



A GOOD WORKER IS HARD TO FIND: THE DETERMINANTS OF SKILLS SHORTAGES IN NEW ZEALAND FIRMS

Geoff Mason^{*}, Penny Mok[†], Peter
Nunns[‡], Philip Stevens[†], Jason
Timmins[§]

^{*} *National Institute of Economic and
Social Research, London*

[†] *Ministry of Economic Development*

[‡] *Formerly of Ministry of Economic
Development*

[§] *Department of Labour*



Abstract

This paper investigates the determinants of firms' 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). We estimate the determinants of firms reporting having vacancies and skill shortage in comparison to the non-skill-related vacancies using probit models and two-stage (Heckman) probit models with selection, to enable us to distinguish between those factors related with the firm posting a vacancy and those that make them difficult to fill. 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 Codes: J24, J31, L60.

Keyword(s): skill shortages, hard-to-fill vacancies, Business Operations Survey, probit.

Disclaimer: The opinions, findings, recommendations and conclusions expressed in this report are those of the author(s). Statistics New Zealand, the Ministry of Economic Development, the Department of Labour and the National Institute of Economics and Social Research, take no responsibility for any omissions or errors in the information contained here. Access to the data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular, business or organisation. The results in this paper have been confidentialised to protect individual businesses from identification. The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information is published or disclosed in any other form, or provided back to Inland Revenue for administrative or regulatory purposes. Any person who had access to the unit-record data has certified that they have been shown, have read and have understood section 81 of the Tax Administration Act 1994, which relates to privacy and confidentiality. Any discussion of data limitations or weaknesses is not related to the data's ability to support Inland Revenue's core operational requirements.

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 information available 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. In addition, we compare these determinants with vacancies that are hard to fill for non-skill-related reasons. We utilise 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 other data in the prototype Longitudinal Business Database (LBD) to investigate the determinants vacancies and external skill shortages. We investigate three types of vacancy: firms that had any vacancies in the last year; vacancies that were hard to fill because the applicants

lacked the necessary skills, qualification or experience which we defined as skill-related reasons; and vacancies that were hard to fill for non-skill-related reasons. 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 and non-skill-related shortage vacancies are by definition hard-to-fill vacancies. In addition, a proportion of firms have vacancies while other firms do not recruit so it is important to account for the factors which influence the reason firms recruit as opposed to those who do not. Therefore, we estimate a two-stage (Heckman) probit model with selection.

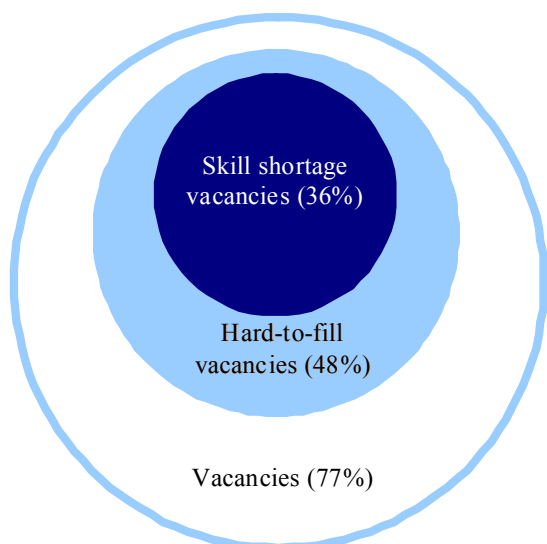
By matching the BSS to the LBD we can consider a broad suite of variables, including firm's size and industry, their market focus, wages, R&D investment, innovation, previous performance (e.g. productivity) and the degree of competition they are subject to.

This paper is organised as follows. We begin in the next section by discussing our data and presenting descriptive statistics. The empirical results are discussed in Section 3. Section 4 discusses and concludes.

2. Data

In this paper we consider all firms that report vacancies of any kind and skill shortage vacancies. The skill shortage vacancies are defined as vacancies that were hard-to-fill because the applicants lacked the necessary skills, qualification or experience, which is a subset of hard-to-fill vacancies. As there are other reasons for having hard-to-fill vacancies, this allows us to distinguish the non-skill-related vacancies from the 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, skill and non-skill-related 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

2.1 Vacancies

Respondents were asked: ‘During the last financial year, has this business had any vacancies?’ (question 14 of Module C). The first row of Table 1 summarises firms’ responses. 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.

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 picture 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).

2.2 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 1: Vacancies, hard-to-fill, skill and non-skill-related shortage vacancies, %

	Business size				Overall
	E <20*	20 ≤ E <50	50 ≤ E <100	E ≥ 100	
Vacancies	71.8	89.4	93.2	94.9	76.6
Hard-to-fill Vacancies	43.2	58.4	65.0	73.4	47.9
Skill Shortage Vacancies	32.1	43.0	48.7	58.7	35.7
Non-skill-related shortage vacancies	11.5	15.6	16.8	14.6	12.6

- 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.

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.

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

Occupation	Business size				Overall
	E <20*	20≤E<50	50≤E<100	E≥100	
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	65.0	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 Skill and non-skill-related 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) (For the full list of responses, see SNZ, 2008). Those that replied having other reasons besides 'applicants lack the work experience the business demands' or 'applicants lack the qualifications or skills the business demands' were

defined as having non-skill-related shortage vacancies (NSR). This unique information allows us to distinguish the determinants of vacancies between skill and non-skill-related reasons.

The third and final row of Table 1 summarises the data on these vacancies. Almost three quarters of firms with hard-to-fill vacancies reported external skill shortages (35.7% compared with 47.9%). The remaining quarter of firms with hard-to-fill vacancies reported having only non-skill-related shortage vacancies (12.6% compared with 47.9%).

We break the skill shortage vacancies 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
<i>Business size</i>		
6-19 Employees	26	26
20-49 Employees	36	37
50-99 Employees	39	40
100+ Employees	48	48
<i>Industry</i>		
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. Econometric results

3.1 Probit model

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

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

where y_i^* is the propensity to have a vacancy, a skill or non-skill-related 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_i^* terms, but the binary realisation of them, therefore we assume:

$$y_i = 0 \quad \text{if} \quad y_i^* < 0$$

$$y_i = 1 \quad \text{if} \quad y_i^* \geq 0 \quad (2)$$

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 (using the `svy: probit` command in Stata).

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 in the variables (e.g. Green, Machin and Wilkinson, 1998) or both (e.g. Haskel and Martin, 2001).

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. In comparison, we have the second specification which has as its second stage the probability of having a non-skill-related shortage vacancy. We describe this model in more detail in 3.4 below.

The variables we include are the employment, employment separations (separations over employment), wages relative to the 1-digit industry, multi-factor 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. Firm productivity is measured through the multifactor productivity, in which is the residual of the regression of value added on capital and labour (Iyer, *et al.* 2010). This measurement is a better representation of the true productivity of the firm. We have estimated the multifactor productivity using the Levinsohn and Petrin (2003) procedure to correct for endogeneity bias as opposed to the OLS procedure. The conventional labour productivity measurement is also used to compare these measurements. The variables themselves are set out in more detail in the Data Appendix to this paper.

In the next section, we present the results of our probit models of reporting vacancies, skill and non-skill-related shortage vacancies. We present two versions of each of these. In section 3.2 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 3.3 we used lagged variables to account for endogeneity issues.

3.2 Contemporaneous variables

There are a number of interesting results that emerge from the results on the probability of a firm posting a vacancy shown in Table 5. Firms with a high degrees of separation relative to their employment are more likely to report vacancies and non-skill-related vacancies, but are no more likely to report skill shortage vacancies. There is no evidence that more productive firms are more likely to have any types of vacancies.

Fast growing firms, in terms of sales, do appear to report less non-skill-related shortage vacancies. There is a similar sign to the coefficient on for column (1), but this is not statistically significant. Firms which pay higher relative wage to their staff report more vacancies and skill shortage vacancies but less likely to report having non-skill-related vacancies. However, these are not always statistically significant.

The probability of a firm reporting a skill shortage vacancy increases with the proportion of professional employees (NB the proportion of labourers is the reference group). Firms with a greater proportion of professionals are less likely to report having non-skill-related vacancies. Similar signs are displayed for proportion of tradesperson to the total employment in column (2) and proportion of technicians to the total employment in column (3). There is evidence that suggests foreign-owned businesses find it easier to fill the skilled vacancies (column (2)). Firms that provide training to their staff are more likely to report having vacancies of all types.

3.3 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 (1a) to (3a) 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 columns (1a) to (3a) of Table 5. Many of the results remain, although there are a few differences. Firms that train their staff are more likely to report having vacancies of any kind. The lagged employment separation shows consistent results using the current variables for all vacancies and non-skill-related vacancies (although not for skill shortage vacancies).

Table 5: Probit regressions using contemporaneous and lagged variables

	Current variables			Lagged variables		
	(1) All Vacancies	(2) Skill shortage vacancies	(3) Non-skill- related vacancies	(1a) All Vacancies	(2a) Skill shortage vacancies	(3a) Non-skill- related vacancies
Employment	-0.145 (0.191)	-0.00342 (0.175)	0.173 (0.195)	0.057 (0.203)	-0.253 (0.204)	0.423* (0.200)
mfp_lp	0.062 (0.084)	-0.0195 (0.066)	0.0885 (0.074)	0.017 (0.092)	-0.051 (0.082)	0.056 (0.094)
relative wage	0.373* (0.163)	0.175 (0.155)	-0.0748 (0.169)	0.196 (0.163)	0.380* (0.183)	-0.204 (0.180)
separation	1.465** (0.465)	0.288 (0.386)	1.340** (0.397)	1.327* (0.549)	0.559 (0.468)	1.145* (0.523)
rnd	-0.203 (0.228)	-0.0166 (0.164)	-0.309 (0.215)	0.273 (0.219)	-0.029 (0.171)	-0.378 (0.220)
innovate	0.164 (0.152)	0.0322 (0.108)	-0.162 (0.123)	0.102 (0.157)	0.059 (0.130)	0.072 (0.157)
invest	0.165 (0.162)	0.0622 (0.112)	0.151 (0.122)	0.027 (0.171)	-0.010 (0.127)	-0.130 (0.139)
international market	0.0700 (0.214)	-0.0159 (0.201)	0.0762 (0.226)	0.116 (0.268)	0.063 (0.241)	0.020 (0.287)
technology	-0.166 (0.131)	-0.0173 (0.114)	0.101 (0.128)	0.019 (0.150)	-0.029 (0.141)	0.152 (0.162)
compare_competitor	0.0857 (0.064)	-0.0169 (0.059)	0.0226 (0.071)	0.040 (0.073)	0.082 (0.066)	-0.133 (0.072)
bus_strategy	0.0259 (0.0495)	-0.00301 (0.047)	0.0442 (0.050)	0.048 (0.054)	0.009 (0.055)	0.045 (0.055)
odi	0.196 (0.312)	-0.117 (0.180)	-0.0609 (0.216)	0.012 (0.322)	-0.290 (0.193)	-0.028 (0.194)
fdi	-0.0661 (0.215)	-0.368* (0.148)	0.180 (0.202)	-0.064 (0.229)	-0.316 (0.163)	-0.249 (0.157)
export	0.134 (0.233)	-0.0725 (0.200)	0.249 (0.225)	0.070 (0.278)	0.079 (0.216)	-0.304 (0.243)
monopoly	-0.276 (0.289)	-0.186 (0.268)	0.535 (0.298)	-0.558 (0.332)	-0.405 (0.322)	0.713* (0.355)
duopoly	-0.0517 (0.149)	0.132 (0.131)	-0.0308 (0.132)	-0.079 (0.166)	0.260 (0.154)	-0.069 (0.155)
competitive	0.0857 (0.147)	0.0637 (0.131)	0.151 (0.158)	-0.287 (0.168)	0.163 (0.150)	0.180 (0.161)
sales growth	-0.0293 (0.180)	0.262 (0.154)	-0.555** (0.185)	-0.045 (0.253)	-0.329 (0.244)	0.481* (0.197)
relative wage growth	0.287 (0.184)	-0.322 (0.170)	0.423** (0.160)	0.269 (0.269)	0.303 (0.229)	-0.123 (0.235)
union	-0.0276 (0.153)	0.205 (0.123)	0.00827 (0.123)	-0.126 (0.155)	0.113 (0.138)	0.041 (0.139)
prop_man	-0.621 (0.538)	-0.186 (0.467)	0.433 (0.502)	-1.052 (0.613)	-0.523 (0.558)	0.563 (0.640)

	Current variables			Lagged variables		
	(1) All Vacancies	(2) Skill shortage vacancies	(3) Non-skill- related vacancies	(1a) All Vacancies	(2a) Skill shortage vacancies	(3a) Non-skill- related vacancies
prop_prof	-0.681 (0.466)	0.885* (0.404)	-0.904* (0.442)	-0.717 (0.567)	0.611 (0.521)	-0.499 (0.520)
prop_tech	-0.493 (0.355)	-0.0367 (0.335)	-0.720* (0.364)	-0.057 (0.461)	0.113 (0.439)	-0.352 (0.418)
prop_trade	0.105 (0.259)	0.764** (0.207)	-0.266 (0.277)	0.030 (0.278)	0.788** (0.239)	-0.328 (0.290)
prop_cleric	-0.407 (0.273)	-0.0164 (0.216)	-0.377 (0.240)	-0.376 (0.297)	0.172 (0.247)	-0.201 (0.268)
train1	0.838** (0.204)	0.244 (0.147)	0.194 (0.192)	0.831** (0.253)	0.212 (0.175)	0.183 (0.247)
train2	0.703** (0.086)	0.383** (0.068)	0.226** (0.081)	0.618*** (0.099)	0.375*** (0.080)	0.172 (0.088)
constant	-1.237 (0.991)	-1.240 (0.752)	-3.132** (0.851)	-0.692 (1.083)	-1.038 (0.935)	-2.739* (1.075)
Observation	2544	2544	2535	1893	1899	1896
Industry Dummies	Yes	Yes	Yes	Yes	Yes	Yes
F test	4.52	4.47	3.04	4.18	3.10	4.85
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000

Standard errors in parentheses

*significant at 5%, ** significant at 1%

Stratified and weighted

[†] lagged one year (2007)

3.4 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 (2)) are a subset of those reporting having a vacancy. 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 skill and non-skill-related shortage vacancies accounting for sample selection. In both models, we model the first stage (the selection equation) as being the probability of reporting having posted a vacancy.

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

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

and a probit equation for the binary outcome:

$$SSV_i = (SSV_i^* > 0) \quad (4)$$

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, non-skill-related vacancies (NSR) takes the value zero if the firm does not have a NSR vacancy, but does have a vacancy. It takes a missing value when the firm does not have any vacancies.

Thus the dependent variable in 0 for observation i is observed if:

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

where

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

When $\rho \neq 0$, standard probit techniques applied to 0 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 0 that is not in the outcome equation. We estimate y^{select} whether a firm had any vacancies. We estimate both SSV and NSR models using current values of the variables and lagged values.

The results of our Heckman probit estimation are presented in Table 6 and Table 7. Looking first at the selection equations in Table 6, firms with more employees and trainings are more likely to have vacancies for both models and in both current and lagged values. Also, firms with more separation workforces are also more likely to have vacancies with the exception of the lagged values in the NSR model. Firms with increasing wages are also more likely to post vacancies (although this effect is not true for the lagged values).

Turning to the outcome equations in Table 7, we see that once we have accounted for their probability of having vacancies, larger firms are no more likely to have skill shortage vacancies. Interestingly, firms that pay higher wages relative to their (1-digit) industry are more likely to have skill shortage vacancies but the relationship is reversed in the non-skill-related vacancies. This suggests

that firms pay higher wage in response to skill-shortage vacancies but it may not be sufficient to clear the market. It is still true when we used lagged relative wages. Instead, firms could response easily to the non-skill-related vacancies by increasing the wages. Training is significant and positive in the SSV equation but is significantly negative in the NSR equations. This suggests that training could easily reduce the probability of firms having non-skill-related vacancies as opposed to SSV. Foreign firms appear to find it easier to find skilled staff but not non-skilled staff. Firms that compete in a monopoly environment are more likely to report having difficulty finding non-skilled staff.

Notably, the coefficients of correlation between the two equations (ρ), are statistically significant at 10 and 5 percent level of significance for equation 4 and 5 respectively using the current variables. This suggests that the sample selection bias exists in a subset of SSV and NSR observations. Thus, the Heckman selection technique is appropriate.

Table 6: Heckman selection model of skill and non-skill-related shortage vacancies, Selection equation

Variable	Skill shortage vacancy		Non-skill-related vacancy	
	(4)	(4a)	(5)	(5a)
	Current	Lagged	Current	Lagged
Employment	0.377*** (0.065)	0.275* (0.159)	0.360*** (0.056)	0.321** (0.136)
separation	1.032** (0.437)	1.131* (0.617)	1.159** (0.457)	0.872 (0.620)
LP	0.126** (0.056)	0.100 (0.062)	0.147*** (0.055)	0.074 (0.046)
sales growth	0.060 (0.079)	0.146 (0.137)	0.046 (0.084)	0.191** (0.097)
Wage growth	0.696*** (0.169)	0.078 (0.103)	0.551*** (0.175)	0.051 (0.097)
technology	-0.047 (0.140)	0.291** (0.134)	-0.063 (0.138)	0.271** (0.128)
innovate	0.233** (0.112)	0.009 (0.157)	0.208* (0.110)	-0.018 (0.138)
union	-0.106 (0.070)		0.019 (0.088)	
train1	0.822*** (0.243)	0.887*** (0.227)	0.931*** (0.239)	0.903*** (0.215)
train2	0.724*** (0.087)	0.660*** (0.131)	0.770*** (0.096)	0.660*** (0.105)
prop_man	-0.285 (0.383)	-0.759* (0.460)	-0.196 (0.381)	-0.214 (0.464)
prop_prof	-0.816*** (0.232)	-0.550 (0.727)	-0.466* (0.244)	-0.370* (0.199)
prop_tech	-0.275 (0.318)	0.050 (0.455)	-0.410 (0.272)	0.046 (0.243)
prop_trade	-0.114 (0.114)	-0.089 (0.463)	0.125 (0.139)	0.014 (0.144)
prop_cleric	-0.373* (0.218)	-0.462 (0.309)	-0.414** (0.208)	-0.323* (0.181)
odi	0.366 (0.245)	0.163 (0.241)	0.273 (0.234)	0.122 (0.180)
fdi	-0.044 (0.161)	-0.060 (0.207)	-0.120 (0.157)	-0.076 (0.162)
Constant	-1.040*** (0.216)	-1.026*** (0.379)	-1.115*** (0.187)	-1.122*** (0.287)
Athrho	3.908* (2.179)	0.580 (2.019)	-0.876** (0.367)	-2.317 (1.552)
Observations	2664	1995	2664	1995

Robust standard errors in parentheses

significant at 10%, ** significant at 5%, *significant at 1%*

Stratified and weighted

All specifications include industry dummies

Table 7: Heckman selection model, Determinants of skill and non-skill related shortage vacancies

Variables	Skill shortage vacancy		Non-skill-related vacancy	
	(4)	(4a)	(5)	(5a)
	Current	Lagged	Current	Lagged
Employment	-0.034 (0.129)	-0.401 (0.291)	0.244** (0.115)	0.194 (0.130)
mfp_lp	0.059 (0.060)	-0.095 (0.085)	-0.045 (0.055)	0.044 (0.073)
relative wage	0.173* (0.106)	0.492*** (0.173)	-0.223** (0.101)	-0.218** (0.098)
rnd	0.086 (0.123)	-0.149 (0.174)	-0.255 (0.186)	-0.342** (0.153)
innovate	-0.004 (0.121)	-0.043 (0.125)	-0.214 (0.147)	0.097 (0.122)
invest	-0.031 (0.090)	-0.011 (0.181)	0.175 (0.114)	-0.118 (0.122)
international market	0.005 (0.129)	0.069 (0.207)	0.112 (0.108)	0.064 (0.130)
technology	0.044 (0.112)	-0.023 (0.165)	0.128 (0.184)	-0.078 (0.194)
compare_competitor	-0.026 (0.055)	0.050 (0.074)	-0.045 (0.051)	-0.090** (0.043)
bus_strategy	0.000 (0.022)	-0.035 (0.040)	0.028 (0.025)	0.037** (0.019)
train1	0.232* (0.137)	0.066 (0.372)	-0.134 (0.128)	-0.409** (0.185)
train2	0.404*** (0.090)	0.371 (0.263)	-0.187** (0.086)	-0.390*** (0.096)
odi	-0.239 (0.195)	-0.385 (0.248)	-0.049 (0.234)	-0.057 (0.178)
fdi	-0.425*** (0.142)	-0.310* (0.163)	0.250 (0.186)	-0.073 (0.153)
export	-0.117 (0.133)	0.094 (0.209)	0.184 (0.174)	-0.276 (0.211)
monopoly	-0.146 (0.245)	-0.507 (0.385)	0.558* (0.333)	0.699** (0.350)
Constant	-1.817*** (0.637)	0.221 (2.043)	-0.172 (0.765)	-0.260 (0.663)
Observations	2664	1995	2664	1995
ρ	0.999*	0.522	-0.704**	-0.981
Wald	3.216	0.082	5.686	2.231
Prob>C	0.073	0.774	0.017	0.135
Chi2	415.774	215.007	100.582	351.183
Prob>C	0.000	0.000	0.000	0.000

Robust standard errors in parentheses

significant at 10%, ** significant at 5%, *significant at 1%*

Stratified and weighted

4. Discussion and conclusion

In this paper we have investigated the influences on firms posting vacancies and whether these prove difficult to fill due to skill and non-skill-related reasons. In particular, we have considered the determinants of external skill gaps or 'skill shortage vacancies' in comparison to the non-skill-related 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 current variables) 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.

One would expect that firms would experience fewer hard-to-fill vacancies if they pay more and undertake more training. This is true for non-skill-related vacancies but the probability of reporting skill shortage vacancies is higher for high-paying and high-training firms. Our results also show that NZ-owned firms experience more difficulties than foreign-owned firms in recruiting skilled workers. This suggests that some firms simply find it very difficult to obtain the skills they require in the country or regional labour markets and/or they have relatively high standards when it comes to recruiting skilled personnel. Our results show that it is important to distinguish between the skilled and non-skilled hard-to-fill vacancies.

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).

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
- Iyer, K., Stevens, P. & Austin, D. (2010), 'Are Non-exporters Locked out of Foreign markets because of Low Productivity?: Evidence from New Zealand Agriculture and Forestry', *Economics Bulletin*, 30(2), pp.1694-1709.
- 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.

Appendix

A1: Probit regression using contemporaneous variables

	(1b) Vacancy	(2b) Skill shortage vacancies	(3b) Non-skill related HTF	(1c) Vacancy	(2c) Skill shortage vacancies	(3c) Non-skill related HTF
Employment	-0.138 [0.203]	-0.0163 [0.176]	0.219 [0.199]	-0.158 [0.179]	-0.00768 [0.161]	0.260 [0.187]
mfp_ols	0.0301 [0.0860]	-0.0236 [0.0649]	0.0806 [0.0723]			
LP				-0.00114 [0.0726]	0.00993 [0.0583]	-0.0148 [0.0831]
Relative wage	0.394* [0.163]	0.178 [0.152]	-0.0769 [0.168]	0.426** [0.146]	0.155 [0.139]	-0.119 [0.159]
New hire	1.402** [0.474]	0.365 [0.403]	1.294** [0.451]	1.442** [0.407]	-0.0134 [0.360]	1.463** [0.423]
Rnd	-0.196 [0.227]	-0.0156 [0.163]	-0.311 [0.214]	-0.254 [0.209]	-0.164 [0.146]	-0.0371 [0.245]
Innovate	0.163 [0.153]	0.0307 [0.108]	-0.165 [0.124]	0.215 [0.135]	0.0714 [0.0964]	-0.140 [0.122]
Invest	0.166 [0.162]	0.0602 [0.112]	0.151 [0.123]	0.148 [0.143]	0.0591 [0.100]	0.0237 [0.110]
International market focus	0.0586 [0.215]	-0.0276 [0.202]	0.0552 [0.223]	0.0860 [0.200]	0.220 [0.191]	-0.105 [0.198]
technology	-0.166 [0.132]	-0.0159 [0.115]	0.109 [0.128]	-0.0765 [0.117]	0.0569 [0.100]	0.154 [0.119]
compare_competitor	0.0955 [0.0633]	-0.0156 [0.0589]	0.0315 [0.0706]	0.0636 [0.0548]	-0.0224 [0.0498]	-0.0468 [0.0695]
bus_strategy	0.0314 [0.0494]	-0.00194 [0.0472]	0.0484 [0.0502]	0.0880 [0.0455]	0.0294 [0.0408]	-0.0199 [0.0522]
Odi	0.202 [0.308]	-0.113 [0.181]	-0.0414 [0.219]	0.224 [0.277]	-0.0331 [0.214]	-0.151 [0.241]
Fdi	-0.0516 [0.215]	-0.366* [0.148]	0.196 [0.202]	-0.0414 [0.188]	-0.225 [0.140]	0.171 [0.205]
export	0.113 [0.232]	-0.0739 [0.200]	0.257 [0.226]	0.0293 [0.211]	0.102 [0.199]	0.0784 [0.220]
monopoly	-0.301 [0.297]	-0.192 [0.271]	0.484 [0.296]	-0.265 [0.262]	-0.264 [0.216]	0.402 [0.254]
duopoly	-0.0467 [0.148]	0.132 [0.132]	-0.0386 [0.133]	-0.0231 [0.132]	0.0424 [0.119]	-0.00903 [0.139]
competitive	0.0891 [0.149]	0.0625 [0.132]	0.145 [0.159]	0.0978 [0.134]	-0.0450 [0.118]	0.198 [0.136]

	(1b) Vacancy	(2b) Skill shortage vacancies	(3b) Non-skill related HTF	(1c) Vacancy	(2c) Skill shortage vacancies	(3c) Non-skill related HTF
Sales growth	-0.0305 [0.179]	0.259 [0.154]	-0.570** [0.191]	0.0773 [0.151]	0.274* [0.122]	-0.482** [0.162]
Relative wage growth	0.102 [0.183]	-0.359* [0.173]	0.252 [0.165]	0.0255 [0.162]	-0.241 [0.145]	0.204 [0.152]
union	-0.0250 [0.154]	0.205 [0.123]	0.0151 [0.124]	-0.124 [0.136]	0.109 [0.111]	0.0807 [0.115]
prop_man	-0.591 [0.541]	-0.189 [0.467]	0.475 [0.500]	-0.356 [0.490]	-0.486 [0.439]	0.278 [0.453]
prop_prof	-0.639 [0.467]	0.902* [0.406]	-0.931* [0.442]	-0.563 [0.403]	0.497 [0.353]	-0.325 [0.398]
prop_tech	-0.455 [0.357]	-0.0181 [0.336]	-0.707 [0.369]	-0.521 [0.317]	-0.128 [0.300]	-0.472 [0.358]
prop_trade	0.119 [0.258]	0.773** [0.208]	-0.244 [0.274]	0.135 [0.236]	0.795** [0.188]	-0.238 [0.272]
prop_clerical	-0.388 [0.278]	-0.00858 [0.218]	-0.339 [0.241]	-0.389 [0.246]	-0.0220 [0.187]	-0.148 [0.213]
train1	0.843** [0.204]	0.245 [0.147]	0.198 [0.193]	0.833** [0.181]	0.227 [0.137]	0.127 [0.168]
train2	0.689** [0.0865]	0.381** [0.0685]	0.216** [0.0804]	0.725** [0.0763]	0.355** [0.0617]	0.198* [0.0808]
Constant	-0.614 [0.381]	-1.446** [0.302]	-2.267** [0.341]	-0.752* [0.354]	-1.359** [0.273]	-2.008** [0.344]
Observation	2544	2544	2535	3129	3129	3120
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
F test	4.31	4.50	2.99	5.99	4.70	2.53
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000

Standard errors in parentheses

**significant at 5%, ** significant at 1%*

Stratified and weighted

A2: Probit regression using lagged variables

	(1d) Vacancy	(2d) Skill shortage vacancies	(3d) Non-skill related HTF	(1e) Vacancy	(2e) Skill shortage vacancies	(3e) Non-skill related HTF
Employment	0.0357 [0.209]	-0.290 [0.203]	0.457* [0.220]	0.146 [0.191]	-0.121 [0.166]	0.374* [0.186]
mfp_ols	0.0325 [0.0898]	-0.0500 [0.0820]	0.0892 [0.0935]			
LP				0.0470 [0.0793]	-0.00989 [0.0604]	0.116 [0.0819]
relative wage	0.222 [0.158]	0.391* [0.174]	-0.205 [0.180]	0.172 [0.146]	0.287* [0.144]	-0.206 [0.148]
New hire	1.601** [0.526]	0.782 [0.484]	1.274* [0.514]	1.587** [0.464]	0.820 [0.418]	1.378** [0.502]
Rnd	0.282 [0.220]	-0.0208 [0.172]	-0.386 [0.222]	0.185 [0.218]	0.0160 [0.166]	-0.298 [0.203]
innovate	0.101 [0.156]	0.0574 [0.130]	0.0911 [0.158]	0.0667 [0.139]	0.136 [0.114]	0.106 [0.135]
invest	0.0182 [0.173]	-0.0211 [0.128]	-0.132 [0.141]	-0.0505 [0.154]	-0.0141 [0.112]	-0.136 [0.133]
International market focus	0.114 [0.272]	0.0434 [0.239]	0.0367 [0.312]	0.101 [0.227]	0.187 [0.207]	-0.0710 [0.273]
technology	0.0137 [0.149]	-0.0268 [0.140]	0.143 [0.163]	0.139 [0.131]	-0.109 [0.120]	0.274 [0.141]
compare_competi tor	0.0425 [0.0714]	0.0831 [0.0652]	-0.126 [0.0735]	0.0431 [0.0606]	0.0466 [0.0541]	-0.129 [0.0659]
bus_strategy	0.0539 [0.0540]	0.0121 [0.0546]	0.0480 [0.0547]	0.100* [0.0478]	0.0251 [0.0473]	0.0483 [0.0497]
odi	0.0303 [0.333]	-0.280 [0.193]	-0.0189 [0.195]	-0.144 [0.260]	-0.312 [0.171]	-0.0598 [0.193]
fdi	-0.0625 [0.232]	-0.313 [0.162]	-0.249 [0.156]	-0.0775 [0.191]	-0.301* [0.136]	0.00622 [0.254]
export	0.0493 [0.278]	0.0708 [0.217]	-0.325 [0.248]	0.0647 [0.250]	0.0295 [0.193]	-0.326 [0.217]
sales growth	-0.0523 [0.251]	-0.320 [0.242]	0.464* [0.196]	0.0980 [0.210]	-0.307 [0.159]	0.337* [0.165]
Relative wage growth	0.129 [0.268]	0.221 [0.234]	-0.213 [0.232]	-0.0472 [0.242]	0.163 [0.182]	-0.277 [0.234]
union	-0.131 [0.156]	0.113 [0.138]	0.0457 [0.139]	-0.186 [0.135]	0.0578 [0.118]	-0.0956 [0.123]
prop_man	-1.120 [0.618]	-0.517 [0.556]	0.562 [0.638]	-0.632 [0.525]	-0.483 [0.510]	0.235 [0.566]

	(1d)	(2d)	(3d)	(1e)	(2e)	(3e)
	Vacancy	Skill shortage vacancies	Non-skill related HTF	Vacancy	Skill shortage vacancies	Non-skill related HTF
prop_prof	-0.777 [0.571]	0.617 [0.523]	-0.557 [0.520]	-0.710 [0.470]	0.488 [0.420]	-0.306 [0.436]
prop_tech	-0.0681 [0.465]	0.128 [0.441]	-0.378 [0.414]	-0.357 [0.390]	0.130 [0.364]	-0.791* [0.399]
prop_trade	0.0421 [0.280]	0.809** [0.238]	-0.321 [0.290]	-0.0152 [0.244]	0.871** [0.214]	-0.513 [0.263]
prop_cleric	-0.403 [0.297]	0.175 [0.247]	-0.227 [0.270]	-0.608* [0.258]	0.0681 [0.210]	-0.297 [0.236]
train1	0.812** [0.252]	0.200 [0.176]	0.153 [0.251]	0.858** [0.225]	0.131 [0.155]	0.111 [0.203]
train2	0.621** [0.0984]	0.373** [0.0808]	0.164 [0.0896]	0.653** [0.0903]	0.331** [0.0697]	0.216** [0.0776]
monopoly	-0.514 [0.330]	-0.404 [0.329]	0.756* [0.358]	-0.406 [0.282]	-0.452 [0.263]	0.427 [0.331]
Duopoly	-0.0654 [0.166]	0.267 [0.154]	-0.0651 [0.154]	-0.0436 [0.148]	0.192 [0.141]	-0.196 [0.140]
Competitive	-0.306 [0.168]	0.155 [0.149]	0.174 [0.160]	-0.149 [0.150]	0.140 [0.130]	0.114 [0.141]
Constant	-0.543 [0.425]	-1.588** [0.350]	-2.241** [0.331]	-0.863* [0.388]	-1.646** [0.298]	-2.004** [0.404]
Observation	1893	1899	1896	2421	2433	2427
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
F test	4.34	3.13	4.96	5.11	4.19	4.54
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000

Standard errors in parentheses

*significant at 5%, ** significant at 1%

Stratified and weighted

A3: Polychoric correlation matrix for Business Strategy

Variables	compare_cost	compare_time	compare_quality	compare_flex	compare_satisf	compare_emp	bus_gs	bus_price
compare_cost	1.00							
compare_time	0.45	1.00						
compare_quality	0.34	0.38	1.00					
compare_flex	0.47	0.42	0.74	1.00				
compare_satisf	0.40	0.40	0.84	0.74	1.00			
compare_emp	0.42	0.39	0.67	0.64	0.74	1.00		
bus_gs	0.10	0.15	0.32	0.32	0.28	0.12	1.00	
bus_price	0.24	0.23	0.15	0.16	0.15	0.14	0.11	1.00
Observation	34677	34677	34677	34677	34677	34677	34677	34677
rho	0.11							
SE rho	0.03							

A4: Principal Component Analysis: Business Strategy

	comp_competitor	bus_strategy
Cumulative proportion of variation explained	0.49	0.63
Variable loadings on Business strategy factor:		
compare_cost	0.61	0.44
compare_time	0.60	0.42
bus_price	0.29	0.72
compare_quality	0.87	-0.27
compare_flex	0.87	-0.14
compare_satisf	0.89	-0.23
compare_emp	0.81	-0.12
bus_gs	0.37	-0.21
Kaiser-Meyer-Olkin measure of sampling adequacy		0.86
observation	34677	34677

A5: Polychoric correlation matrix for Training

Variables	train_ new	train_ change	train_ exist	cust	com	team	trade	manag e	prof	market	oral	write	numer acy
train_new	1.00												
train_change	0.70	1.00											
train_exist	0.72	0.63	1.00										
cust	0.62	0.45	0.50	1.00									
com	0.56	0.48	0.54	0.48	1.00								
team	0.63	0.51	0.58	0.68	0.42	1.00							
trade	0.50	0.33	0.48	0.28	0.24	0.41	1.00						
manage	0.54	0.56	0.54	0.52	0.55	0.63	0.33	1.00					
prof	0.48	0.37	0.57	0.36	0.53	0.35	0.33	0.58	1.00				
market	0.39	0.34	0.38	0.72	0.44	0.53	0.24	0.54	0.38	1.00			
oral	0.52	0.44	0.54	0.65	0.40	0.72	0.31	0.61	0.30	0.54	1.00		
write	0.48	0.36	0.51	0.50	0.47	0.63	0.25	0.56	0.43	0.53	0.90	1.00	
numeracy	0.49	0.35	0.46	0.51	0.40	0.58	0.25	0.47	0.30	0.40	0.76	0.82	1.00
Observation	26244	26244	26244	26244	26244	26244	26244	26244	26244	26244	26244	26244	26244
rho	0.82												
SE rho	0.03												

A6: Principal Component Analysis: Training

	Train1	Train2
Cumulative proportion of variation explained	0.54	0.64
Variable loadings on Training factor:		
train_new	0.80	0.32
train_change	0.69	0.37
train_exist	0.78	0.33
trade	0.49	0.38
manage	0.79	0.08
prof	0.61	0.37
com	0.68	0.23
cust	0.77	-0.15
team	0.82	-0.14
market	0.68	-0.21
oral	0.82	-0.45
write	0.80	-0.44
numeracy	0.73	-0.44
Kaiser-Meyer-Olkin measure of sampling adequacy		0.77
Observation		26244

Variables

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* basis¹. The BOS is conducted using two-way stratified sampling, with stratification on rolling-mean-employment (RME) and two-digit industry according to the ANZSIC system². 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 surveys³ – and thus reduces respondent load and

¹ In 2005 and 2009 this was a 'Business Practices Module' and in 2006 an 'Employment Practices Survey'. The 2007 module was on 'International Engagement'.

² Note that there was some minor additional stratification conducted at the three-digit level.

³ 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

makes 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 .

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.

Market focus (**international market**)

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

⁴ 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).

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 *international market* takes the value of 1 if the respondent indicates that their market focus is ‘international’ and zero otherwise.

Training (**train1, train2**)

This variable relates to question 25 and 26 of Module C. Respondents were asked: ‘During the last financial year, have proportion of this business’s staff participated in training?’ and ‘During the last financial year, did this business’s staff participate in the following types of training provided or funded by this business?’. The training estimates are obtained through the principal-component-analysis.

Business strategy (**compare_competitor, bus_strategy**)

This variable relates to question 41 of Module A, question 6 and 8 of Module C. Respondents were asked: ‘How do you think this business compares to its major competitors on each of the following?’ Question 6 and 8 respectively asked: ‘Which of the following best describes this business’s goods or services?’ and ‘How often is this business able to obtain a higher price than competitors for its main goods or services?’ The business strategy estimates are obtained through the principal-component-analysis.

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**)

Competition is measured through binary variable monopoly. These variables relate to question 47 of Module A. Data item (A4700). The respondents were asked ‘How would you describe this business’s competition?’ The variable *monopoly* takes the value of 1 if the response is yes, zero otherwise.

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 7 Staff occupation/role variables

Occupation/role	Data item	Variable (divided by total employment)
Managers <i>(i.e. those who supervise staff or determine policy and future direction)</i> ⁵	C1002	<i>prop_man</i>
Professionals <i>(i.e. those who have specific expertise, but no managerial responsibility)</i> ⁶	C1003	<i>prop_prof</i>
Technicians and associate professionals <i>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)</i>	C1101	<i>prop_tech</i>
Tradespersons and related workers <i>Tradespersons and related workers perform tasks requiring trade specific technical knowledge. Include all apprentices and trade supervisors (e.g. electrician, mechanic, hairdresser, baker).</i>	C1201	<i>prop_trade</i>
Clerical, sales and service workers <i>(i.e. those who perform administrative, sales or customer service tasks)</i>	C1302	<i>prop_cleric</i>
Labourers, production, transport or other workers <i>(i.e. those who operate vehicles or equipment or perform manual tasks)</i>	C1203	<i>Baseline</i>

⁵ In Module A, where Managers 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’

⁶ 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)’.

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.

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 – accessions and separations

Labour separation (*SE*) is measured as the annualised number of separations to the firm divided by RME.

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

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

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 (*LP*)

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). This variable is used as a comparison to the multifactor productivity variable measured using the Levinsohn and Petrin (2003) and OLS procedures. The *LP* and multifactor productivity variables (*mfp_lp*, *mfp_ols*), are also shown in the Appendix.