A Good Worker is Hard to Find: The determinants of skills shortages in New Zealand Firms

Geoff Mason, Penny Mok, Peter Nunns, Philip Stevens and Jason Timmins

Presented to the New Zealand Association of Economists Annual Conference, Auckland, July 2010

Ministry of Economic Development
Occasional Paper 10/##

June 2010

Ministry of Economic Development
Manatū Ōhanga

Ministry of Economic Development Occasional Paper 10/##

A Good Worker is Hard to Find: The determinants of skills shortages in New Zealand Firms

---

Date: June 2010

Author:

Geoff Mason, National Institute of Economic and Social Research, London

Penny Mok, Ministry of Economic Development

Peter Nunns, Formerly of Ministry of Economic Development

Philip Stevens, Ministry of Economic Development

Jason Timmins, Department of Labour

Acknowledgements

A special thanks to all those involved in the design of the BIS module, including Hamish Hill, Meighan Ragg, Elizabeth Chisholm and Belinda Buxton from Statistics New Zealand, plus Rosie Byford and Richard White from MoRST and Sid Durbin from Treasury.

Contact: Occasionalpapers@med.govt.nz

Disclaimer

Abstract

This paper aims to investigate the determinants of external skill shortages – that is, vacancies that are hard to fill for skill-related reasons within and across industries. This paper utilises a specially-designed survey, the Business Strategy and Skills (BSS) module of the Business Operations Survey 2008 (BOS 2008). We estimate the determinants of firms reporting having, vacancies, hard-to-fill vacancies and skill shortages using probit models and two-stage (Heckman) probit models with selection. We consider a broad suite of variables, including firm's size and industry, their market focus, R&D investment, innovation, previous performance (e.g. productivity), the degree of competition they are subject to.

JEL Classification: J24, J31, L60

Keywords: skill shortages, hard-to-fill vacancies, Business Operations Survey, probit

Table of Contents

List of Figures

List of Tables

A Good Worker is Hard to Find: The determinants of skills shortages in New Zealand Firms

1 Introduction

There is widespread concern that a shortage of workers with particular skills is detrimental to the functioning of the New Zealand economy. Whilst there is a great deal of information at the aggregate level on skills issues, not much is known on how these issues affect individual firms in New Zealand. International evidence suggests that the availability of individuals with the appropriate types and levels of skills have a major impact on the success of firms. Skill shortages directly constrain production and prevent firms from meeting demand and using available input efficiently with consequences for lower productivity (Haskel and Martin, 1993a; Stevens, 2007; Tan, *et al.*, 2007). Indirectly, skill shortages inhibit innovation and use of new technologies which are skill-intensive activities. This may have longer-term impacts on the way firms do business, in terms of their location, size, structure, production methods and product strategy (Durbin, 2004; Mason and Wilson, 2003; Mason, *et al.,* 2005). Thus, analysing how these skill shortages manifest themselves and developing policies to address them is critically important if New Zealand is to raise productivity in industry and improve its international competitiveness.

The interrelationship between the skills of the workforce and the emergence and performance of successful firms is central to many governments' policies, as is creating more high value-added firms. The success of such policies depends upon having a workforce with the appropriate skills. However, it is important to be aware of the crucial interactions between skills and other factors, such as the degree and nature of competition, the business' strategy, and the nature of the products or services themselves – all of which are likely to vary across firms (Mason, 2005).

This paper aims to inform our understanding of the importance of skills to firms. In particular we investigate the determinants of external skill shortages – that is, vacancies that are hard to fill for skill-related reasons. This paper utilises a specially-designed survey, the Business Strategy and Skills (BSS) module of the Business Operations Survey 2008 (BOS 2008). The BOS is an annual omnibus business survey collecting annual financial and employment data, and qualitative information on firm performance, information on innovation and communication technology use. The BSS module was designed to investigate the nature of businesses' current and future strategies, their market focus, skills requirements, internal and external skill gaps and training responses.

In this paper we combine the BSS module with data from other sections of the current and previous years' BOS and the prototype Longitudinal Business Database (LBD) to investigate the determinants vacancies and external skill shortages. We investigate three, increasingly focussed, types of vacancy: We analyse the probability that the firm had any vacancies in the last year. We consider those that proved hard to fill. Finally, we focus our analysis on vacancies that were hard to fill because the applicants lacked the necessary skills, qualification or experience. The LBD includes information from tax and survey-based financial data, merchandise and services trade data, a variety of sample surveys on business practices and outcomes. This allows us to link the responses of the BSS module to a wealth of information on firms.

We use two methods to investigate the determinants of a firm reporting the three types of vacancy. First, we estimate separate probits for the probability of a firm reporting each of the vacancy types. However, the mechanisms causing firms to report each of the various types of vacancies are likely to be interrelated. All skill shortage vacancies are by definition hard-to-fill vacancies. Therefore, we estimate a two-stage (Heckman) probit model with selection.

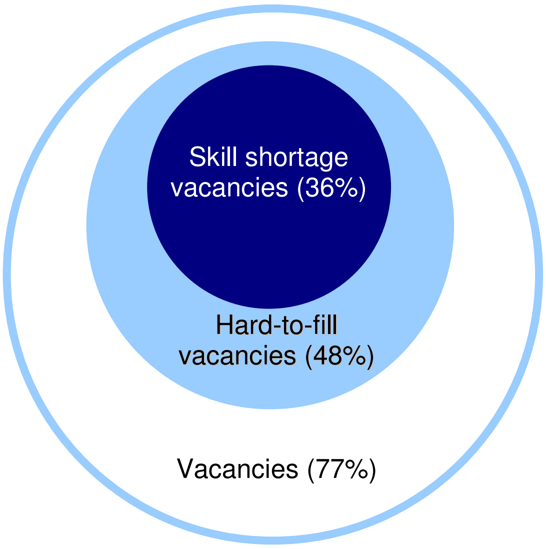
By matching the BSS to the LBD we can consider a broad suite of variables, including firm's size and industry, their market focus, wages, R&D investment, innovation, previous performance (e.g. productivity) and the degree of competition they are subject to.

2 Data and Preliminary Analysis

The data come from Statistics New Zealand's prototype Longitudinal Business Database (LBD). The primary data source is the Business Operations Survey. This is a survey of over five and a half thousand firms employing six or more people[1]. As part of the "Impact of Skills on New Zealand Firms" project, the team designed a module for survey, entitled "Business Strategy and Skills" (BSS). This module collected information about businesses' market focus, current and future business strategy, staff breakdown, vacancies, internal skill gaps and training. Data from the BOS is linked to data from Goods and Services Tax (GST) returns, financial accounts (IR10) and aggregated Pay-As-You-Earn (PAYE) returns.

In this paper we consider all firms that report vacancies of any kind and two subsets – hard-to-fill vacancies and skill shortage vacancies. The overall percentage of firms reporting each type of vacancy is depicted in Figure 1. More detail on the construction and patterns of our measures of vacancies, hard-to-fill vacancies and skill shortage vacancies are set out in the following sections.

Figure 1 Vacancies, hard-to-fill and skill shortage vacancies



- *Figure shows the percentage of firms that report each type of vacancy.*
- *Figures based in sample strata and weights*
- *Note that figures for the percentage of businesses with vacancies and hard-to-fill vacancies will not match the tables in the Statistics New Zealand Hot of the Press release because: (a) we use a slightly different sample and (b) we do not use imputed values*

[1] Note that this is the size of the firm *at the time of sampling*. Because we use actual employment numbers, some of our firms have fewer than six employees.

2.1 Vacancies

Respondents were asked: 'During the last financial year, has this business had any vacancies?' (C14[2]). The first row of Table 1 summarises these data. Overall, 76.6% of firms reported that they had posted a vacancy. As one might expect, given the greater number of employees, the likelihood of posting a vacancy increases with firm size.

Table 1 Vacancies, hard-to-fill and skill shortage vacancies, %

	Business size				Overall
	$E<20$*	$20\leq E<50$	$50\leq E<100$	$E\geq 100$	
Vacancies	71.8	89.4	93.2	94.9	76.6
Hard-to-fill Vacancies	43.2	58.4	65.2	73.4	47.9
Skill Shortage Vacancies	32.1	43.1	48.7	58.6	35.7

- *Table shows percentage of firms reporting each type of vacancy*
- *Figures based in sample strata and weights*
- *Business size (E) is measured by rolling mean employment, or RME.*
- ** Note that the figure for business size being fewer than 20 RME is not all firms in the total business population with fewer than 20 RME, but rather firms in the BOS sample. For more on these see the Data Appendix.*
- *Note that figures for number of businesses with vacancies and hard-to-fill vacancies will not match the tables in the Statistics New Zealand Hot of the Press release because: (a) we use a slightly different sample; (b) we do not use imputed values; and (c) we use rolling mean employment (RME) from the 2008 financial year, rather than 2007.*

We can break the reporting of vacancies down by occupation. Respondents that reported they had posted vacancies in the last year were asked a follow-up question: 'During the last financial year, how many vacancies has this business had for the following roles?' (C15). The responses to this question are set out in Table 2. It is for 'clerical, sales and services workers' that the greatest proportion of firms had vacancies, followed by 'labourers, production, transport or other workers'. This reflects the greater number of staff in these occupations. This is not quite true across all firm sizes. 'Managers' is the second most popular category for firms with more than one hundred employees (and also, marginally, for those with between 50 and 99 employees).

[2] That is, question 14 of Module C

Table 2 Businesses reporting vacancies, %

Occupation	*Business size*				Overall
	E <20*	20≤*E*<50	50≤*E*<100	*E*≥100	
Managers	10.2	23.3	40.0	62.4	15.8
Professionals	11.5	19.1	29.0	42.6	14.7
Technicians and associate professionals	8.7	19.3	25.6	38.5	12.4
Tradespersons and related workers	20.3	25.5	27.5	36.2	22.1
Clerical sales and service workers	28.6	45.2	58.9	72.8	34.5
Labourers, production, transport or other workers	26.2	41.1	47.3	53.3	30.7
All occupations	71.8	89.4	93.2	94.9	76.6

- *Table presents data from questions C14: 'During the last financial year, has this business had any vacancies?' and C15: 'During the last financial year, how many vacancies has this business had for the following roles?'*
- *Table shows percentage of firms reporting each type of vacancy*
- *Figures based in sample strata and weights*
- *Business size (E) is measured by rolling mean employment, or RME.*
- ** Note that the figure for business size being fewer than 20 RME is not all firms in the total business population with fewer than 20 RME, but rather firms in the BOS sample. For more on these see the Data Appendix.*
- *Note that figures for number of businesses with vacancies and hard-to-fill vacancies will not match the tables in the Statistics New Zealand Hot of the Press release because: (a) we use a slightly different sample; (b) we do not use imputed values; and (c) we use rolling mean employment (RME) from the 2008 financial year, rather than 2007.*

2.2 Hard-to-fill vacancies

Respondents were asked: 'During the last financial year, was this business easily able to fill all vacancies with suitable applicants?' (C16). Those whom answered 'no' to this question were classified as having a hard-to-fill vacancy. The second row of Table 1 (repeated at the bottom of Table 3) summarises these data. Well over half of firms that have vacancies find them hard to fill (47.9% compared to 76.6%). Again, the probability of having a hard-to-fill vacancy increases with firm size, with almost three-quarters of firms with rolling mean employment of one hundred or more having hard-to-fill vacancies.

Respondents that reported that they found some vacancies hard to fill were asked: 'For this business, which roles were hard to fill?' (C18). 'Tradespersons and related workers' were the occupations that most businesses had recruitment difficulties with overall (Table 3). However, this once again reflects the greater number of small (6-19 employees) firms.

'Managers' were the role for which most large (100+) firms found difficult to fill vacancies. Given that managers represent a relatively small proportion of total staff, and one that has an important impact on firm performance (Bloom and Van Reenen, 2007, 2010; UTS, 2010), this is a worrying result.

Table 3 Businesses reporting hard-to-fill vacancies, %

	Business size				
Occupation	*E* <20*	20≤*E*<50	50≤*E*<100	*E*≥100	Overall
Managers	6.0	12.7	19.2	31.6	8.7
Professionals	8.5	14.0	18.7	27.0	10.6
Technicians and associate professionals	5.4	11.1	14.5	23.2	7.4
Tradespersons and related workers	15.8	18.5	17.9	21.6	16.6
Clerical sales and service workers	10.8	15.2	19.2	25.4	12.5
Labourers, production, transport or other workers	13.5	19.0	20.9	24.4	15.1
All occupations	43.2	58.4	65.2	73.4	47.9

- *Table presents data from questions C16 'During this last financial year, was this business easily able to fill all vacancies with suitable applicants?' and C18: 'Mark all that apply/ for this business, which roles were hard to fill?'*
- *Table shows percentage of firms reporting each type of vacancy*
- *Figures based in sample strata and weights*
- *Business size (E) is measured by rolling mean employment, or RME.*
- *Note that the figure for business size being fewer than 20 RME is not all firms in the total business population with fewer than 20 RME, but rather firms in the BOS sample. For more on these see the Data Appendix.*
- *Note that figures for number of businesses with vacancies and hard-to-fill vacancies will not match the tables in the Statistics New Zealand Hot of the Press release because: (a) we use a slightly different sample; (b) we do not use imputed values; and (c) we use rolling mean employment (RME) from the 2008 financial year, rather than 2007.*

2.3 External skill shortages – Skill shortage vacancies

Respondents that had hard-to-fill vacancies were asked 'For which of the following reasons did this business find it hard to fill vacancies?' (question C17). They were given twelve categories, from which they could choose as many as they wished. Those that replied 'applicants lack the work experience the business demands' or 'applicants lack the qualifications or skills the business demands' were defined as having skill shortage vacancies (SSVs)[3].

The final row of Table 1 summarises the data on these. Almost three quarters of firms with hard-to-fill vacancies reported external skill shortages (35.7% compared with

[3] For the full list of responses, see SNZ (2008).

47.9%). We break this down into the two constituent parts and present them by firm size and industry in Table 4.

Table 4 Skill-related reasons for hard-to-fill vacancies, by size and industry

	Applicants lack work experience	Applicants lack qualifications or skills
Business size		
6-19 Employees	26	26
20-49 Employees	36	37
50-99 Employees	39	40
100+ Employees	48	48
Industry		
Agriculture, forestry and fishing	20	15
Mining	23	20
Manufacturing	31	30
Electricity, gas, water and waste services	20	17
Construction	41	44
Wholesale trade	31	26
Retail trade	22	23
Accommodation and food services	35	32
Transport, postal and warehousing	33	32
Information media and telecommunications	29	27
Financial and insurance services	32	23
Rental, hiring and real estate services	18	16
Professional, scientific and technical services	34	35
Administrative and support services	23	25
Education and training	21	44
Health care and social assistance	21	28
Arts and recreation services	15	12
Other services	28	31
Overall	29	29

- *Table presents data from questions C16 'During this last financial year, was this business easily able to fill all vacancies with suitable applicants?' and C18: 'Mark all that apply/ for this business, which roles were hard to fill?'*
- *Source: SNZ (2009)*
- *Note that the percentages in this table are taken from 'SNZ (2009) and are not exactly comparable with the Table 1 and Table 2 and Figure 1. For more on this, see the footnotes to Table 1 and Table 2.*

There are almost no differences between the two factors of skill shortage vacancies across business size, but there are some differences across industry. In 'education and training' and 'healthcare and social assistance' it is lack of qualification or skills that is the problem (particularly in the former industry). For financial and insurance services it is lack of experience that is the greater problem. For some industries it is certification, gained in institutions or on-the-job is the most important, for others they do not play this role.

3 Models and Results

In this section we explore the determinants of vacancies, hard-to-fill vacancies and skill shortage vacancies.

Our first empirical model is very simple. We suppose that the propensity of firm i to post a vacancy, to find it hard-to-fill, or to have a skill shortage vacancy can be expressed as

(1) $$y_i^* = X_i\beta - \varepsilon_i$$

where y^* is the propensity to have a vacancy, a hard-to-fill vacancy, or a skill shortage vacancy, X_i is a $(1\times k)$ vector of k explanatory variables, β is an $(k\times 1)$ vector of parameters and ε_i. We do not observe the y^* terms, but the binary realisation of them, therefore we assume:

(2) $$\begin{aligned} y_i &= 0 \quad if \quad y_i^* < 0 \\ y_i &= 1 \quad if \quad y_i^* \geq 0 \end{aligned}$$

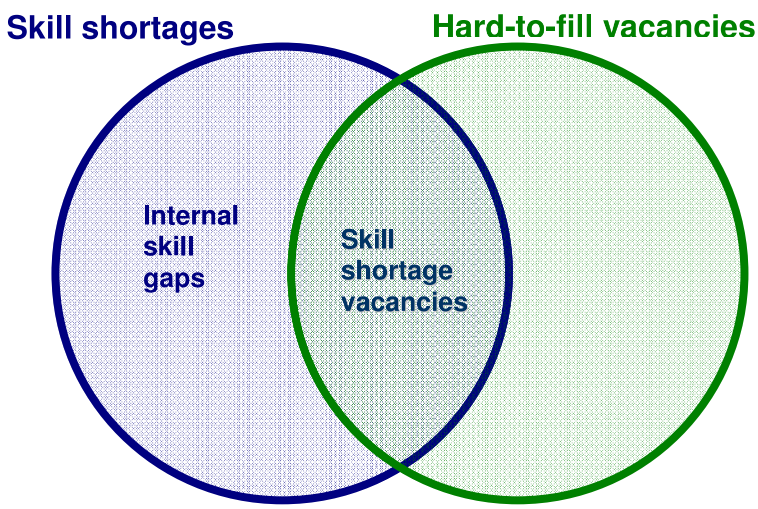
We estimate (2) using a probit model. Because the data was collected using a stratified sample, our models are estimated using sampling weights and correction for stratification[4]. We estimate our models in two forms. First we use contemporaneous variables. This ensures maximum sample size. However, because some of the variables are likely to be endogenous, we also estimate a model using lagged values of the independent variables.

An important issue for modelling is the relationship between the dependant variables. As is clear from Figure 1, all skill shortage vacancies are, by definition, hard-to-fill vacancies, which are of course vacancies. Other analyses have either ignored the relationship between these (e.g. Mason and Stevens, 2003) or relied on the imprecision

[4] using the `svy: probit` command in Stata

in the variables (e.g. Green, Machin and Wilkinson, 1998) or both (e.g. Haskel and Martin, 2001). For example, Green, Machin and Wilkinson (1998) modelled the relationship by estimating bivariate probits (essentially a pair of seemingly unrelated regressions) of 'skill shortages' and hard-to-fill vacancies. However, their measure of 'skill shortages' is rather more general than ours. In the Employers' Manpower and Skills Practices Survey (EMSPS) the respondents were asked 'Would you say that this establishment has experienced a 'skill shortage' in the last 12 months, or not?' contains not just what we would define as skill shortage vacancies, but also internal skills gaps. Indeed, one of the key issues considered by Green *et al.* (1998) is precisely what people mean by 'skill shortages'.

Figure 2 Skill shortages and Hard-to-fill vacancies in Green *et al.* (1998)



We cannot account for the relationship between the terms using a bivariate probit model, because of the fact that the sample for one equation is exactly the same as one of the outcomes of the other. Therefore, in addition to our probit models, we estimate two types of Heckman selection model. The first has the probability of reporting a vacancy as the first stage regression and the probability of reporting a skill shortage vacancy as the second stage. The second specification has as its first stage the probability of having a hard-to-fill vacancy. The first of these two specifications is our preferred one, as the act of posting a vacancy and a firm's ability to fill it appear to be two qualitatively different things. On the other hand, the difference between a hard-to-fill vacancy and a skill shortage vacancy is merely one of classification. Nevertheless, we include both specifications for completeness. We describe this model in more detail in 3.2 below.

The variables we include are the employment, employment turbulence rate (the sum of hires and separations divided by employment), net employment growth rate (hires less separations over employment), wages relative to the 4-digit industry, labour productivity, the growth in sales, whether any employees at the firm are covered by a collective employment agreement, the firms geographic market focus, indicators of ODI and FDI, the nature of competition, whether they have invested in expansion, undertaken or funded R&D, whether they provide training, the occupational make-up of their workforce and industry dummies. The variables themselves are set out in more detail in the Data Appendix to this paper.

3.1 Simple estimates of the probability of reporting vacancies

In this section, we present the results of our probit models of reporting vacancies, hard-to-fill vacancies and skill shortage vacancies. We present two versions of each of these. In section 3.1.1 we use contemporaneous variables from the BOS and other parts of the LBD. This allows us to keep our sample size as large as possible. In section 3.1.2 we used lagged variables to account for endogeneity issues.

3.1.1 Contemporaneous variables

The results for our estimation of the probability of a firm posting a vacancy are set out in Table 5. As one would expect, larger firms (in terms of employment) have more vacancies. This scale impact is stronger for general vacancies (column (1)) than it is for those which are hard to fill (column (2)). Firms with a highly turbulent labour force (lots of hires and fires relative to their employment) are more likely to report vacancies, but are no more likely to report hard-to-fill or skill shortage vacancies. There is no evidence that more productive firms are more likely to have vacancies, although this effect is only statistically significant in the case of skill shortage vacancies.

Fast growing firms, in terms of sales, do not appear to report more vacancies generally, but do report more skill shortage vacancies, suggesting that fast growing firms may well be constrained by shortages of key skills. There is a similar sign to the coefficient on for columns (1) and (2), but these are not statistically significant.

Firms that have innovated (introduced a new product, service, process or marketing method) are more likely to post a vacancy and to find vacancies hard-to-fill for a skill-related reason. This may be because the key skills required by highly innovative firms are hard to come by. There is some weak evidence that businesses with an international focus find it harder to fill there vacancies, but *mark_int* is only significant in model (2).

Table 5 Results – Probits of vacancies using contemporaneous variables

	(1)	(2)	(3)
	All Vacancies	Hard-to-fill Vacancies	Skill shortage vacancies
$\ln(E)$	0.439*** (0.055)	0.245*** (0.036)	0.196*** (0.034)
Turbulence	0.547*** (0.165)	0.280* (0.153)	-0.033 (0.148)
NEG	-0.269 (0.562)	-0.667 (0.532)	-0.420 (0.540)
$\ln(w_i)$- $\ln(w_j)$	0.099 (0.156)	0.155 (0.130)	0.064 (0.133)
$\Delta\ln(w_i)$	0.506 (0.339)	-0.387 (0.289)	-0.373 (0.296)
$\ln(LP_i)$	0.065 (0.051)	0.032 (0.046)	0.056 (0.047)
$\Delta\ln(S)$	0.122 (0.131)	0.065 (0.108)	0.215** (0.106)
R&D	-0.061 (0.150)	-0.131 (0.123)	-0.190 (0.119)
invest	0.183 (0.115)	0.083 (0.088)	0.109 (0.088)
innovate	0.317*** (0.096)	0.087 (0.078)	0.144* (0.078)
union	0.026 (0.105)	0.182** (0.086)	0.177** (0.090)
mark_int	0.111 (0.151)	0.241* (0.136)	0.166 (0.156)
train	0.831*** (0.105)	0.682*** (0.103)	0.657*** (0.114)
odi	0.182 (0.142)	-0.118 (0.119)	-0.219* (0.122)
fdi	0.256 (0.228)	-0.031 (0.174)	-0.002 (0.201)
monopoly	-0.210 (0.201)	-0.047 (0.168)	-0.205 (0.166)
duopoly	0.019 (0.118)	-0.040 (0.099)	-0.046 (0.099)
compet	-0.054 (0.109)	-0.032 (0.093)	-0.087 (0.095)

	(1)	(2)	(3)
	All Vacancies	Hard-to-fill Vacancies	Skill shortage vacancies
prop_man	0.676* (0.397)	-0.488 (0.331)	-0.358 (0.339)
prop_prof	-0.120 (0.326)	0.351 (0.275)	0.625** (0.279)
prop_tech	-0.162 (0.266)	-0.125 (0.237)	0.228 (0.242)
prop_trade	0.165 (0.197)	0.514*** (0.168)	0.785*** (0.163)
prop_cleric	-0.115 (0.206)	-0.020 (0.158)	0.085 (0.155)
Constant	-2.199*** (0.604)	-1.854*** (0.513)	-2.325*** (0.529)
Industry Dummies	Yes	Yes	Yes
Observations	4413	4413	4413
F test	7.240	6.729	6.953
Prob>F	0.000	0.000	0.000

Standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*
stratified and weighted

Firms with a collective agreement are no more likely to have vacancies, but do appear to find those vacancies they have harder to fill. This may be related to the industries in which there is greater union coverage, although note that all regressions include industry dummies.

Firms that train are significantly more likely to report all three types of vacancies. Reasons for this might include the fact that training is potentially an internal response to a skill gap whereas looking for new staff is an external one. In later versions of this paper, we will use a more sophisticated measure of training that accounts for whom it is that is being trained – new staff, existing staff in new roles and existing staff in their existing roles – and what proportion of staff.

It is interesting to note that the proportion of staff in each type only really enter into the two types of hard-to-fill vacancies (columns (2) and (3))[5]. With firms with a higher proportion of staff in professional and trades occupations having vacancies that are hard to fill, with the results in (3) suggesting that this is because they cannot find people with these skills.

[5] Although the proportion of managers does appear to be significant in column (1) at the 10% level.

3.1.2 Lagged variables

For some of the results in Table 5, it is unclear in which direction causality lies. For example, firms may raise wages or provide training in response to difficulties in finding staff. In this section, therefore, we estimate the three probits in columns (1) to (3) using variables from the previous year (2007). All the variables from Module A of the BOS are available in every year of the survey and so can be included in our analysis. The training and market focus variables, however, come from the one-off BIS. Since the BAI, IR10 and LEED are available over the whole period of the LBD (2000-2008), we can include the lagged employment, wage, productivity and sales variables also. The results of explaining the probability of each of the three types of vacancies with lagged variables are presented in Table 6. Many of the results remain, although there are a few differences. Once more, larger firms are more likely to report each of the types of vacancy. When we include lagged employment turbulence, it appears as significant in both the model for all vacancies and hard-to-fill vacancies (although not for skill shortage vacancies). Previous wage growth now strongly predicts the probability of a vacancy. It is insignificant in the specifications for hard-to-fill and skill shortage vacancies, (2a) and (2b).

Table 6 Results – Probits of vacancies using lagged variables

	(1a)	(2a)	(3a)
	All vacancies	Hard-to-fill vacancies	Skill-shortage vacancies
$\ln(E)$	0.537*** (0.063)	0.191*** (0.047)	0.248*** (0.040)
Turbulence	0.611*** (0.208)	0.540*** (0.200)	0.232 (0.176)
NEG	1.478* (0.786)	1.151 (0.775)	1.286** (0.594)
$\ln(w_i)$- $\ln(w_j)$	-0.296 (0.201)	0.063 (0.193)	0.298 (0.204)
$\Delta\ln(w_i)$	0.899** (0.387)	-0.087 (0.368)	-0.143 (0.332)
$\ln(LP_i)$	0.069 (0.057)	0.085 (0.056)	0.084 (0.051)
$\Delta\ln(S)$	0.218 (0.163)	-0.149 (0.152)	-0.195 (0.119)
R&D	0.200 (0.179)	-0.114 (0.147)	0.055 (0.145)
invest	0.066 (0.118)	-0.113 (0.103)	0.030 (0.090)
innovate	0.386*** (0.107)	0.094 (0.095)	0.159* (0.090)
union	-0.055 (0.112)	0.193* (0.105)	0.109 (0.098)
odi	0.116 (0.162)	-0.199 (0.136)	-0.189 (0.127)
fdi	0.191 (0.248)	-0.255 (0.172)	-0.131 (0.154)
monopoly	-0.081 (0.234)	0.142 (0.215)	-0.140 (0.210)
duopoly	-0.152 (0.126)	0.015 (0.123)	-0.076 (0.116)
compet	-0.096 (0.124)	-0.035 (0.123)	-0.101 (0.108)
prop_man	0.339 (0.441)	-0.779* (0.461)	-0.323 (0.397)
prop_prof	-0.016 (0.384)	0.737** (0.359)	0.638** (0.316)
prop_tech	-0.183 (0.316)	0.142 (0.300)	0.289 (0.281)
prop_trade	0.176 (0.213)	0.812*** (0.214)	0.879*** (0.181)
prop_cleric	-0.220 (0.219)	0.143 (0.209)	0.084 (0.178)
Constant	-1.833*** (0.665)	-1.461** (0.634)	-2.305*** (0.567)
Industry dummies F test Prob>F	Yes 4.875 0.000	Yes 3.028 0.000	Yes 4.812 0.000

Standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*
stratified and weighted
includes industry dummies

3.2 Heckman selection models of vacancies

In the previous models we have ignored the relationship between the equations. However, as Figure 1 clearly shows, firms reporting a skill shortage vacancy (i.e. with a dependent variable that has a value of 1 in column (3)) are a subset of those reporting a hard-to-fill vacancy, who in turn are a subset of those that posted a vacancy of any kind. Therefore, in this section we consider a model that attempts to separate out the mechanisms that influence them. We estimate a Heckman selection model with a binomial (probit) model in both 'stages' (Heckman, 1979; Van de Ven and Van Pragg, 1981). We estimate two models the probability of reporting a skill shortage vacancy accounting for sample selection. The second stage model in both is the probability of reporting a skill shortage vacancy. In the first model, we model the first stage (the selection equation) as being the probability of reporting having posted a vacancy. In the second model, we model the selection as being whether a firm reported any hard-to-fill vacancies.

Thus we have a latent equation for skill shortage vacancies (SSVs):

(3) $$SSV_i^* = X_i'\beta' - \varepsilon_{1i}'$$

and a probit equation for the binary outcome:

(4) $$SSV_i = \left(SSV_i^* > 0\right)$$

This variable SSV_i takes one of three values and some of these will vary depending on which model we estimate. If the firm reports a skill shortage, SSV takes the value of 1. In the first model, it takes the value zero if the firm does not have a skill shortage vacancy, but does have a vacancy. It takes a missing value when the firm does not have any vacancies.

In the second model, SSV takes the value zero if the firm does not have a skill shortage vacancy, but does have a hard-to-fill vacancy. It takes a missing value when the firm does not have any hard-to-fill vacancies.

Thus the dependent variable in (4) for observation i is observed if:

(5) $$y_i^{select} = \left(Z_i\gamma + \varepsilon_{2i} > 0\right)$$

where

$$\begin{aligned} &\varepsilon_1 \sim N(0,1) \\ &\varepsilon_2 \sim N(0,1) \\ &corr(\varepsilon_1, \varepsilon_2) = \rho \end{aligned} \tag{6}$$

When $\rho \neq 0$, standard probit techniques applied to (4) yield biased results. The selection model provides consistent, asymptotically efficient estimates for all the parameters in the model. For the model to be identified, we should have at least one variable in the selection equation (5) that is not in the outcome equation.

We estimate two versions of this model, one where y^{select} is whether a firm had any vacancies and the other where y^{select} is whether a firm had any hard-to-fill vacancies. Our preferred model is the model where the sample in the second stage is all firms posting any type of vacancy and thus the selection equation is the probability of reporting a vacancy. This is because the mechanism whereby a firm posts a vacancy will be very different from the influences on whether it cannot find skilled staff to fill vacancies. The first will be driven largely by internal factors that determine the quantity of labour it requires (such as the numbers of quits and growth in demand due to firm growth) and the quality of labour (i.e. what *types* of staff (and skills) it requires because of its business strategy – e.g. whether it is highly innovative). The ability to fill the vacancy will depend much more on the external labour market. Nevertheless, for completeness we include results of considering the selection as operating over having hard-to-fill vacancies and the second stage equation as estimating influences on which firms with hard-to-fill vacancies are experiencing them because of skills. As with the models in section 3.1, we estimate both models using current values of the variables and lagged values.

The results of our Heckman probit estimation are presented in Table 7 and Table 8. Looking first at the selection equations (Table 7), the results are what we would expect from the first two columns of Table 5 and Table 6. Firms with more employees are more likely to have vacancies and hard hard-to-fill vacancies. Also, firms with more turbulent workforces are also more likely to have vacancies (although this is only significant in our preferred models (columns (4) and (4a)). Firms with increasing wages are also more likely to post vacancies (although this effect is not true for hard-to-fill vacancies) and so on.

Turning to the outcome equations in Table 8, we see that once we have accounted for their probability of having vacancies, larger firms are no more likely to have skill shortage vacancies. They do not appear to suffer any particular problems viz. a viz. finding skilled labour than smaller firms over and above that caused by the fact that they do it more frequently. One might say the probability of a skilled vacancy being filled is the same. This is not quite true, but in future work we will investigate this explicitly by weighting the SSV equation by the number of vacancies á la Mason and Stevens (2003). Moreover, large firms appear to be less likely to report a hard-to-fill vacancy more generally (although this result is not significant when we use lagged variables). This may suggest that larger firms have an economy of scale advantage in filling a given vacancy (greater visibility in the labour market, a dedicated human resources department etc.) that balances the fact that they demand greater numbers of skilled labour. This is something we shall investigate further in later work.

Firms that pay higher wages relative to their (4-digit) industry are more likely to have hard-to-fill and skill shortage vacancies. This may be because higher wages are a response to hiring difficulties. However it is still true when we used lagged relative wages. Innovation remains significant and positive in the SSV equation (columns (4) and (4a)) as does training. Firms that have ownership in overseas businesses appear to find it easier to find staff. This may be something to do with the appeal of firms that are expanding overseas.

Table 7 Heckman selection model of skill shortage vacancies, Selection equation

	(4)	(4a)	(5)	(5a)
	Any vacancies		*HTF vacancy*	
	Current	Lagged	Current	Lagged
$\ln(E)$	0.467*** (0.052)	0.482*** (0.062)	0.245*** (0.028)	0.279*** (0.040)
Turbulence	0.541** (0.212)	0.568*** (0.173)	0.106 (0.106)	0.577 (0.614)
NEG	0.292 (0.309)	0.754 (0.649)	-0.126 (0.269)	0.907 (0.653)
$\Delta\ln(w_{it})$	0.798*** (0.214)	0.666* (0.398)	-0.196 (0.231)	0.215 (0.789)
invest	0.096 (0.071)	0.060 (0.080)	0.077 (0.086)	-0.064 (0.204)
innovate	0.362*** (0.071)	0.388*** (0.087)	0.125 (0.079)	0.175 (0.111)
union	-0.070 (0.071)	-0.061 (0.109)	0.186* (0.097)	0.126 (0.198)
mark_int	0.039 (0.133)	0.089 (0.137)	0.196 (0.192)	0.162 (0.154)
train	0.798*** (0.063)	0.750*** (0.087)	0.771*** (0.102)	0.808*** (0.150)
odi	0.133 (0.141)	0.038 (0.160)	-0.152* (0.090)	-0.128 (0.159)
fdi	0.446** (0.193)	0.434* (0.228)	-0.058 (0.112)	-0.269 (0.225)
prop_man	0.659 (0.403)	0.408 (0.449)	-0.421 (0.402)	-0.688 (0.481)
prop_prof	-0.258 (0.318)	-0.232 (0.344)	0.551** (0.244)	0.450 (0.358)
prop_tech	-0.121 (0.279)	-0.306 (0.282)	0.147 (0.109)	0.058 (0.492)
prop_trade	-0.229 (0.149)	-0.333*** (0.125)	0.657*** (0.076)	0.571 (0.781)
prop_cleric	-0.240** (0.121)	-0.383** (0.179)	-0.018 (0.100)	-0.081 (0.248)
invest_nk	0.091 (0.219)	-0.129 (0.310)	-0.039 (0.262)	-0.314 (0.289)
union_nk	-0.310** (0.158)	-0.135 (0.164)	-0.308 (0.217)	-0.310* (0.186)
odi_nk	0.373 (0.444)	0.180 (0.397)	0.534* (0.295)	0.846*** (0.327)
fdi_nk	-0.628 (0.482)	0.096 (0.492)	-0.569** (0.251)	-0.438 (0.305)
Constant	-1.441*** (0.205)	-1.330*** (0.267)	-1.693*** (0.149)	-1.849*** (0.327)
athrho	3.274 (5.324)	1.447*** (0.312)	-5.223 (18.918)	0.055 (3.315)
Observations	4,554	3,393	4,665	3,471

Robust standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*
Weighted and stratified
All specifications include industry dummies

Table 8 Heckman selection model, Determinants of skill shortage vacancies

	(4)	(4a)	(5)	(5a)
	Any vacancies		*HTF vacancy*	
	Current	Lagged	Current	Lagged
Skill shortage vacancies				
$\ln(E)$	0.066 (0.079)	0.018 (0.082)	-0.425*** (0.097)	-0.411 (0.622)
NEG	-0.107 (0.391)	1.279** (0.541)	-0.022 (0.410)	0.942 (1.219)
$\ln(w_i)$-$\ln(w_j)$	0.123** (0.055)	0.161** (0.068)	0.296*** (0.088)	0.354** (0.150)
$\ln(LP_i)$-$\ln(LP_j)$	0.015 (0.031)	0.031 (0.033)	-0.054 (0.058)	0.058 (0.071)
$\Delta\ln(S)$	0.169** (0.067)	-0.188*** (0.070)	0.258** (0.112)	-0.350 (0.224)
R&D	-0.144 (0.097)	-0.026 (0.120)	-0.122 (0.145)	0.245 (0.190)
innovate	0.174** (0.078)	0.135 (0.104)	0.014 (0.085)	-0.010 (0.356)
mark_int	0.124 (0.183)	0.268* (0.158)	-0.168** (0.085)	0.158 (0.463)
train	0.657*** (0.088)	0.683*** (0.159)	-0.368*** (0.098)	0.049 (1.776)
odi	-0.299*** (0.095)	-0.272*** (0.104)	-0.151 (0.185)	-0.289 (0.530)
invest_nk	-0.043 (0.311)	-0.167 (0.320)	-0.087 (0.293)	0.394 (0.637)
union_nk	-0.411** (0.191)	-0.232 (0.188)	0.118 (0.225)	0.062 (0.844)
odi_nk	0.228 (0.357)	0.288 (0.406)	-0.259 (0.404)	-0.189 (1.211)
fdi_nk	-0.459* (0.271)	-0.161 (0.285)	0.448 (0.275)	0.281 (0.752)
Constant	-1.287*** (0.348)	-1.089*** (0.391)	1.506*** (0.581)	1.581 (5.942)
Observations	4,554	3,393	4,665	3,471
χ^2 Prob> χ^2	610.999 0.000	347.060 0.000	666.658 0.000	2253.339 0.987
ρ Wald Prob>χ2	0.997 0.378 0.539	0.895*** 21.543 0.000	-1.000 0.076 0.783	0.055 0.000 0.000

Robust standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*
Weighted and stratified
All specifications include industry dummies

One final note to make about our results is that ρ, the coefficient of correlation between the two equations, is statistically significant only in column (4a), the specification with the first stage modelled as reporting any vacancies and using lagged variables. The value of ρ is similar in column (4), but is insignificant. One reason for this not being significant for columns (5) and (5a) is that the model with the hard-to-fill vacancies as the selection equation is badly identified. This accords with our *a priori* expectation that the separation between the more internal, firm-specific decision of posting a vacancy and the more external, labour-market influenced probability of having a skill shortage vacancy is a better model.

4 Conclusions

In this paper we have investigated the influences on firms posting vacancies and whether these prove difficult to fill. In particular, we have considered the determinants of external skill gaps or 'skill shortage vacancies'. In particular, we have considered a Heckman selection model of the determinants of skill shortage vacancies that accounts for the fact that we only observe these external skill shortages for firms that have vacancies. This model allows for there to be a relationship between these two things. Our preferred specification (using lags to overcome problems of endogeneity) suggests that it is important to account for this correlation. Once we have accounted for this interrelationship, we find that some of the variables that at first glance appear to 'cause' skill shortage vacancies, may not be related to the difficulty of filling vacancies *per se*, but rather the likelihood of having a vacancy in the first place.

These results, however, are preliminary and there are a number of ways in which we will seek to develop the analysis. We need to think carefully about which variables should enter into the selection and which are direct determinants of skill shortage vacancies. This is true not only from the perspective of economic interpretation, but also from the statistical perspective of identification. Our focus with respect to the skill shortage equation is on labour market variables. Some of these will have a geographic element, such as measures of relative wages from a regional perspective or perhaps other indicators of labour market tightness such as unemployment rates. In Grimes, Ren and Stevens (2009), we matched single plant firms in the BOS to regions and multi-plant firms to their predominant regions, using the Business Frame and the plant level information on employment held in the LEED dataset (see also Maré, 2008). In a future version of this paper, we will consider creating a more sophisticated measure that is based on employment weighted local labour market variables sourced from either LEED or possibly even the census (see, for example, Maré, Fabling and Stillman, 2010).

In a similar way to Grimes, Ren and Stevens (2008), we can also use the BOS to create aggregate variables (similar to Grimes *et al.*'s industry knowledge intensity variable) that may improve our estimation. It might even be possible, for example, decompose occupation-level wage rates from total wage bills and numbers of staff by occupation.

Other ways we shall be looking to improve our analysis is in our measurement of training and business strategy. We can break the former down into training for new staff, training for existing staff in new roles and that for existing staff in their existing roles. We also have limited (categorical) information on the proportion of each type of staff undertaking each type of training. This data is only available for 2008, but results from the companion piece to this (Mason, Mok, Nuns, Stevens and Timmins, 2010) might help to inform which variables will be useful to instrument this variable.

Appendix

The data come from Statistics New Zealand's prototype Longitudinal Business Database (LBD).

The LBD is built around the Longitudinal Business Frame (LBF), to which are attached, among other things, Goods and Services Tax (GST) returns, financial accounts (IR10) and aggregated Pay-As-You-Earn (PAYE) returns, all provided by the Inland Revenue Department (IRD). The full LBD is described in more detail in Fabling, Grimes, Sanderson and Stevens (2008) and Fabling (2009). The survey data considered in this paper relate to the Business Operations Survey (BOS).

The administrative data we use have four sources: the Linked Employer Employee Database, the Business Activity Indicator (BAI) dataset, and IR10 forms. These are described in more detail in the Data Appendix.

The Business Operations Strategy

The Business Operation Survey (BOS) is an annual three part modular survey, which began in 2005. The first module is focussed on firm characteristics and performance. The second module alternates between biennial innovation and business use of ICT collections. The third module is a contestable module that enables specific policy-relevant data to be collected on an *ad hoc* basis[6]. The BOS is conducted using two-way stratified sampling, with stratification on rolling-mean-employment (RME) and two-digit industry according to the ANZSIC system[7]. The survey excludes firms with fewer than six RME and firms in the following industries: M81 Government Administration, M82 Defence, P92 Libraries, Museums and the Arts, Q95 Personal Services, Q96 Other Services, and Q97 Private Households Employing Staff. The 2008 survey achieved an 81.1% response rate (after adjusting for ceases), a total of 5,543 usable responses, representing a population of 36,075 firms.

The BOS is something approaching best practice in such surveys internationally. It has removed replication of surveys[8] – and thus reduces respondent load and makes

[6] In 2005 and 2009 this was a 'Business Practices Module' and in 2006 an 'Employment Practices Survey'. The 2007 module was on 'International Engagement'.

[7] Note that there was some minor additional stratification conducted at the three-digit level.

[8] Prior to the BOS, surveys tended to occur on a fairly ad hoc basis – one assumes when policy-makers were considering a particular issue. Thus there was a Business Practices Survey in 2001, an Innovation Survey in 2003 and a Business Finance Survey in (2004). Elements of each of these are considered either every year as part of the Business Performance Module (Module A) or every two or more years (i.e.

sampling simpler. It is explicitly designed with a panel element, enabling more sophisticated analysis to be undertaken allowing us to better understand issues of causality and – as the panel element increases – dynamic issues[9].

The data we use here has been edited by SNZ to remove any coding errors. A common edit is for financial data where there are components and totals. If there is no total amount (e.g. 'operating revenue', Q10), but all the components (e.g. 'Sales of goods and services', Q8; and 'other operating revenue', Q9) contain data, then the total is calculated from the components. If the total does not equal the sum of the components, then an alert is displayed and this may be manually edited (e.g. if the figures in one number are clearly transposed, or there is a scanning error).
We do not use SNZ-imputed values in cases of item non-response where it is impossible to obtain them by simple edit rules (e.g. more than one expenditure categories are missing).

The 'Business Strategy and Skills' (BIS) Module

The Business Strategy and Skills (BIS) module of the 2008 Business Operations Survey was produced as part of the 'Impact of Skills on New Zealand Firms' project. This project involved the Ministry of Economic Development, the Department of Labour, New Zealand Treasury and the Ministry of Research, Science and Technology and was partly funded by the Cross Departmental Research Pool. The module was designed by the project team in conjunction with Statistics New Zealand and Geoff Mason, from the National Institute of Economic and Social Research in London.

The BOS Variables

Collective agreements (union)

This variable relates to question 36 of Module A: 'As at the end of the last financial year, what percentage of this business's employees were covered by a collective employment agreement?' (Data item A3600.) The variable union takes the value of 1 if the respondent reports any value above zero and zero otherwise.

---

the Innovation Module is run every other year and the Business Practices Module was run in 2005 and is scheduled to repeat in 2009).

[9] The panel element is in fact larger than it first seems as there is considerable overlap with previous surveys, such as the 2001 Business Practices Survey (Fabling, 2007a).

Market focus (*mark_int*)

This variable relates to question 2 of Module C: 'In the last 2 financial years, what market accounted for the largest proportion of this business's total sales of goods or services?' Data item (C0200). Respondents could answer one of either 'local', 'national' or 'international'. The variable *mark_int* takes the value of 1 if the respondent indicates that their market focus is 'international' and zero otherwise.

Training (*train*)

This variable relates to question 22 of Module C. Data item (C2200). Respondents were asked: 'During the last financial year, have the staff of this business received training of any type?' The variable *train* takes the value of 1 if the response is yes, zero otherwise.

Ownership of overseas businesses (*odi*)

This variable relates to question 25 of Module A. Data item (A2500). The question asked is 'As at the end of the last financial year, did this business hold any ownership interest or shareholding in an overseas located business (including its own branch, subsidiary or sales office)? The variable *odi* takes the value of 1 if the response is yes, zero otherwise.

Foreign ownership of business (*fdi*)

This variable relates to question 26 of Module A. Data item (A2600). The question is 'As at the end of the last financial year, did any individual or business located overseas hold an ownership interest or shareholding in this business?' The variable *fdi* takes the value of 1 if the response is yes, zero otherwise.

Competition (*monopoly*, *oligopoly*, *compete*)

Competition is measured through binary variables monopoly, oligopoly and compete. These variables relate to question 47 of Module A. Data item (A4600). The respondents were asked 'How would you describe this business's competition?' They were given five potential responses. From these were created four binary variables (with 'many competitors, none dominant the baseline category). These are outlined in Table 9.

Table 9 Competition variables

Response	***Variable***
Captive market/no effective competition	*monopoly*
No more than one or two competitors	*oligopoly*
Many competitors, several dominant	*Baseline category*
Many competitors, none dominant	*compete*
Don't know	*comp_nk*

Investment in expansion (*invest*)

This variable relates to question 21 of Module A. Data item (A2100). The question was 'For the last financial year, did this business invest in its expansion?' Respondents were asked to include: 'purchase of one or more business assets (e.g. land, buildings, equipment)'; 'development or introduction of new or significantly improved goods, services or processes'; and 'entry into new markets'. They were asked to not include: 'increases in turnover for existing business'; or 'ongoing operational expenses'. The variable *invest* takes the value of 1 if the response is yes, zero otherwise.

Research and Development (*R&D*)

This variable relates to question 23 of Module A. Data item (A2300). The question is: 'For the last financial year, did this business undertake or fund any research and development (R&D) activities?' The respondents are asked to include: 'any activity characterised by originality: it should have investigation as its primary objective, and an outcome of gaining new knowledge, new or improved materials, products, services or process'; or 'the buying abroad of technical knowledge or information'. They were asked to not include: 'market research'; 'efficiency studies'; or 'style changes to existing products'. The variable *rnd* takes the value of 1 if the response is yes, zero otherwise.

Occupational breakdown of staff (*prop_man*, *prop_prof*, *prop_tech*, *prop_trade*, *prop_cleric*)

This variable relates to questions 10-13 of Module C. Note that every year in Module A, businesses are asked to provide a breakdown of their staff by four occupations (A3201-A3204). Respondents are asked to copy the totals from Module A into boxes in Module C. They are then asked to provide a further breakdown of two of these ('Managers and professionals' and 'all other occupations'). We calculate the proportion of the workforce in each of the occupations, with 'labourers, production, transport or other workers' as the baseline category.

Table 10 Staff occupation/role variables

Occupation/role	Data item	Variable (divided by total employment)
Managers *(i.e. those who supervise staff or determine policy and future direction)* [10]	C1002	*prop_man*
Professionals *(i.e. those who have specific expertise, but no managerial responsibility)*[11]	C1003	*prop_prof*
Technicians and associate professionals *Technicians and associate professionals perform complex technical or administrative tasks, often in support of professionals or managers (e.g. technical officer, building inspector, legal executive)*	C1101	*prop_tech*
Tradespersons and related workers *Tradespersons and related workers perform tasks requiring trade specific technical knowledge. Include all apprentices and trade supervisors (e.g. electrician, mechanic, hairdresser, baker).*	C1201	*prop_trade*
Clerical, sales and service workers *(i.e. those who perform administrative, sales or customer service tasks)*	C1302	*prop_cleric*
Labourers, production, transport or other workers *(i.e. those who operate vehicles or equipment or perform manual tasks)*	C1203	*Baseline*

Innovation (*innovate*)

This variable relates to question 42 of Module A. Data item (A4200). The question is: 'In the last financial year, did this business develop or introduce any new or significantly improved: goods or services; operational processes; organisational/managerial processes; marketing methods?' The variable *innovate* takes the value of 1 if the response is yes, zero otherwise.

[10] In Module A, where Mangers and professionals are grouped together, there are separate descriptions of managers and professionals. In addition to the description of managers given in the question in Module C, respondents are also offered two examples: 'General Manager' and 'Finance Manager'

[11] In Module A, respondents are also offered a different description: 'Professionals perform analytical, conceptual or creative tasks with skills equivalent to a bachelor degree or higher (e.g. accountant, engineer, journalist, computer programmer)'.

LEED/PAYE Data

Our data on employment come from the Linked Employer-Employee Database. It has two components, counts of employees and working proprietors.

Employees

Employment is measured using an average of twelve monthly PAYE employee counts in the year. These monthly employee counts are taken as at 15th of the month. This figure excludes working proprietors and is known as Rolling Mean Employment (RME).

Working proprietors

The working proprietor count is the number of self-employed persons who were paid taxable income during the tax year (at any time). In LEED, a working proprietor is assumed to be a person who (i) operates his or her own economic enterprise or engages independently in a profession or trade, and (ii) receives income from self-employment from which tax is deducted.

From tax data, there are five ways that people can earn self-employment income from a firm:

- As a sole trader working for themselves (using the IR3 individual income tax form [this is used for individuals who earn income that is not taxed at source]);
- Paid withholding payments either by a firm they own, or as an independent contractor (identified through the IR348 employer monthly schedule);
- Paid a PAYE tax-deducted salary by a firm they own (IR348);
- Paid a partnership income by a partnership they own (IR20 annual partnership tax form [this reports the distribution of income earned by partnerships to their partners] or the IR7 partnership income tax return);
- Paid a shareholder salary by a company they own (IR4S annual company tax return [this reports the distribution of income from companies to shareholders for work performed (known as shareholder-salaries)]).

Note that it is impossible to determine whether the self-employment income involves labour input. For example, shareholder salaries can be paid to owner-shareholders who were not actively involved in running the business. Thus there is no way of telling what labour input was supplied, although the income figures do provide some relevant information (a very small payment is unlikely to reflect a full-year, full-time labour input).

Labour turbulence and growth – accessions and separations

Labour turbulence ($Turbulence$) is measured as the annualised number of accessions to the firm (A) plus the separations (S) divided by RME. That is:

(7) $$Turbulence_i = \frac{(A_i + S_i)}{RME_i}$$

Net employment growth (NEG) is measured as the change in total employment (RME plus working proprietors) total employment. It is formally equivalent to the number of accessions less the number of separations except that working proprietors are not included in accessions and separations. That is

(8) $$NEG_{it} = \frac{(E_{it} - E_{it-1})}{E_{it-1}}$$

Wages

Wages are calculated as 'total employee gross earnings' from the LEED database, divided by RME (i.e. excluding working proprietors). This data comes from the Employers Monthly Schedule (EMS).

Business Activity Indicator (BAI) and Financial Accounts (IR10)

The Business Activity Indicator uses GST data from the Inland Revenue matched to the Statistics NZ Business Frame. The BAI data come from the Goods and *Services Tax return form*, GST 101. In order to create the BAI dataset, Statistics NZ temporarily apportion the data down to a monthly frequency, apportion returns across GST group members and apply limited imputation in cases where a single return appears to be missing. As noted in Fabling *et al.* (2008), the GST-based sales and purchases data is potentially contaminated by capital income and expenditure. In particular this includes sales of second-hand assets and businesses, purchases of land, buildings, plant, machinery and businesses. For more on this subject see section 5.4 of Fabling *et al.* (2008).

We calculate the change in stocks from page 1 of the IRD form Accounts information IR10 form. More information on what should appear in the IR10 form can be found in the IRD guide IR10G.

Sales

The sales data in the BAI relate to 'Total sales and income for the period (including GST and any zero-rated supplies).' This is adjusted using data on zero-rated sales as follows

(9) $$S_E = \frac{8}{9}(S_I - Z) + Z$$

where S_E = Sales excluding GST, S_I = Sales including GST, Z = zero rated sales.

Purchases

The purchases data in the BAI also come from the *Goods and services tax return form*, GST 101. They relate to 'Total purchases and expenses (including GST) for which tax invoicing requirements have been met' as include an estimate for imported goods and the use of private goods and services in taxable activity.

Change in stocks

The change in stocks data comes from the IR10 financial accounts form. It is calculated as closing stocks less opening stocks.

Labour Productivity

Labour productivity is calculated from the BAI, IR10 and LEED data. Value added is calculated as sales minus purchases from the BAI adjusted for the change in stocks from the IR10. The variable LP is the log of value added less the log of total RME (rolling mean employees plus the count of working proprietors).

References

Ashton, D., Green, F., Sung, J. and James, D. (2002), 'The evolution of education and training strategies in Singapore, Taiwan and S. Korea: a development model of skill formation,' *Journal of Education and Work*, 15.1, pp. 5-30.

Bennett, Jessica and Seamus McGuinness, (2009), 'Assessing the impact of skill shortages on the productivity performance of high-tech firms in Northern Ireland', *Applied Economics*, vol. 41(6), pages 727-737.

Bosworth, Derek, (1993), 'Skill Shortages in Britain', *Scottish Journal of Political Economy*, 40.3, pp. 241-271

Bloom, Nicholas and John Van Reenen, (2007), 'Measuring and Explaining Management Practices Across Firms and Countries', *Quarterly Journal of Economics*, Vol. 122, No. 4, pp. 1351–1408.

Bloom, Nicholas and John Van Reenen, (2010), 'Why Do Management Practices Differ across Firms and Countries?', *Journal of Economic Perspectives*, Volume 24, Number 1, pp. 203–224.

Durbin, Sid, (2004), 'Workplace Skills, Technology Adoption and Firm Productivity: A Review', New Zealand Treasury Working Paper 04/16.

Fabling, Richard and Arthur Grimes, (2007), 'HR Practices and Firm Performance: What matters and who does it?', MED Occasional Paper 07/02

Forth, John, and Geoff Mason, (2004), 'Information and Communication Technology (ICT) Skill Gaps and Company-Level Performance: Evidence from the ICT Professionals Survey 2000-01', National Institute of Economic and Social Research, London. NIESR Discussion Paper No. 236. Downloadable at: http://www.niesr.ac.uk/pubs/dps/dp236.PDF

Green, F.; Machin, S. and Wilkinson, D. (1998), 'The meaning and determinants of skills shortages', *Oxford Bulletin of Economics and Statistics*, 60(2)

Grimes, Arthur, Cleo Ren and Philip Stevens, (2009), 'The Need for Speed: Impacts of Internet Connectivity on Firm Productivity', Motu Working Paper 09-15

Haskel, J. and Martin, C. (1993a), 'Do Skill Shortages Reduce Productivity: Theory and Evidence from the United Kingdom', *Economic Journal*, 103, pp. 386-394

Haskel, J. and Martin, C. (1993b), 'The causes of skill shortages in Britain', *Oxford Economic Papers*, 45(4), pp. 573-588

Haskel, Jonathan and Christopher Martin, (2001), 'Technology, wages and skill shortages: evidence from UK micro data', *Oxford Economic Papers*, 53, pp. 642-658.

Heckman, James, (1979), 'Sample selection bias as a specification error', *Econometrica*, 47, pp. 153-161

Maré, David C., (2008), 'Labour Productivity in Auckland Firms', MED Occasional Paper Number 08/09. Downloadable at: http://www.med.govt.nz/templates/MultipageDocumentTOC____38579.aspx

Maré, David C., Richard Fabling and Steven Stillman, (2010), 'Immigration and Innovation', draft Motu Working Paper, Motu Economic and Public Policy Research.

Mason, Geoff, (2005), 'In Search of High Value Added Production: How Important are Skills? Investigations in the Plastics Processing, Printing, Logistics and Insurance Industries in the UK', Research Report RR663, Department for Education and Skills, London.

Mason, Geoff, Penny Mok, Peter Nuns, Philip Stevens and Jason Timmins, (2010), 'The Training's A-Comin' In: An analysis of training decisions using microdata', paper presented at the 2010 NZAE Annual Conference, Auckland.

Mason, Geoff and Rob Wilson, (eds.), (2003), 'Employers Skill Survey: New Analyses and Lessons Learned', Report NALL1, Department for Employment and Skills, London. Downloadable at: http://www.dfes.gov.uk/research/data/uploadfiles/RR663.pdf

Mason, Geoff and Philip Stevens, (2003), 'The determinants of HTF and SSV in key occupational groups in Employers Skill Survey', Chapter 7 of Mason and Wilson (eds.) (2003).

Razzak, Weshah A. and Jason C. Timmins, (2008), 'A macroeconomic perspective on skill shortages and the skill premium in NZ', *Australian Economic Papers*, 47(1), pp. 74-91.

SNZ (2009), 'Business Operations Survey: 2008 Tables', Statistics New Zealand, http://stats.govt.nz/methods_and_services/access-data/tables/business-op-survey-2008.aspx

Stevens, Philip Andrew, (2007), 'Skill shortages and firms' employment behaviour', *Labour Economics*, vol. 14(2), pages 231-249.

Tan, Hong, Yevgeniya Savchenko, Vladimir Gimpelson, Rostislav Kapelyushnikov and Anna Lukyanova, (2007), 'Skills Shortages and Training in Russian Enterprises', IZA Discussion Paper No. 275.

UTS (2009), 'Management Matters in New Zealand - how does manufacturing measure up?', Report for Ministry of Economic Development by University of Technology, Sydney. Downloadable at: http://www.med.govt.nz/templates/MultipageDocumentTOC____43278.aspx

Van de Ven, W.P.M.M., and B.M.S. Van Pragg, (1981), 'The demand for deductibles in private health insurance: A probit model with sample selection, *Journal of Econometrics*, 17, pp. 229-252.