 | 28 |
| 1997 | 239,453 | 12 | 12 | 25 | 31 |
| 1998 | 249,246 | 12 | 12 | 23 | 29 |
| 1999 | 257,103 | 12 | 12 | 25 | 31 |
| 2000 | 265,058 | 15 | 15 | 23 | 31 |
| 2001 | 265,867 | 15 | 15 | 23 | 33 |
| 2002 | 268,558 | 10 | 11 | 20 | 28 |
| 2003 | 256,430 | 4 | 5 | 19 | 23 |
| All years | 2,429,543 | 11 | 11 | 22 | 29 |

Table 2 shows the totals of all firms' sales, purchases, value-added and hours worked after filling in every year. Total value-added ranges from between 80.9 and 96.2 percent of GDP over the period. Total sales and total purchases in each year are on average approximately 3.5 and 2.5 times total value-added respectively. Total hours worked per year averages 2,730 million for the entire period (or around 1.4 million full time equivalent employees[11]) which represents on average around 90 percent of economy-wide total hours worked (ranging from about 83.3 percent in 1994 to 94.5 percent in 1999).

9 This situation may occur because i) SNZ are unable to determine whether non-response to the ABFU is genuine non-response or because the enterprise has ceased operating; or ii) the enterprise continues to file GST returns as it sells off assets even through it has ceased trading.

10 This is in addition to any historical imputation that Statistics New Zealand may have undertaken.

11 One full time equivalent employee represents 1920 hours of labour (i.e. 40 hours per week for 48 weeks per year).

***Table 2  Firm variable totals: 1994 to 2003***

| Year | Total sales ($million) | Total purchases ($million) | Total value-added ($million) | Percentage of economy-wide GDP | Total hours worked (millions) | Percentage of economy-wide hours worked |
|---|---|---|---|---|---|---|
| 1994 | 248960 | 177793 | 71038 | 81.3 | 2271 | 83.3 |
| 1995 | 269904 | 195003 | 74778 | 81.6 | 2492 | 87.2 |
| 1996 | 284381 | 207051 | 76915 | 80.9 | 2627 | 89.0 |
| 1997 | 296996 | 213584 | 82596 | 84.3 | 2702 | 91.4 |
| 1998 | 301134 | 215433 | 84798 | 86.7 | 2747 | 93.3 |
| 1999 | 310654 | 222401 | 86964 | 86.4 | 2808 | 94.5 |
| 2000 | 331368 | 235233 | 94621 | 89.8 | 2855 | 93.9 |
| 2001 | 347086 | 241402 | 103592 | 96.2 | 2870 | 92.7 |
| 2002 | 340712 | 232790 | 105771 | 93.9 | 2957 | 93.1 |
| 2003 | 344502 | 232513 | 108684 | 92.9 | 2971 | 91.9 |

## 2.3 Weighting

The means, medians and standard deviations of key variables used in our analysis are presented in Table 3. The table compares estimates that are unweighted (i.e. each firm-year observation contributing equally), weighted by the firm's annual total hours, and weighted by the firm's mean annual total hours over its lifetime. Weighting better reflects the productive capacity of different firms, by giving greater importance to large versus small firms. Weighting generally increases the means and medians for employment hours, sales per hour, purchases per hour, and value-added per hour. The choice between a time varying and a time invariant weight also makes a difference to the summary statistics. All our subsequent analysis is weighted, using the mean annual total hours worked for each firm over its lifetime within the sample period.

***Table 3  Comparison of weighted and unweighted labour productivity and components***

| | Unweighted | | | Weighted (hours) | | | Weighted (mean hours) | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Mean | Median | SD | Mean | Median | SD | Mean | Median | SD |
| Hours[a] | 11.28 | 3.08 | 124.63 | 1389.13 | 54.82 | 3483.80 | 1317.47 | 47.73 | 3392.06 |
| Sales / hour[b] | 100.82 | 38.76 | 1875.80 | 112.45 | 55.71 | 720.57 | 151.26 | 57.98 | 1539.06 |
| Purchases / hour[b] | 68.39 | 20.33 | 1685.36 | 80.11 | 30.75 | 656.63 | 112.21 | 32.40 | 1449.26 |
| Labour productivity[b] | 32.47 | 14.31 | 492.49 | 32.40 | 19.91 | 201.94 | 39.13 | 20.74 | 284.74 |

Notes to Table 3: a = thousands, b = single units.

## 2.4 Other Issues

There are a number of potential problems with the various data sources used and the data set constructed from them that may, among other things, reduce the reliability of the estimated proxy for labour productivity. One is the fact that sales and purchases of capital goods are included in the variables that measure firms' sales and purchases. The typically infrequent and lumpy nature of capital transactions means that measures of value-added for individual firms could be quite volatile from period to period and could even be large negative values, with corresponding negative measures of labour productivity. In fact about 13 percent of firm-year observations on labour productivity are negative.

Another potential problem could arise with the assumed timing of transactions used to calculate value-added. For all firms in every year, value-added is calculated as sales in that year less purchases in the same year. For many firms, however, sales may in fact lag purchases. Further, the timing of sales versus purchases may differ across industries or even across firms within industries.

There are other data limitations. First, small enterprises that have GST sales below $30,000 are excluded from the BD. Second, company restructures and changes of ownership that are accompanied by new GST registrations will result in enterprise births and deaths even though these pertain to existing enterprises. Therefore, enterprise births and deaths may reflect administration changes in addition to genuine business start ups and closures.[12] Third, employment data are for a point in time, while sales and purchases data are on an annual basis. Fourth, for entering and exiting firms in particular, sales and purchases data were not always available for the entire year of entry or exit and had to be annualised.

These limitations mean that any results using this data should be interpreted with a degree of caution. They also highlight the need for careful examination of the properties of these data.

---

12 Although the BD does not control for 'false' births and deaths owing to enterprise administrative changes, the development of the Linked Employer Employee Database (LEED) is attempting to do this.

# 3 Summary statistics

In this section summary measures are presented for key variables to give a profile of the productivity and size of firms. Section 3.1 discusses how the mean values and standard deviations of labour productivity and its components vary across industries. The distribution of labour productivity within industries is also discussed. Cohorts of firms are examined in Section 3.2 for any systematic differences in labour productivity. In Section 3.3 the distribution of labour productivity is examined in more detail. All sales per hour, purchases per hour and labour productivity per hour are weighted by mean annual total hours worked.

## 3.1 Aggregate and industries

The means and standard deviations of sales per hour, purchases per hour and labour productivity for all one digit industries and the aggregate are given in Table 4. Average aggregate labour productivity for the years 1994 to 2003 was 39 dollars. That is, an average of 39 dollars of value-added was generated for every hour worked. This was made up of an average of 151 dollars of sales less 112 dollars of purchases for every hour worked.

***Table 4 Aggregate and industry average means and standard deviations of labour productivity and components***

| | Sales / hour | | Purchases / hour | | Labour productivity | |
|---|---|---|---|---|---|---|
| Industry | Mean | SD | Mean | SD | Mean | SD |
| Agriculture, forestry & fishing | 88 | 282 | 66 | 244 | 22 | 162 |
| Mining | 201 | 1184 | 136 | 397 | 65 | 888 |
| Manufacturing | 137 | 484 | 109 | 463 | 28 | 120 |
| Electricity, gas and water supply | 724 | 3905 | 559 | 3749 | 165 | 456 |
| Construction | 102 | 340 | 71 | 266 | 31 | 117 |
| Wholesale trade | 393 | 1208 | 332 | 1086 | 61 | 449 |
| Retail trade | 127 | 295 | 111 | 297 | 16 | 68 |
| Accommodation, cafes and restaurants | 51 | 94 | 35 | 68 | 16 | 60 |
| Transport and storage | 170 | 683 | 94 | 483 | 76 | 296 |
| Communication services | 167 | 324 | 94 | 243 | 73 | 196 |
| Finance and insurance | 142 | 957 | 106 | 870 | 36 | 520 |
| Property and business services | 160 | 1979 | 103 | 1773 | 57 | 478 |
| Government administration & defence | 224 | 710 | 125 | 605 | 99 | 276 |
| Education | 170 | 2378 | 147 | 2289 | 23 | 100 |
| Health and community services | 51 | 83 | 24 | 54 | 27 | 46 |
| Cultural and recreational services | 116 | 743 | 74 | 564 | 43 | 276 |
| Personal and other services | 49 | 71 | 25 | 55 | 24 | 32 |
| Aggregate | 151 | 1539 | 112 | 1449 | 39 | 285 |

There is considerable variation in average sales, purchases and labour productivity across industries. Electricity gas and water supply has the highest mean for all variables: 14.8 times the lowest industry (Personal and other services) for sales per hour, 23.3 times the lowest industry (Health and community services) for purchases per hour, and 10.3 times the lowest industry (Retail trade) for labour productivity. There is also considerable variation within industries, with the standard deviation on each variable being several times the mean for most industries.

Differences in labour productivity across industries are, for the most part, not surprising given that there will be differences in capital intensity. Industries with relatively high levels of labour productivity include for example Electricity, gas and water supply, Transport and storage, and Communication services. Industries with relatively low levels of labour productivity include Accommodation, cafes and restaurants, and Retail trade.

Table 5 shows the distribution of labour productivity within industries. For each one digit industry several percentiles of the distribution are given. It is immediately apparent that there exists a great deal of heterogeneity in firms' labour productivity within all industries. For the aggregate, the median firm is more than twice as productive as the firm at the 25th percentile. The firm at the 75th percentile is almost twice as productive again. Some industries have much greater variation than others. Firms in the Mining, electricity, gas and water supply and in the Wholesale trade industries for example have much higher variability in labour productivity than those in the Retail trade and in the Accommodation, cafes and restaurants industries.

***Table 5 Distribution of labour productivity***

| | Percentile | | | | | | |
|---|---|---|---|---|---|---|---|
| Industry | 1 | 5 | 25 | 50 | 75 | 95 | 99 |
| Agriculture, forestry & fishing | -115 | -17 | 3 | 11 | 24 | 96 | 424 |
| Mining | -528 | -121 | 5 | 20 | 46 | 156 | 416 |
| Manufacturing | -83 | -15 | 12 | 22 | 36 | 84 | 218 |
| Electricity, gas and water supply | -155 | -36 | 37 | 97 | 217 | 662 | 909 |
| Construction | -31 | 0 | 12 | 20 | 30 | 65 | 401 |
| Wholesale trade | -213 | -23 | 12 | 28 | 50 | 220 | 1062 |
| Retail trade | -30 | -3 | 6 | 14 | 22 | 43 | 89 |
| Accommodation, cafes and restaurants | -34 | -3 | 6 | 13 | 21 | 40 | 101 |
| Transport and storage | -51 | -2 | 14 | 27 | 70 | 305 | 964 |
| Communication services | -116 | 0 | 28 | 32 | 146 | 181 | 292 |
| Finance and insurance | -155 | -17 | -1 | 4 | 28 | 130 | 780 |
| Property and business services | -92 | -4 | 11 | 27 | 46 | 137 | 840 |
| Government administration & defence | -7 | 5 | 26 | 35 | 47 | 512 | 1615 |
| Education | -2 | 3 | 7 | 14 | 29 | 39 | 70 |
| Health and community services | -12 | 1 | 14 | 26 | 31 | 55 | 155 |
| Cultural and recreational services | -126 | -9 | 7 | 19 | 39 | 118 | 630 |
| Personal and other services | -19 | -1 | 10 | 23 | 36 | 42 | 86 |
| Aggregate | -69 | -4 | 9 | 21 | 35 | 110 | 481 |

Variation in labour productivity outcomes isn't really any less within one digit industries than it is across the economy. For eleven out of seventeen one digit industries the median firm is more than twice as productive as the firm at the 25th percentile. For eight industries the firm at the 75th percentile is more than twice as productive again. Between the 75th and 25th percentiles the difference in labour productivity can be quite marked. For instance, in Mining the firm at the 75th percentile is more than nine times as productive as the firm at the 25th percentile. There are likely to be several reasons for such variation within industries. Firms may be at different stages of their life cycles,[13] employ very different production technologies, buy and sell large chunks of capital or may simply be better than others at converting labour input into value-added.

Although not shown here, a similar degree of heterogeneity is present in the sales per hour and purchases per hour of firms within the same industry. One difference however between

[13] Young firms for example may have relatively large purchases while they build up inventories, dying firms on the other hand may have relatively high sales.

the distribution of labour productivity and those of sales per hour and purchases per hour is that a significant proportion of observations on labour productivity are negative, in total approximately 13 percent. In almost all industries, at least 5 percent, and in the case of Finance and insurance more than 25 percent, of observations are negative. While negative labour productivity is theoretically quite possible it does cause difficulties when calculating labour productivity growth for individual firms.

There is now a large body of evidence for many countries that productivity dispersion is very large across industries and across firms within the same industry. In their survey of longitudinal studies of firm productivity, Bartelsman and Doms (2000, p. 278) comment that "Of the basic findings related to productivity and productivity growth uncovered by recent research using micro data, perhaps most significant is the degree of heterogeneity across establishments and firms in nearly all industries examined." They point out moreover, that while some of this dispersion may reflect "dirty data", there are several reasons to believe that it reflects real productivity differences. These reasons include the long history of research based on diverse data sets and that this heterogeneity in productivity appears in both developed and developing countries where the extent of statistical error is likely to vary. Moreover, relative productivity across plants has been shown to be correlated with wages, technology use, export success and knowledge differences (see for example Criscuolo, Haskel and Slaughter, 2005). Further, high productivity firms have been found to have higher growth and are less likely to exit. The wide dispersion in sales, purchases and labour productivity that we find amongst New Zealand firms is therefore not surprising and is consistent with international evidence.

## 3.2 Cohorts

Our longitudinal firm database contains firms that were in operation throughout the entire sample period as well as firms that were born and firms that died during the period. It is possible that firms belonging to different cohorts display different levels of labour productivity and that this variation has contributed to the heterogeneity observed in the previous section. Table 6 gives the means and standard deviations of labour productivity for, as well as the number of members in, each possible entry and exit cohort between 1994 and 2003.

Each row of the table provides information on all cohorts of firms that are first observed in a particular year. Columns of the table provide information on all cohorts of firms that are last observed in a particular year. Along diagonals, cohorts are observed for the same length of time. For example, the left most diagonal of the table shows information for all cohorts of firms that were observed for only one year between 1994 and 2003.

The average labour productivity of firms belonging to different cohorts varies considerably, ranging from 15 dollars per hour worked, for the cohort of firms that were first observed in 2001 and last observed in 2002, to 158 for the cohort of firms that were first observed in 1998 and last observed in 2001. Firms that we observe for the entire sample period also have relatively high average labour productivity of 42 dollars per hour worked. Their cohort is ranked 6$^{th}$ out of the 55 cohorts shown in Table 6. The standard deviation of firms' labour productivity within cohorts also varies considerably, ranging from 121 dollars per hour worked for firms entering and exiting in 1998 to 1451 dollars per hour worked for firms entering and exiting in 1997.

It is very hard to identify any pattern in either the means or the standard deviations of labour productivity across cohorts of firms simply by looking at Table 6. Regressions on both cohort means and standard deviations using each cohorts entry year and its age, conditional on entry year, as explanatory variables as well as indicators for whether the cohorts were left and right censored (whether or not the cohorts were first or last observed in the first and last years of our sample period respectively) did not reveal any statistically significant patterns in the data.

***Table 6 Average labour productivity for all cohorts***

| | | Exit-year | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Entry-year | | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 |
| 1994 | Mean | 22 | 56 | 39 | 34 | 28 | 27 | 37 | 27 | 36 | 42 |
| | SD | 122 | 667 | 241 | 277 | 210 | 394 | 342 | 341 | 211 | 233 |
| | N | 6120 | 13178 | 11216 | 11334 | 10049 | 9515 | 9650 | 8563 | 8200 | 95583 |
| 1995 | Mean | | 33 | 28 | 21 | 24 | 19 | 25 | 33 | 21 | 35 |
| | SD | | 407 | 231 | 164 | 235 | 352 | 128 | 230 | 484 | 253 |
| | N | | 1785 | 3663 | 3665 | 3115 | 2699 | 2464 | 2258 | 1988 | 14487 |
| 1996 | Mean | | | 20 | 18 | 21 | 39 | 25 | 28 | 25 | 44 |
| | SD | | | 160 | 150 | 331 | 514 | 135 | 165 | 259 | 417 |
| | N | | | 1541 | 3436 | 3191 | 2715 | 2608 | 2183 | 1851 | 13750 |
| 1997 | Mean | | | | 39 | 20 | 22 | 29 | 26 | 32 | 34 |
| | SD | | | | 1451 | 191 | 211 | 241 | 191 | 602 | 250 |
| | N | | | | 1613 | 3073 | 2712 | 2363 | 2061 | 1718 | 12117 |
| 1998 | Mean | | | | | 21 | 33 | 24 | 158 | 27 | 34 |
| | SD | | | | | 121 | 234 | 164 | 789 | 195 | 208 |
| | N | | | | | 2250 | 3502 | 3283 | 2898 | 2360 | 15490 |
| 1999 | Mean | | | | | | 23 | 29 | 26 | 25 | 40 |
| | SD | | | | | | 225 | 345 | 160 | 227 | 242 |
| | N | | | | | | 2101 | 3606 | 3472 | 2656 | 17620 |
| 2000 | Mean | | | | | | | 25 | 28 | 25 | 35 |
| | SD | | | | | | | 265 | 302 | 269 | 240 |
| | N | | | | | | | 1964 | 3479 | 3491 | 22181 |
| 2001 | Mean | | | | | | | | 47 | 15 | 36 |
| | SD | | | | | | | | 376 | 332 | 343 |
| | N | | | | | | | | 1091 | 2989 | 22683 |
| 2002 | Mean | | | | | | | | | 45 | 31 |
| | SD | | | | | | | | | 476 | 180 |
| | N | | | | | | | | | 1325 | 27261 |
| 2003 | Mean | | | | | | | | | | 27 |
| | SD | | | | | | | | | | 807 |
| | N | | | | | | | | | | 14533 |

## 3.3 Association between labour productivity and its components

The dispersion in labour productivity across firms may be related to the size of the firm, as proxied by their total hours, sales or purchases. To evaluate this possibility, we grouped firms into quartiles according to their level of labour productivity in any given year, where quartile 1 is comprised of the lowest productivity firms and quartile 4 the highest. Given a firm's position in the labour productivity distribution at a particular point in time, its position in the distribution of other variables at the same point in time is then examined. Table 7 provides these relationships.

Table 7A shows how the distribution of labour productivity relates to that of total hours worked or the labour input of firms. The left most column of the table indicates a firm's position in the labour productivity distribution while the uppermost row of the table indicates a firm's position in the distribution of total hours worked. Quartile 1 in both cases refers to the lowest quartile of the distribution and quartile 4 the highest. Each row and column of the table must sum to one as once in a labour productivity or labour input quartile a firm must lie somewhere on the labour input or labour productivity distribution respectively. Figures in Table 7 are averages for the ten years between 1994 and 2003 and indicate the proportion of firms in various labour input quartiles given they belong to a particular labour productivity quartile.

If there is a positive association between a firm's labour productivity and labour input there should be a relatively high proportion of firms in the diagonal elements of the table. Conversely, if there is a negative association between a firm's labour productivity and labour input there should be a relatively high proportion of firms in the off-diagonal elements. If no association exists between labour productivity and labour input, the proportions should be similar across the table.

In general, relatively high proportions of firms in labour productivity quartiles belong to the same quartile of the labour input distribution. The proportion of firms in labour input quartiles tends to decline the more dissimilar that labour input quartile is to the labour productivity quartile. For example, of those firms in the bottom labour productivity quartile around 36 percent are also in the bottom labour input quartile while only 16 percent are in the top quartile. The pattern exhibited in panel A of Table 7 suggests there is a weak positive association between labour productivity and labour input.

***Table 7 Association between labour productivity and measures of firm size (contemporaneous: average of 10 years)***

***A: Labour productivity and total hours***

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.36 | 0.26 | 0.22 | 0.16 |
| 2 | 0.29 | 0.33 | 0.24 | 0.14 |
| 3 | 0.16 | 0.23 | 0.25 | 0.35 |
| 4 | 0.19 | 0.18 | 0.29 | 0.35 |

***B: Labour productivity and sales***

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.50 | 0.23 | 0.17 | 0.11 |
| 2 | 0.30 | 0.35 | 0.23 | 0.12 |
| 3 | 0.12 | 0.24 | 0.29 | 0.34 |
| 4 | 0.08 | 0.18 | 0.31 | 0.43 |

***C: Labour productivity and purchases***

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.43 | 0.26 | 0.17 | 0.14 |
| 2 | 0.30 | 0.34 | 0.23 | 0.13 |
| 3 | 0.14 | 0.23 | 0.30 | 0.33 |
| 4 | 0.12 | 0.17 | 0.31 | 0.41 |

The relationship between the distributions of labour productivity and sales and between labour productivity and purchases are shown in panels B and C of Table 7 respectively. In general, higher proportions of firms belong to the diagonal elements of these tables than is the case for the labour productivity and total hours worked relationship. On average, the proportion of firms who are in the same quartile for labour productivity as they are for sales is around 39 percent. This compares to around 32 percent of firms being in the same quartile for labour productivity as they are for labour input. On average the proportion of firms who are in the same quartile for labour productivity as they are for purchases is around 37 percent.

There appears to be a positive association between labour productivity and each of sales, purchases and labour input, all of which can be thought of as a crude measure of the scale of a firm's production. This is perhaps suggestive of increasing returns to scale. However, to be certain that this reflects scale effects, one would need to control for the effects of other factors likely to impact on firm labour productivity such as capital, inputs.

The relationship between the distributions of labour productivity and its numerator, value-added, is shown in Table 8. On average the proportion of firms who are in the same quartile for labour productivity as they are for value-added is around 47 percent. This relatively high average is mostly due to 70 percent of firms in the bottom quartile and 48 percent of firms in the top quartile of the labour productivity distribution also being in the bottom and top quartiles of the distribution of value-added respectively.

***Table 8 Association between labour productivity and value-added (contemporaneous: average of 10 years)***

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.70 | 0.17 | 0.10 | 0.03 |
| 2 | 0.22 | 0.40 | 0.27 | 0.11 |
| 3 | 0.07 | 0.25 | 0.31 | 0.37 |
| 4 | 0.02 | 0.18 | 0.32 | 0.48 |

Table 9 shows how the distribution of labour productivity relates to that of the sales per hour and purchases per hour of firms respectively. In Table 9A, on average the proportion of firms who are in the same quartile for labour productivity as they are for sales per hour is around 48 percent. Recall from Section 2.1 that there are alternative ways to construct labour productivity from this database, one of which is simply to divide total sales by total hours worked. On average more than half the firms are in different sales per hour quartiles relative to their labour productivity quartile. This suggests that the choice of labour productivity measurement may impact on conclusions.

The relationship between the distributions of labour productivity and purchases per hour is shown in 9B. On average the proportion of firms who are in the same quartile for labour productivity as they are for purchases per hour is around 37 percent.

***Table 9 Association between labour productivity and sales per hour worked and purchases per hour (contemporaneous: average of 10 years)***

***A: Labour productivity and sales per hour***

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.65 | 0.12 | 0.10 | 0.13 |
| 2 | 0.32 | 0.34 | 0.20 | 0.14 |
| 3 | 0.03 | 0.44 | 0.36 | 0.18 |
| 4 | 0.00 | 0.10 | 0.34 | 0.56 |

***B: Labour productivity and purchases per hour***

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.47 | 0.17 | 0.16 | 0.20 |
| 2 | 0.30 | 0.27 | 0.25 | 0.18 |
| 3 | 0.15 | 0.34 | 0.31 | 0.20 |
| 4 | 0.09 | 0.22 | 0.27 | 0.42 |

# 4 Time-series properties and transition rates

This section presents time series analysis of firm labour productivity and its components. The aim is to gain a better understanding of the variability and noise present in the data and the persistence of firms' labour productivity in particular. We describe the patterns using two metrics. First, Section 4.1 presents the autocorrelations of, and cross-autocorrelations between, sales per hour, purchases per hour and labour productivity. Second, Section 4.2 describes the 1-year, 4-year and 9-year transition probabilities between quartiles (and also entry and exit status) for each of sales per hour, purchases per hour and labour productivity. The group of firms we analyse is all firms in our sample, regardless of their industry or length of time in operation.

## 4.1 Autocorrelations

Simple measures of the association between the sales per hour, purchases per hour, total hours and labour productivity are developed by estimating the correlation between these variables and by estimating the association of a variable with its own values in previous and future time periods using autocorrelation analysis.

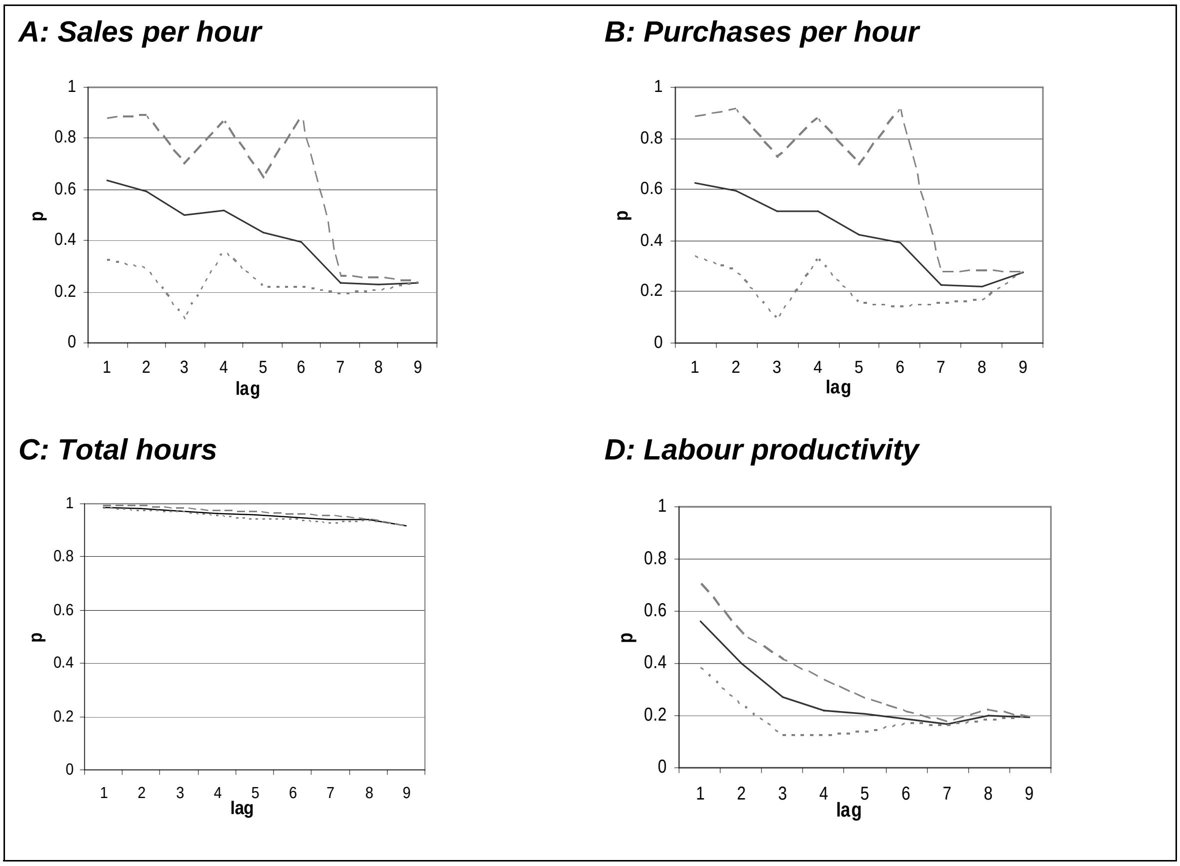
Autocorrelation patterns for sales per hour, purchases per hour, total hours and labour productivity are shown in Figure 1. The centre line shows the average autocorrelation for various lag lengths (where lags are in years). With ten years of data, one for each of 1994 to 2003, there are nine one period autocorrelations from which to derive an average, one for each pair of consecutive years. Similarly there are eight two period autocorrelations and so on. The top line shows the maximum autocorrelation for each lag length. The bottom line shows the minimum autocorrelation for each lag length. Obviously the three lines meet at the same point for the nine period autocorrelation (lag 9 years).

From Panel A in Figure 1 it is apparent that the relationship between a firm's sales per hour in different periods becomes steadily weaker as the length of time between periods increases. For example, the average autocorrelation between sales per hour in consecutive periods is around 0.65 compared to 0.25 for the nine period autocorrelation. Autocorrelations for the same lag length can vary considerably. The highest autocorrelation between consecutive years was 0.87 and occurred between 2002 and 2003 while the lowest was 0.32 and occurred between 2000 and 2001.

The autocorrelation pattern for purchases per hour is shown in panel B of Figure 1. The picture looks very similar to that for sales per hour. The relationship between a firm's purchases per hour in different periods becomes steadily weaker as the length of time between periods increases, tending to about 0.25 by the seven year lag.

Panel C of Figure 1 shows the autocorrelation pattern for firms' total hours. The labour input of firms appears very persistent. The average autocorrelation between firms' total hours in consecutive periods is around 0.98 and is still more than 0.91 for the nine period autocorrelation. This suggests that labour hoarding is quite common and that rather than vary labour inputs in response to variations in demand they tend to compensate in other ways, for example through changes in inventories and variations in the rate of capital utilisation and therefore labour productivity. It may also partly reflect the extent to which observations for this variable are imputed, approximately 22 percent.

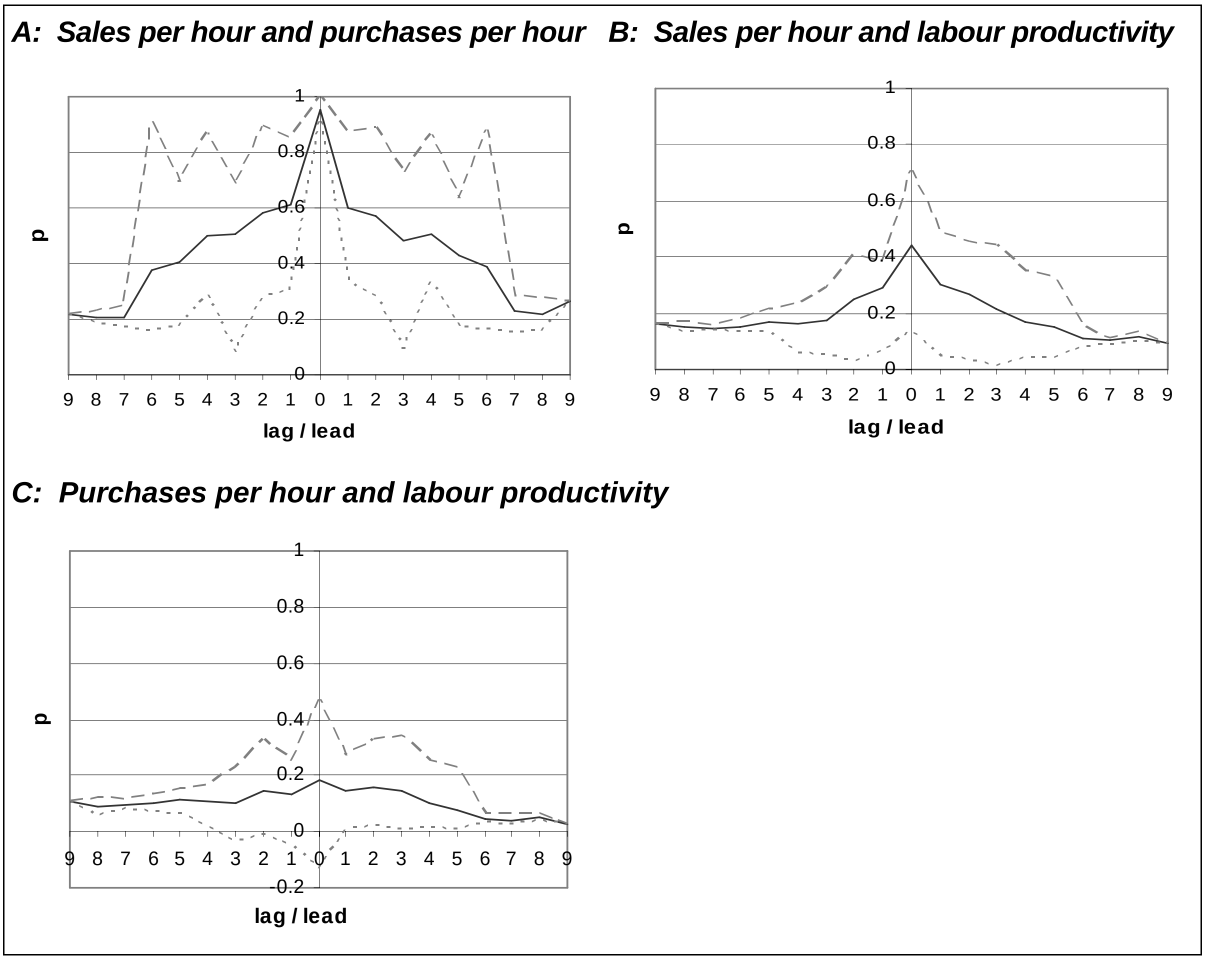
***Figure 1 Autocorrelations for sales per hour, purchases per hour, total hours and labour productivity***




Panel D of Figure 1 shows the autocorrelation pattern for labour productivity. The pattern here differs from those of sales per hour and purchases per hour. The relationship between a firm's labour productivity in different periods is not as strong and weakens more rapidly as the length of time between period's increases. It is interesting that sales per hour of a firm, an alternative measure of labour productivity, is less volatile than the value-added measure of labour productivity.

Cross-autocorrelation patterns are shown in Figure 2. The cross-autocorrelation pattern for sales per hour and purchases per hour is shown in panel A. The leftward part of the chart shows the cross-autocorrelations between sales per hour in year *t* and purchases per hour in years prior to year *t.* The rightward part of the chart shows the cross-autocorrelations between sales per hour in year *t* and purchases per hour in future years. The two parts are quite symmetrical. Given the potential problem mentioned in Section 2 concerning the appropriate timing of sales and purchases when constructing value-added, it is interesting to note that the highest correlation between sales per hour and purchases per hour is the contemporaneous one.

***Figure 2  Cross-autocorrelations***




The cross-autocorrelation pattern for sales per hour and labour productivity is shown in panel B.  It looks very different to that for sales per hour and purchases per hour.  The relationship is very weak, even contemporaneously.  This further reinforces the point that the choice of measure for labour productivity may materially impact on results.

The cross-autocorrelation pattern for purchases per hour and labour productivity is shown in Panel C of Figure 2.  The relationship looks to be even weaker than that for sales per hour and labour productivity.  In some cases the correlation between purchases per hour and labour productivity is even negative.

Because of firm entry and exit, the sample of firms we have used to calculate the autocorrelations and cross-autocorrelations presented in this section are not constant over the sample period.  This may have had some influence over the patterns we observe. To check for this possibility, we produced the same autocorrelations and cross-autocorrelations for the group of firms that were in operation in every year between 1994 and 2003.  The autocorrelation patterns for this group of firms are similar to those presented in this section for the full sample of firms.  We also compared autocorrelation patterns across different industries and, although the patterns did vary across industries, these differences did not appear to be systematic.  For example, we were unable to discern systematic differences between industries that were likely to be capital intensive and those that were labour intensive, or between industries that exhibited high rates of entry and exit compared to industries that exhibited low rates of entry and exit.

## 4.2 Quartile transitions

Transition matrices can be derived to provide some insight into how a firm's sales per hour, purchases per hour and labour productivity typically evolve over time in relation to other firms. The transition tables that follow are derived as follows. In every year between 1994 and 2003 firms are grouped into quartiles according to their level of sales per hour, purchases per hour or labour productivity. In each case quartile 1 is comprised of the lowest quarter of firms in the distribution and quartile 4 the highest. Given a firm's position in the distribution of a variable at a particular point in time, its position in that same distribution one, four and nine years later are then examined. It is not necessarily the case that firms are alive in both reference periods. To allow for this possibility, firms not in operation in any period are grouped into a category labelled 'Out'. Examination of transitions between this category and the various quartiles of a variable's distribution allow inferences to be drawn about the prevalence of firm entry and exit, and the persistence of relative sales per hour, purchases per hour and labour productivity respectively.

Table 10 describes the transition (relative) frequencies between quartiles of the sales per hour distribution. Panel A describes the average "1-year" transitions (i.e. between consecutive years); similarly, panels B and C describe the average "4-year" and "9-year" (i.e. between 1994 and 2003) transitions. The left most column in each panel indicates a firm's position in the sales per hour distribution in a particular year, or whether it was not yet in operation ('Out'). Similarly, the top row of each panel indicates the firm's position in the distribution 1-, 4- or 9-years later (or if it was no longer in operation). Each row of the table must sum to one because at any particular point in time a firm must either be in a quartile of the sales per hour distribution or not in operation. The table entries are averages across the relevant years between 1994 and 2003 and indicate the proportion of firms in various sales quartiles, or that are no longer in operation, given that they previously belonged to a particular sales quartile or were not yet in operation.

If there is a strong positive association between firms' sales per hour in consecutive periods there should be a relatively high proportion of firms in the diagonal elements of the table. Conversely, if there is a negative association between firms' sales per hour in consecutive periods there should be a relatively high proportion of firms in the off-diagonal elements. If no association exists between sales per hour in consecutive periods, the frequencies should be similar across the cells.

If there is a positive association between the prevalence of firm entry or exit and sales, the proportion of firms in the cells of the bottom row and right most column of the table should increase from left to right and top to bottom respectively, ignoring the cell at the bottom right of the table.[14] Conversely, if there is a negative association between the prevalence of firm entry or exit and sales, the proportion of firms in cells should decrease from left to right and top to bottom respectively. If no association exists, frequencies should be similar across the bottom row and similar across the right most column of the table.

Panel A of table 10 shows that, in general, relatively high proportions of firms in sales per hour quartiles in a particular period belong to the same quartile of the sales per hour distribution a year later. The proportion of firms in the sales per hour quartiles a year later tends to decline the more distant it is from the sales per hour quartile a year earlier. For example, of firms initially in the bottom sales per hour quartile, around 77 percent are also in the bottom sales per hour quartile a year later while only 1 percent are in the top quartile.

---

[14] The figure in this cell shows the proportion of firms that were not in operation in either of the reference periods.

Panel B of Table 10 shows lower proportions of firms on the diagonal for the 4-year transitions compared to the 1-year transitions. On average, the proportion of firms who are in the same quartile for sales per hour as they were four years earlier is around 54 percent. This compares to around 73 percent of firms being in the same quartile for sales per hour as they were one year earlier. The negative association between the prevalence of both entry and exit and sales per hour also seems to be weaker, although not surprisingly far higher proportions of firms enter and exit over four years than over one year.

The 9-year transitions in panel C of Table 10 show that about 40 percent of firms in existence in 1994 are in the same sales per hour quartile in 2003. This is lower again than the equivalent average for the 4-year transitions. It is also now difficult to see any relationship between the prevalence of firm entry and exit, and sales per hour. The patterns in Table 10 suggest there is a positive association between sales per hour in consecutive periods. There also seems to be a negative association between the prevalence of both entry and exit, and sales per hour. In other words firms appear more likely to enter and exit with low levels of sales per hour.

***Table 10 Sales per hour quartile transitions***

| Start-year quartile | End-year quartile 1 | 2 | 3 | 4 | Out |
|---|---|---|---|---|---|
| | | | **A: Average 1-year relative frequencies** | | |
| 1 | 0.77 | 0.11 | 0.02 | 0.01 | 0.09 |
| 2 | 0.12 | 0.66 | 0.16 | 0.02 | 0.04 |
| 3 | 0.03 | 0.16 | 0.67 | 0.12 | 0.03 |
| 4 | 0.02 | 0.02 | 0.11 | 0.82 | 0.03 |
| Out | 0.05 | 0.04 | 0.03 | 0.03 | 0.86 |
| | | | **B: Average 4-year relative frequencies** | | |
| 1 | 0.53 | 0.15 | 0.04 | 0.02 | 0.26 |
| 2 | 0.13 | 0.48 | 0.20 | 0.03 | 0.16 |
| 3 | 0.07 | 0.15 | 0.48 | 0.17 | 0.13 |
| 4 | 0.05 | 0.03 | 0.14 | 0.65 | 0.14 |
| Out | 0.14 | 0.12 | 0.10 | 0.09 | 0.54 |
| | | | **C: Average 9-year Relative Frequencies** | | |
| 1 | 0.41 | 0.12 | 0.06 | 0.03 | 0.38 |
| 2 | 0.09 | 0.32 | 0.24 | 0.05 | 0.30 |
| 3 | 0.07 | 0.13 | 0.35 | 0.18 | 0.27 |
| 4 | 0.04 | 0.03 | 0.12 | 0.50 | 0.31 |
| Out | 0.21 | 0.21 | 0.15 | 0.15 | 0.28 |

Table 11 similarly shows the 1-, 4- and 9-year transition probabilities for purchases per hour. Apart from a weaker relationship between firm entry and purchases per hour than that for sales per hour, the patterns here look very similar. On average the proportions of firms that are in the same quartile for purchases per hour as they were one, four and nine years earlier are around 73, 53 and 38 percent respectively. As for sales per hour, the amount of movement by firms across the distribution of purchases rises over time.

*Table 11 Purchases per hour quartile transitions*

| Start-year quartile | End-year quartile 1 | 2 | 3 | 4 | Out |
|---|---|---|---|---|---|
| | **A: Average 1-year relative frequencies** | | | | |
| 1 | 0.76 | 0.13 | 0.02 | 0.01 | 0.09 |
| 2 | 0.14 | 0.66 | 0.16 | 0.01 | 0.04 |
| 3 | 0.03 | 0.15 | 0.67 | 0.11 | 0.03 |
| 4 | 0.01 | 0.02 | 0.11 | 0.83 | 0.03 |
| Out | 0.04 | 0.03 | 0.03 | 0.03 | 0.86 |
| | **B: Average 4-year relative frequencies** | | | | |
| 1 | 0.53 | 0.18 | 0.04 | 0.02 | 0.24 |
| 2 | 0.14 | 0.47 | 0.21 | 0.03 | 0.16 |
| 3 | 0.06 | 0.14 | 0.48 | 0.17 | 0.15 |
| 4 | 0.04 | 0.04 | 0.12 | 0.65 | 0.15 |
| Out | 0.14 | 0.12 | 0.10 | 0.10 | 0.54 |
| | **C: Average 9-year relative frequencies** | | | | |
| 1 | 0.38 | 0.17 | 0.09 | 0.02 | 0.33 |
| 2 | 0.11 | 0.32 | 0.24 | 0.03 | 0.29 |
| 3 | 0.06 | 0.11 | 0.32 | 0.19 | 0.31 |
| 4 | 0.04 | 0.03 | 0.10 | 0.51 | 0.32 |
| Out | 0.21 | 0.20 | 0.15 | 0.15 | 0.28 |

Table 12 shows 1-, 4- and 9-year transitions for labour productivity. The patterns here look somewhat different to those for sales per hour and purchases per hour. In general, lower proportions of firms belong to the diagonal elements of these tables than those for sales or purchases per hour. On average the proportions of firms that are in the same quartile for labour productivity as they were one, four and nine years earlier are around 62, 44 and 33 percent respectively. The negative association between the prevalence of firm entry and exit, and labour productivity looks to be slightly stronger however. In other words, firms look more likely to enter or exit with low levels of labour productivity, a characteristic also noted by Law and McLellan (2005).

*Table 12 Labour productivity quartile transitions*

| Start-year quartile | End-year quartile 1 | 2 | 3 | 4 | Out |
|---|---|---|---|---|---|
| | **A: Average 1-year relative frequencies** | | | | |
| 1 | 0.67 | 0.15 | 0.04 | 0.05 | 0.09 |
| 2 | 0.17 | 0.54 | 0.19 | 0.06 | 0.04 |
| 3 | 0.05 | 0.19 | 0.57 | 0.16 | 0.03 |
| 4 | 0.04 | 0.06 | 0.17 | 0.70 | 0.03 |
| Out | 0.05 | 0.04 | 0.02 | 0.03 | 0.86 |
| | **B: Average 4-year relative frequencies** | | | | |
| 1 | 0.43 | 0.17 | 0.06 | 0.08 | 0.26 |
| 2 | 0.18 | 0.35 | 0.19 | 0.10 | 0.18 |
| 3 | 0.09 | 0.18 | 0.44 | 0.17 | 0.12 |
| 4 | 0.07 | 0.08 | 0.20 | 0.52 | 0.12 |
| Out | 0.14 | 0.14 | 0.09 | 0.09 | 0.54 |
| | **C: Average 9-year relative frequencies** | | | | |
| 1 | 0.33 | 0.14 | 0.07 | 0.08 | 0.39 |
| 2 | 0.18 | 0.24 | 0.15 | 0.07 | 0.37 |
| 3 | 0.08 | 0.15 | 0.34 | 0.17 | 0.26 |
| 4 | 0.07 | 0.06 | 0.21 | 0.42 | 0.24 |
| Out | 0.19 | 0.22 | 0.16 | 0.15 | 0.28 |

In summary, the relative rates of firms' sales per hour and purchases per hour display quite high year-to-year persistence. Around 73 percent of firms remain in the same quartile of sales per hour and purchases per hour from one year to the next. Even after nine years around 40 percent of firms are in the same quartile. Relative rates of labour productivity are somewhat less persistent for one year transitions, although around 62 percent of firms remain in the same quartile after one year. After four years this percentage drops to around 43 and after 9 years it drops to around 33.

Although these transitions rates suggest a high degree of persistence, the limited amount of international evidence that does exist suggests that persistence is to be expected. For instance, Bartelsman and Dhrymes (1998) find that for US manufacturing plants, more than one-third of plants remain in the same quintile after five years. Baily, Hulten and Campbell (1992) used a sample of 23 four-digit SIC industries and, when weighted by employment in each plant, found that about 20 percent of employment remains in its original position after ten years. Compared to rates of transition for US manufacturing, New Zealand firm transition rates may be a little lower, but what is clear is that a high degree of persistence is not unusual.

# 5 Modelling

In this section we develop and describe a simple stylised statistical error components model that is (broadly) consistent with some of the stylised facts pertaining to the autocorrelation properties of the data described in Section 4, particularly Section 4.1. There is a large literature on modelling firm productivity dynamics (e.g. see Sutton, 1997, for an overview). However, as with the previous results, we do not explicitly consider firm entry and exit decisions and effects. Rather, our modest objective here is simply to calibrate this model to these stylised facts, in order to provide a *possible* interpretation of the sources of variation observed in the GST-based measures of labour productivity. We also remain largely agnostic on the extent to which the observed variations reflect (true) productivity variation or data quality issues.

The 'stylised facts' of the data that the model will attempt to replicate can be summarised as follows:

1. Autocorrelations in sales per hour ($S$), purchases per hour ($P$), and labour productivity per hour ($lp$) each range from (trivially) 1 contemporaneously, then fall sharply to around 0.6 on average at 1-year lags, and then gradually to around 0.2 at 9-year lags.
2. Cross-autocorrelations between $S$ and $P$ are also high (about 0.95 on average) contemporaneously, fall sharply to about 0.6 at lag-1, and then gradually to around 0.2 at lag-9.
3. Cross-autocorrelations between $S$ and $lp$ are about 0.4 contemporaneously, and then fall gradually to about 0.1 at lag-9.
4. Cross-autocorrelations between $P$ and $lp$ are about 0.2 contemporaneously and fall to near-0 by lag-9.

## 5.1 The model

The model we develop here focuses on firm sales per hour ($S$) and purchases per hour ($P$), and treats labour productivity ($lp$) as simply the difference between $S$ and $P$. There are two basic characteristics of the model. First, in a univariate context, our statistical modelling of $S$ and $P$ allows each to have three components of firm-level variation: a time-invariant permanent component, a persistent (but declining) component, and a purely transitory component. Loosely speaking, the first two components may be thought of as being associated, respectively, with permanent differences in business activity across firms, perhaps due to firm-specific technologies, patents, and location and/or human capital (dis)advantages; persistent, though non-permanent business activity shocks, perhaps due to cyclical or other temporary shocks (e.g. droughts and/or exchange rate movements) that differentially affect firms and industries. The third component may, similarly, reflect purely transitory shocks to firm activity; alternatively, this component may be due to purely random measurement error (or 'noise') in the data.

In particular, to capture these univariate features, we assume that $S$ and $P$ satisfy the following statistical processes:

$$P_{it} = \lambda_i^P + \theta_{it}^P + \eta_{it}^P, \tag{1a}$$

and

$$S_{it} = \lambda_i^S + \theta_{it}^S + \eta_{it}^S, \tag{1b}$$

for firm-*i* in year-*t*, where (for *j=P,S*) $\lambda_i^j$ represents a time-invariant firm-specific component, $\theta_{it}^j$ is a time-varying persistent component, and $\eta_{it}^j$ is a purely transitory component. We assume throughout this exercise that all shocks and other variables here are normally distributed. Noting that $lp$ is simply the difference between $S$ and $P$ implies

$$lp_{it} = S_{it} - P_{it} = (\lambda_i^S - \lambda_i^P) + (\theta_{it}^S - \theta_{it}^P) + (\eta_{it}^S - \eta_{it}^P), \tag{2}$$

In terms of the autocorrelation patterns described above, the permanent component ($\lambda_i^j$) provides a permanent correlation across different lags, the purely transitory component ($\eta_{it}^j$) contributes only contemporaneously, while the component ($\theta_{it}^j$) facilitates variation across correlation lags. Given the raw autocorrelations for $S$ and $P$ exhibit steady (geometric) decay after lag-1, we suppose that both $S$ and $P$ follow stationary first-order autoregressive, AR(1), processes:

$$\theta_{it}^P = \rho^P \theta_{it-1}^P + \nu_{it}^P, \tag{3a}$$

and

$$\theta_{it}^S = \rho^S \theta_{it-1}^S + \nu_{it}^S, \tag{3b}$$

where $\rho^P$ and $\rho^S$ are coefficients which capture the persistence in the AR(1) processes, and $\nu_{it}^P$ and $\nu_{it}^S$ are innovations. Furthermore, given the similarity of the autocorrelation patterns for $S$, $P$ (and $lp$), we assume that $\rho^P$ = $\rho^S$ = $\rho$ . A convenient consequence of this is that the second component of the $lp$ process also follows the same AR(1) process – i.e. $(\theta_{it}^S - \theta_{it}^P) = \rho(\theta_{it-1}^S - \theta_{it-1}^P) + (\nu_{it}^S - \nu_{it}^P)$ .

Second, and more substantively, to the extent that a firm's GST sales and purchases capture and reflect relevant aspects of the firm's production process, we would expect these measures to be closely related via the firm's business activity. We incorporate this feature by allowing the respective components of $S$ and $P$ to be related through common business activity shocks, and we also assume that such business activity shocks are related via firms' value-added or "mark-up" from purchases to sales.

In particular, first, we assume the permanent components of differences are due entirely to common business activity differences and related by

$$\lambda_i^P = \lambda_i, \text{ and } \lambda_i^S = (1+\mu).\lambda_i \tag{4}$$

where $\mu$ represents the firm's relative value-added between purchases and sales. Second, we assume the sales and purchases' innovations to the AR(1) components include a common business activity innovation, which is similarly affected by value-added between purchases to sales:

$$\nu_{it}^P = \alpha^{1/2}.\nu_{it} + (1-\alpha)^{1/2}.\upsilon_{it}^P \tag{5a}$$

and

$$\nu_{it}^S = (1+\mu).\beta^{1/2}.\nu_{it} + (1-\beta)^{1/2}.\upsilon_{it}^S \tag{5b}$$

where $\alpha$ and $\beta$ represent the relative contributions of the common business activity innovation to the AR(1) shocks of purchases and sales respectively. Finally, we similarly assume that the purely transitory shocks to $S$ and $P$ have common business activity components plus idiosyncratic components:

$$\eta_{it}^P = \gamma^{1/2}.\eta_{it} + (1-\gamma)^{1/2}.\psi_{it}^P \tag{6a}$$

and

$$\eta_{it}^S = (1+\mu)\delta^{1/2}.\eta_{it} + (1-\delta)^{1/2}.\psi_{it}^S \tag{6b}$$

where $\gamma$ and $\delta$ represent the relative contributions of the common business activity innovation to the AR(1) shocks of purchases and sales respectively.

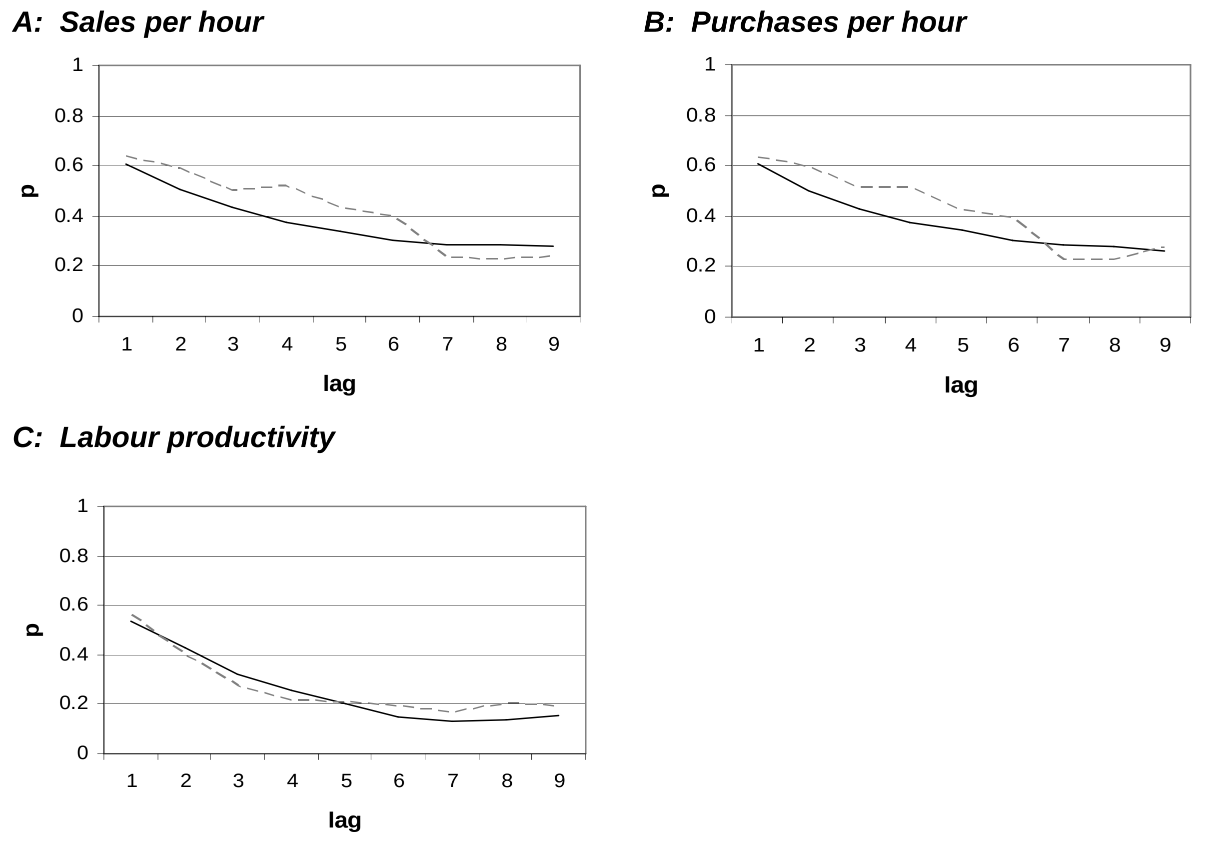
Given this set-up for the model, we choose parameter values as follows. First, based on the result (see Table 4) that sales per hour are on average around 35 percent higher than purchases per hour, we adopt a relative value-added rate of $\mu$=0.35. The AR(1) correlation coefficient $\rho$=0.7. A strong implication of stylised fact 2 is that, contemporaneously, $S$ and $P$ are highly correlated. For this reason, we set $\alpha, \beta, \gamma,$ and $\delta = 0.9$, so that the components common to $S$ and $P$ ($\upsilon_{it}$ and $\eta_{it}$) dominate the idiosyncratic components ($\omega_{it}{}^S$ and $\omega_{it}{}^P$, and $\psi_{it}{}^S$ and $\psi_{it}{}^P$).

All innovations, both common, and specific to, $S$ and $P$ respectively, including the permanent business activity component, $\lambda_i$, are normally distributed with mean zero and variance one. These innovations are randomly generated for one thousand hypothetical firms for twenty periods. They are then used to generate $S$, $P$ and $lp$ according to the model described above. To allow the autoregressive process to stabilise, only the last ten periods of synthetic data are used to produce autocorrelation charts similar to those in section 4.1.

## 5.2 Results

Figure 3 compares the autocorrelation patterns generated by the model (the solid line) and the actual data (the dashed line) for sales per hour (Figure 3A), purchases per hour (Figure 3B) and labour productivity (Figure 3C). The patterns are very similar, especially for labour productivity, where the autocorrelations are relatively lower and decline more rapidly as the lag length increases than do those of sales per hour and purchases per hour.

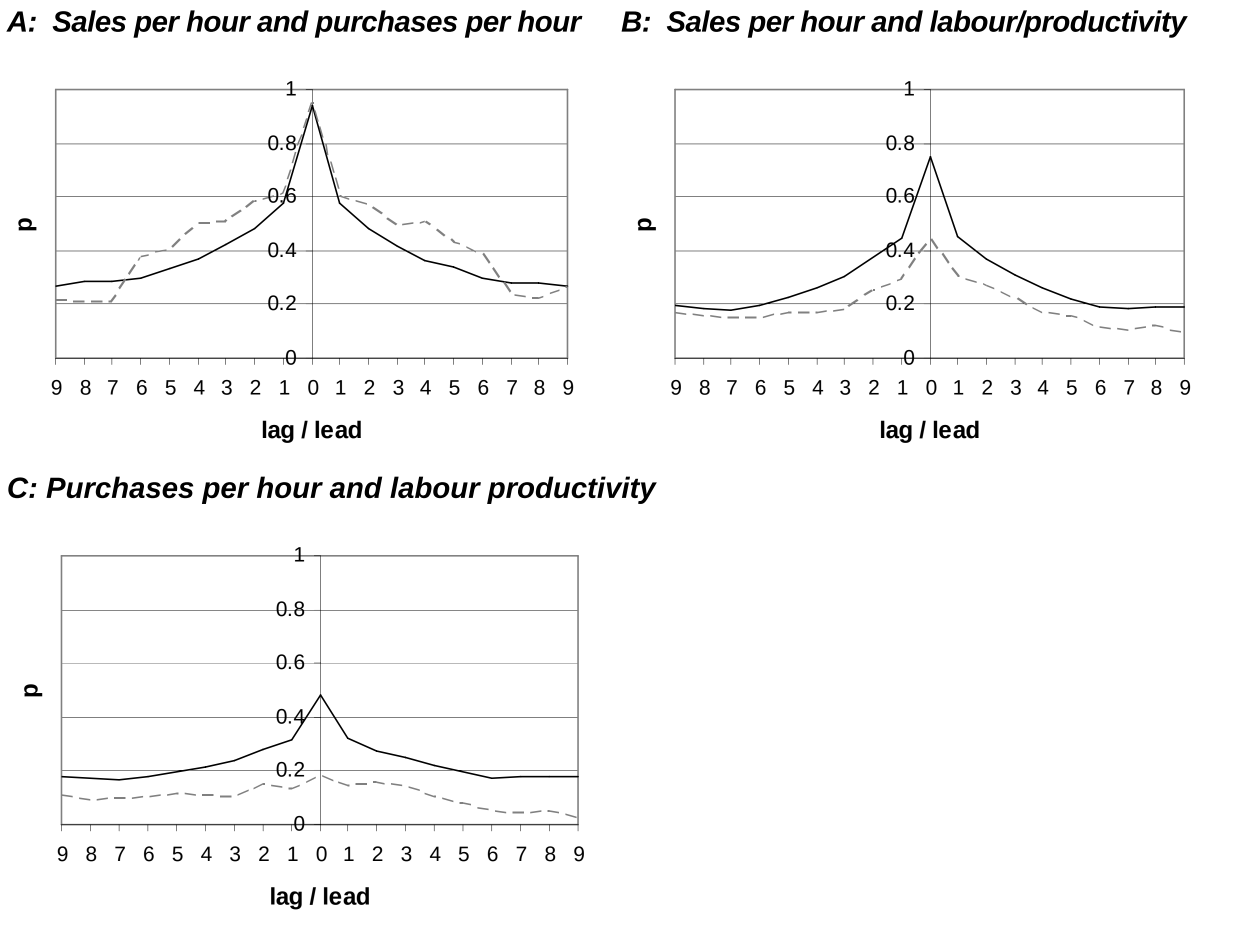
***Figure 3 Calibrated autocorrelations***




The cross-autocorrelations generated by the model (the solid line) and the actual data (the dashed line) are compared in Figure 4. Figure 4A shows the cross-autocorrelations between sales per hour and purchases per hour. Again, the patterns are remarkably similar. The calibrated model captures the high contemporaneous correlation between the two variables, the abrupt drop between lag-0 and lag-1 and the gradual decline thereafter, described in stylised fact 2.

Figures 4B and 4C compare cross-autocorrelations between sales per hour and labour productivity, and purchases per hour and labour productivity generated by the model (the solid line) and the actual data (the dashed line). The calibrated model does not do quite as well at replicating stylised facts 3 and 4 as it did with stylised facts 1 and 2. The cross-autocorrelations generated by the model are somewhat higher than those from the actual data. However, the model does capture the relative difference between cross-autocorrelations, with those between purchases and labour productivity being much lower than those between sales and labour productivity.

*Figure 4 Calibrated cross-autocorrelations*




It is also possible to calculate transition probabilities for the calibrated model, similar to those of Section 4.2. These are presented in the Appendix. While patterns in the various transition probabilities generated by the model and the actual data do share some similarities, it is fair to say that the model does not perform as well in this regard as it did in terms of replicating observed autocorrelation and cross-autocorrelation patterns. This is not particularly surprising however, as the model was calibrated to stylised facts 1 to 4. Even so, this may suggest that the choice of the normal distribution from which to draw each of the models various shocks may be inappropriate. A distribution with more weight in its tails and around its mean may perform better.

# 6 Concluding discussion

Internationally, longitudinal firm unit record databases provide a rich source of information for analysing firm productivity. Statistics New Zealand's Business Demography (BD) and Goods and Services Tax (GST) databases have similar potential yet little is known about their properties. The aim of this study has been to gain a better understanding of the characteristics of the BD and GST data with particular attention being given to measures of labour productivity and its components.

Results reveal a great deal of heterogeneity in each of sales per hour, purchases per hour and labour productivity across industries. Differences across cohorts are less pronounced. Consistent with the findings in the international literature, within each industry or cohort, differences across firms are large. For example, in the case of one industry the firm at the 75th percentile of the labour productivity distribution is more than nine times as productive as the firm at the 25th percentile. There is some evidence to suggest that this difference may in part be due to scale effects as we found positive associations between various measures of firm size and labour productivity. Variation in firm outcomes across the whole economy appears on average to be no less marked. Another constant across industries and cohorts is the presence of negative labour productivity outcomes, constituting nearly 13 percent of all firm-year observations on labour productivity.

On average, measured labour productivity is 35 percent of purchases per hour, although this varies substantially across firms and industries. The univariate time-series properties of sales per hour, purchases per hour and labour productivity are remarkably similar. The autocorrelations in each drop from 1 contemporaneously to about 0.6 at a 1-year lag, and then decay more gradually to around 0.2 after 9-years. Sales per hour and purchases per hour are highly correlated contemporaneously and have similar cross-autocorrelations as either univariate series. The cross-autocorrelations between sales per hour and labour productivity are positive and range from approximately 0.4 to 0.1, while the correlations between purchases per hour and labour productivity are also positive but somewhat lower.

An alternative way to examine the univariate time-series properties of sales per hour, purchases per hour and labour productivity is to calculate transition probabilities for movement of firms between quartiles of each variables respective distribution over varying lengths of time. Results for transition probabilities and autocorrelations are consistent, with the probability of a firm remaining in the same quartile of each variables distribution over time declining steadily as the interval between period's increases. In addition, they reveal that firm entry and exit are both more prevalent for firms at the bottom of the labour productivity distribution.

In order to understand the processes driving the firm data, a simple statistical model for sales and purchases per unit of employment is developed to calibrate to the stylised empirical facts. Reassuringly, both the empirical and model results imply a large fraction of firms' sales and purchases covary contemporaneously. However, the temporal persistence in such differences across firms is much lower. The results are consistent with 10–20 percent permanent variation across firms, and as much as 30 percent being due to purely transitory activity.

With regard to the quality of the BD and GST database, results and the strong performance of our model suggest there is some signal in the data. However, the empirical results also show substantial variation in autocorrelations across pairs of years separated by the same lag-length. Thus, although the results are somewhat encouraging with regards to the information content of the data, the year-to-year variation also demands caution. Hopefully this simple statistical model will help form the basis of economic models of firm productivity to come.

# References


Bartelsman, Eric J and Mark Doms (2000) “Understanding productivity: Lessons from longitudinal microdata”, *Journal of Economic Literature*, 38 (3): 569-594.

Bartelsman, Eric J and Phoebus J Dhrymes (1998) “Productivity dynamics: U.S. manufacturing plants 1972-1986”, *Journal of Productivity Analysis*, 9 (1): 5-34.

Baily, Martin N, Eric J Bartelsman and John C Haltiwanger (1996) “Downsizing and productivity growth: Myth or reality?”, *Journal of Small Business Economics*, 8 (4): 259-278.

Baily, Martin N, Charles Hulten and David Campbell (1992) “The distribution of productivity in manufacturing plants”, *Brookings Papers on Economic Activity: Microeconomics,* Washington DC: 187-249.

Buckle, Robert A and John A Carlson (2000) "Inflation and asymmetric price adjustment" *Review of Economics and Statistics*, 82 (1): 157-160.

Buckle, Robert A and John A Carlson (1998) "Inflation and asymmetric output adjustment by firms" *Economic Inquiry*, 36: 215-228.

Buckle, Robert A and Chris S Meads (1991) "How do firms react to surprising changes to demand? A vector autoregressive analysis using business survey data." *Oxford Bulletin of Economics and Statistics* 53(4): 451-466.

Criscuolo, Chiara, Jonathan E Haskel and Matthew J Slaughter (2005) "Global engagement and the innovation activities of firms." Cambridge MA, National Bureau of Economic Research, Working Paper No 11479.

Dunne, Timothy, Mark J Roberts and Larry Samuelson (1988) "Patterns of firm entry and exit in US manufacturing industries." *RAND Journal of Economics* 19(4): 495-515.

Dixon, Sylvia, David C Maré and Jason Timmins (2005) "Changes in the distribution of firm size in New Zealand" Wellington, Motu Economic & Public Policy Research Report to the Ministry of Economic Development, August.

Fabling, Richard and Arthur Grimes (2003) "Practice makes profit: Business practices and firm success." Wellington, Ministry of Economic Development and *Journal of Small Business Economics* (Forthcoming)

Fabling, Richard and Arthur Grimes (2004) "Three secrets of firm success: Innovation, innovation, innovation." Wellington, Ministry of Economic Development.

Foster, Lucia, John Haltiwanger and Cornell J Krizan (1998) "Aggregate productivity growth: Lessons from microeconomic evidence" Cambridge MA, National Bureau of Economic Research, Working Paper No 4059.

Griliches, Zvi and Haim Regev (1995) "Firm productivity in Israeli industry" *Journal of Econometrics* 65: 175-203.

Law, David and Nathan McLellan (2005) "The contributions from firm entry, exit and continuation to labour productivity growth in New Zealand." Wellington, New Zealand Treasury, Working Paper No 05/01.

Maré, Dave (2005) "Concentration, specialisation and agglomeration of firms in New Zealand." Wellington, Motu Economic & Public Policy Research, Motu Working Paper 05-12.

Maré, Dave and Jason Timmins (2005) "Geographic concentration and firm productivity." Wellington, Motu Economic & Public Policy Research, Forthcoming Motu Working Paper.

Mills, Duncan and Jason Timmins (2004) "Firm dynamics in New Zealand: A comparative analysis with OECD countries." Wellington, New Zealand Treasury, Working Paper No 04/11.

OECD (2004) Understanding economic growth: A macro-level, industry level, and firm-level perspective, (Paris: OECD and Palgrave MacMillan*).* pp165.

Statistics New Zealand (2006) "Productivity statistics: 1988-2005", Productivity Information paper, Wellington, April. http://www2.stats.govt.nz/domino/external/pasfull/pasfull.nsf/web/Paper+Productivity+Information+Paper+1988–2005?open

Sutton, John (1997), "Gibrat's Legacy", *Journal of Economic Literature* 35: 40-59.

# Appendix

***Appendix Table 1 Calibrated quartile transitions***

*Sales per hour: average 1-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.53 | 0.28 | 0.12 | 0.08 |
| 2 | 0.28 | 0.31 | 0.27 | 0.14 |
| 3 | 0.12 | 0.27 | 0.33 | 0.28 |
| 4 | 0.07 | 0.14 | 0.28 | 0.51 |

*Sales per hour: average 4-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.42 | 0.26 | 0.19 | 0.13 |
| 2 | 0.30 | 0.30 | 0.22 | 0.17 |
| 3 | 0.17 | 0.26 | 0.26 | 0.32 |
| 4 | 0.11 | 0.18 | 0.33 | 0.38 |

*Sales per hour: average 9-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.34 | 0.27 | 0.24 | 0.15 |
| 2 | 0.28 | 0.30 | 0.20 | 0.22 |
| 3 | 0.23 | 0.20 | 0.28 | 0.28 |
| 4 | 0.15 | 0.23 | 0.28 | 0.34 |

*Purchases per hour: average 1-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.57 | 0.24 | 0.12 | 0.06 |
| 2 | 0.24 | 0.30 | 0.33 | 0.14 |
| 3 | 0.14 | 0.30 | 0.27 | 0.30 |
| 4 | 0.05 | 0.16 | 0.28 | 0.50 |

*Purchases per hour: average 4-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.43 | 0.26 | 0.14 | 0.17 |
| 2 | 0.27 | 0.30 | 0.25 | 0.18 |
| 3 | 0.17 | 0.27 | 0.28 | 0.28 |
| 4 | 0.13 | 0.18 | 0.32 | 0.37 |

*Purchases per hour: average 9-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.35 | 0.28 | 0.21 | 0.16 |
| 2 | 0.30 | 0.24 | 0.26 | 0.20 |
| 3 | 0.23 | 0.23 | 0.26 | 0.29 |
| 4 | 0.12 | 0.25 | 0.27 | 0.35 |

*Labour productivity: average 1-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.50 | 0.26 | 0.17 | 0.06 |
| 2 | 0.26 | 0.32 | 0.25 | 0.17 |
| 3 | 0.16 | 0.27 | 0.29 | 0.28 |
| 4 | 0.07 | 0.15 | 0.29 | 0.48 |

*Labour productivity: average 4-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.32 | 0.25 | 0.26 | 0.17 |
| 2 | 0.27 | 0.30 | 0.22 | 0.21 |
| 3 | 0.26 | 0.24 | 0.24 | 0.26 |
| 4 | 0.16 | 0.21 | 0.28 | 0.36 |

*Labour productivity: average 9-year relative frequencies*

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.25 | 0.28 | 0.23 | 0.24 |
| 2 | 0.29 | 0.25 | 0.26 | 0.20 |
| 3 | 0.24 | 0.24 | 0.26 | 0.26 |
| 4 | 0.22 | 0.23 | 0.25 | 0.30 |

Labour productivity and sales per hour (contemporaneous – average of 10 years)

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.63 | 0.25 | 0.10 | 0.02 |
| 2 | 0.27 | 0.39 | 0.25 | 0.09 |
| 3 | 0.07 | 0.26 | 0.38 | 0.29 |
| 4 | 0.03 | 0.09 | 0.28 | 0.60 |

Labour productivity and purchases per hour (contemporaneous – average of 10 years)

| Quartile | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.49 | 0.26 | 0.15 | 0.10 |
| 2 | 0.28 | 0.26 | 0.29 | 0.16 |
| 3 | 0.13 | 0.28 | 0.29 | 0.30 |
| 4 | 0.10 | 0.20 | 0.27 | 0.44 |