Worker-Firm Heterogeneity and Matching:

An analysis using worker and firm fixed effects estimated from LEED

David C. Maré
Statistics New Zealand and
Motu Economic and Public Policy Research

and

Dean R. Hyslop
Statistics New Zealand and
The Treasury

November 2006

Acknowledgements

This work was undertaken while the authors were on secondment to Statistics New Zealand. The research was funded by the Department of Labour and The Treasury. We thank Sarah Crichton, Walter Davis, Sylvia Dixon, Richard Fabling, Tas Papadopoulos, Steve Stillman and participants at Statistics New Zealand Linked Employer-Employee Data (LEED) Research Forum for discussions and valuable comments. Any views expressed are those of the authors and do not purport to represent those of Statistics New Zealand, the Treasury or Motu Economic and Public Policy Research. Any remaining errors are the sole responsibility of the authors.

The tables in this paper contain information about groups of people so that the confidentiality of individuals is protected. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person or firm. The results are based in part on tax data supplied by Inland Revenue (IRD) to Statistics New Zealand under the Tax Administration Act. These tax data must be used only for statistical purposes, and no individual information is provided back to IRD for administrative or regulatory purposes. Any discussion of data limitations or weaknesses is in the context of using the LEED for statistical purposes, and is not related to the ability of the data to support IRD's core operational requirements. Careful consideration has been given to the privacy, security and confidentiality issues associated with using tax data in this project. A full discussion can be found in the LEED Project Privacy Impact Assessment paper (Statistics New Zealand, 2003).

Abstract

This paper uses Statistics New Zealand's Linked Employer-Employee Data (LEED) over the six year period April 1999–March 2005 to derive and analyse estimates of two-way worker and firm fixed effects components of job earnings rates. The fixed effects estimates reflect the portable earnings premium that each worker receives in whichever firm they work for, and the time-invariant premium that each firm pays to all the workers it employs. Our main estimates use full-time equivalent annual earnings for <u>each</u> job-year observation weighted by its effective employment, which involves about 18.7 million job-year observations for 2.8 million employees and 320,000 firms. Our analysis focuses on three issues. First, how much of the variation in job earnings rates is attributable to observable worker demographic factors (age and sex), unobserved worker effects and unobserved firm effects? We find that worker effects account for about one half, worker demographics one quarter, and firm effects 10–25 percent of the variance in job earnings. Second, how much compositional change occurred during this period of substantial employment growth? As measured by changes in the annual averages, worker and firm effects declined by about 5 and 1 percent, respectively, over the period. Third, what is the aggregate pattern of sorting of workers and firms across jobs? The correlation between worker and firm effects is 0.12, which is higher than international estimates and implies a tendency for high-earning workers to work for high-paying firms. A primary dimension along which sorting occurs is the full-time / part-time employment dimension. The results are qualitatively robust to various sensitivity tests, including unweighted estimation across all jobs, using only workers' main jobs held in each year, jobs of workers estimated to be employed full-time during the year, and excluding jobs in firms that have a low degree of connectivity to other firms. The estimated correlation between worker and firm effects is higher based on unweighted jobs (0.18) and more-connected firms (0.17), but lower based on main job (0.06) and full-time workers (-0.01).

1. Introduction

Workers' earnings rates may vary because of systematic differences across workers and/or differences across the firms they work for.[1] Disentangling the sources of variation in earnings is important for understanding in several areas, including earnings inequality, productivity differences across firms, the impacts of alternative remuneration policies, etc. In addition, if there are complementarities between workers and firms, then we would expect high-earning workers to be concentrated in high-paying firms.[2] In the presence of worker and/or firm components of earnings differences, patterns of worker and firm matching across jobs may affect the degree of earnings inequality (Acemoglu, 1997, Burgess et al, 2004) and/or the relative performance of dense labour markets (Andersson et al, 2005). Assessing the relative importance of alternative sources of variation in job earnings, as well as the strength of the matching, requires the simultaneous estimation of unobserved worker and firm effects from longitudinal job-level data that allow workers and firms to be linked (Abowd and Kramarz, 1999a).[3]

This paper uses linked employer-employee earnings data to address these two sets of issues for the first time in New Zealand, using Statistics New Zealand's Linked Employer-Employee Data (LEED) over the period 1999–2005. The paper has two main objectives. The primary objective is to estimate the regression-adjusted joint (two-way) worker and firm fixed effects associated with the full-time equivalent (FTE) annualised job earnings rate, controlling for worker observed (age and sex) demographic differences. We first document the cross-sectional variation in job-earnings rates over

[1] For example, earnings differences reflect differences in workers' observable characteristics such as education and experience, as well as unobservable characteristics such as innate ability and effort. Similarly, earnings differences may reflect firm productivity differences that arise from differences in entrepreneurial skill, product or industry rents, or labour market frictions such as search and turnover costs. Groshen (1991a) provides an extensive discussion of potential sources of establishment wage differentials. Groshen (1991b), Lane et al (2001) and Davis and Haltiwanger (1991) all document substantial plant-level differences in earnings.

[2] International estimates of the patterns of matching between workers and jobs are mixed. Abowd, Kramarz and Pérez-Duarte (2003) have estimated that the correlation is approximately zero in the US and negative in France, and show that the negative pattern in France is a consequence of strong negative correlation *within* industries and firm-size classes. They characterise this result that 'good workers are employed by bad firms' as a puzzle, and show such an empirical finding can be consistent with a model with positive assortative matching and labour market frictions. See also Abowd, Kramarz, Lengermann and Pérez-Duarte (2004).

[3] In particular, if worker and firm contributions are correlated, standard one-way estimates based on either worker-level or firm-level data will be biased, reflecting the extent to which high-paying firms employ high-earning workers. Specifically, one-way estimates of worker effects will capture the average effects of the firms in which a worker is employed, while one-way firm effects will reflect in part the average person effect of its employees (Abowd, Kramarz and Margolis, 1999).

the sample, and then consider various econometric identification and estimation issues associated with estimating the worker and firm effects of interest.

Our second objective is to describe and analyse the patterns of interaction between these estimated components of the job earnings rate. In this analysis we focus on three broad issues: the degree of worker and firm earnings heterogeneity, compositional changes over the period, and the strength of assortative matching of workers and firms across jobs. First, we document the degree of heterogeneity, as measured by the variability of each component of earnings, both across the full population and also within alternative subgroups defined by worker and firm characteristics. That is, how much earnings variability across jobs reflects differences in workers' observed demographics, versus unobserved systematic worker effects, firm effects and/or idiosyncratic worker-firm job effects?

Second, we describe the effects of compositional change over the period, by focusing on patterns over time and for subgroups characterised by their entry and exit behaviour. Our sample period 1999–2005 was one of strong cyclical economic and employment growth.[4] If the composition of workers varies systematically over the business cycle according to their productivity levels, such compositional change will act to bias downwards measured earnings and productivity growth during business cycle upswings.[5]

Third, we document the degree of assortative matching between workers and firms, again both across the full population and within various subgroups. Our main focus here is to what extent high-earning workers work for high-paying firms. The basic ideas are captured by Becker (1973), who highlights

[4] For example, GDP growth averaged 3.9 percent per annum between 1999 and 2005, the unemployment rate fell from about 7.5 percent to less than 4 percent, the employment rate rose from 61 to 65 percent and the labour force participation rate rose from 65 to 67 percent. Employment in LEED mirrors these changes, with the number of workers employed annually increasing 16 percent over the period.

[5] The basic hypothesis underlying this idea is that more productive workers are employed throughout the business cycle, while less productive workers are marginalised during recessions and drawn into employment during upturns. Solon, Barsky and Parker (1994) show that the usual macroeconomic finding of weakly cyclical wages over the business cycle is substantially affected by composition bias: controlling for compositional changes in employment over the business cycle, real wages are strongly procyclical. There are potentially two distinct factors in the composition effects story: first, a direct effect associated with the changing labour productivity composition of the workforce; and second, an indirect effect associated with changing capital intensity over the business cycle, as labour utilisation changes – i.e. capital dilution holding labour productivity constant.

the central role of complementarity in household production as a basis for positive assortative matching of marriage partners. In the job earnings context, a positive correlation between worker and firm effects can arise if there are sufficient complementarities in production. More recently, Shimer and Smith (2000) generalise Becker's frictionless result by deriving the conditions under which positive matching occurs in a model with search frictions.[6]

The paper is organised as follows. In the next section we provide a brief overview of the LEED and a discussion of the derivation of the variables we use in the analysis. In section 3 we outline the econometric framework adopted, and discuss various identification and estimation issues encountered. We also highlight our main two-way fixed effects results, and compare and contrast these to alternative estimates taking account of the issues discussed. Section 4 contains the presentation and discussion of the results pertaining to the level of heterogeneity, the compositional changes over the period and the matching of workers and firms. The paper concludes with a summary discussion.

2. Data

The analysis in this paper uses Statistics New Zealand's Linked Employer-Employee Data (LEED). The LEED uses information from tax and statistical sources to construct a record of paid jobs.[7] Since April 1999, all employers in New Zealand are required to file a monthly record with Inland Revenue (IRD) called an Employer Monthly Schedule (EMS), which lists all paid employees at that firm during the month, the earnings they received and the amount of tax that was deducted at source. Two types of recipients are covered by EMS: those who have Pay-As-You-Earn (PAYE) tax deducted, who are employees; and those who pay withholding tax, who are a subset of the self-employed. Because the selection and coverage of which self-employed workers have tax withheld is unknown, we use only information on PAYE-deducted (employee) jobs.

[6] The conditions become more restrictive, but continue to require complementarities in production. Whereas Becker's (1973) result requires strict supermodularity of the production function (all agents have higher productivity when they match with high-productivity agents), Shimer and Smith's (2000) result in a search model requires log-supermodularity of the first derivative and cross-derivative of the production function (p. 356).

7 See Statistics New Zealand (2003), Kelly (2003), and Crichton, Stillman and Hyslop (2005) for more detailed discussions of the LEED.

Firms (employers) and workers (employees) are identified by unique confidentialised identifiers based on their respective IRD tax numbers. For workers, this represents a single identifier over time, enabling workers to be tracked longitudinally and across the firms that they work for. In the IRD data, employers are identified as the legal or administrative unit to which the EMS return relates, and do not equate to any consistent conception of a firm. That is, legal and/or other administrative changes can trigger a change in an employer's IRD identifier, with no effective change in the economic structure of the firm. For this reason, we use a version of the LEED that has allocated EMS returns to geographic units, identified by a unique identifier (the Permanent Business Number, PBN) in Statistics New Zealand's Longitudinal Business Frame (LBF) (Seyb, 2003), and adopt such geographic units as our concept of firms.

In addition to regular firm-worker employment jobs being identified in the LEED, several other relationships involving PAYE tax deductions can also be identified by particular "employer" identifiers. These are working-age social welfare taxable benefits;[8] earnings-related accident compensation payments from the Accident Compensation Corporation (ACC); Student Allowance payments (SA); Paid Parental Leave (PPL) payments; and New Zealand Superannuation (NZS) retirement pensions. In what follows, we make a distinction between LEED *earnings* from employment-jobs and other LEED *income* from these other (non-employment) sources.

Conceptually, the LEED covers the universe of PAYE employment relationships and earnings in New Zealand over the period. In addition, there is limited information on the characteristics of workers and firms: age, sex, and location of workers; and industry and location of firms. However, there are some significant weaknesses with the LEED. Perhaps the main weakness of the LEED for the current analysis is that it contains no information on hours worked. The EMS returns report only monthly earnings for each employee. As a result, we cannot accurately distinguish low hourly <u>wage</u> rates from low hourly <u>employment</u> intensity. Similarly, high earnings may result from either a high wage rate or high employment intensity.

[8] The major working-age benefits are the Unemployment, Domestic Purposes, Sickness and Invalids benefits. Although receipt of a taxable working-age benefit is identified, the specific benefit-type is not separately identifiable from the LEED data.

In order to provide a partial adjustment for the lack of hours information, we develop an algorithm to estimate each worker's relative *employment intensity*. This algorithm takes into account both the worker's monthly LEED earnings from employment and any earnings-tested income they receive from other sources; the algorithm also allocates their total employment across their (multiple) jobs. We first assume that each worker can have up to one unit of employment intensity in any month, and their employment is zero in any month that they have no LEED earnings. A worker's total monthly employment intensity is reduced either if their total monthly earnings are less than full-time minimum wage earnings, and/or if they receive any earnings-tested LEED 'non-work payments' income.[9] In the case of low earnings, we estimate an individual's employment intensity as the ratio of their actual to full-time minimum wage monthly earnings;[10] while, in the presence of 'non-work payments', we estimate the employment intensity as the fraction of earnings to total LEED income (i.e. earnings plus non-work payments). Specifically, we estimate individual-*i*'s employment intensity in month-*m*, e_{im}, as

$$e_{im} = \min\left\{1, \frac{earn_{im}}{(earn_{im} + non\_earn_{im})}, \frac{earn_{im}}{FT\_mw\_earn_{im}}\right\} \qquad (1)$$

where $earn_{im}$ is *i*'s total LEED employment earnings in month-*m*, non_earn_{im} is their total (earnings-tested) non-work income in month-*m*, and $FT\_mw\_earn_{im}$ is the full-time minimum wage earnings level applicable to them in month-*m*. As hourly wages generally exceed both minimum wages and non-work income rates, these adjustments likely overstate the employment intensity of part-time workers and those receiving non-work payments relative to full-time workers.

In order to give a sense of the reliability and possible bias in this measure, we have compared the estimated average employment intensity and the fraction estimated to be full-time with analogous estimates using Household Labour Force Survey (HLFS) data for workers over the sample period. The results are discussed in Appendix 1 and summarised in appendix Table A1. In summary, we

[9] That is, if a person receives any (working-age) non-work payments (i.e. working-age benefit, ACC, SA or PPL payments), we infer that they were not-working for at least part of the month. We do not include NZS income as a non-work payment, as its eligibility does not depend on employment status.
[10] We assume 40 hours per week, 4.35 weeks/month, and apply the relevant minimum wage based on age and period – e.g. full-time minimum wage earnings for adults in 2002/03 were $1392 = $8/hour * 40 hours/week * 4.35 weeks/month.

believe the results provide some assurance that, first, the LEED employment intensity construct has similar properties to analogous survey estimates and, second, in the absence of any direct hours measure, it provides a useful first-order adjustment for estimating differing levels of employment intensity across workers.

For workers with multiple jobs, each worker's total monthly employment intensity (or "effective employment") is allocated across the jobs they held in that month in proportion to the earnings from each job, to give their effective monthly employment in those jobs. We define a *job* as a unique firm-worker (i.e. PBN-employee) combination, and a *job-month* as a unique firm-worker-month combination. That is, worker-*i*'s effective employment in firm-*j* (i.e. job-*ij*) in month-*m* is

$$e_{ijm} = \frac{earn_{ijm}}{earn_{im}} * e_{im} \tag{2}$$

where $earn_{ijm}$ is worker-*i*'s LEED earnings from firm-*j* in month-*m*. Aggregating the job-level effective employment of each worker within a firm-*j*, gives the firm's total effective monthly employment

$$e_{jm} = \sum_{i=1}^{N_{jm}} e_{ijm} \ . \tag{3}$$

Also, summing either a worker's, a job's or a firm's monthly effective employment across months in a year provides our estimate of the <u>annual</u> effective employment of the worker, job or firm, which we express in annual terms (i.e. a full-time, full-year worker has annual effective employment of 1, etc.). For example, summing across the 12 months in year-*t*, gives

$$e_{ijt} = \sum_{m=1}^{12} \frac{e_{ijm}}{12} \tag{4}$$

job-*ij*'s annual effective employment in year-*t*. Finally, based on these estimates of worker, job and/or firm annual effective employment, we estimate the corresponding full-time-equivalent (FTE) annual earnings rate as the relevant annual earnings divided by estimated annual effective employment. For example, job-*ij*'s FTE annual earnings rate in year-*t*, is

$$Y_{ijt} = \frac{earn_{ijt}}{e_{ijt}} \tag{5}$$

where $earn_{ijt}$ is job-*ij*'s total LEED earnings in year-*t*.

We use all the available data on PAYE employee jobs in New Zealand during the six March-years from April 1999 to March 2005.[11] Table 1 provides a summary of the annual data: for all years pooled, the first year (1999/2000), the last year (2004/05) and the percentage change between the first and last years. The three panels summarise job-level, worker-level and firm-level data, respectively; the four left columns provide unweighted means, while the four right columns provide analogous means weighted by FTE employment. Over the six-year period, there are 18,676,300 distinct job-year observations and 8,018,300 FTE job-years associated with 9,729,900 jobs, worked by 2,776,400 workers (employees) in 322,700 firms (PBNs). On average over the six year period, workers have 3.5 jobs (different firms), firms employ about 30 different workers, and there are 1.9 annual observations per job. In addition, there are 11,716,400 worker-year observations (on average 4.2 per worker) and 1,211,202 firm-year observations (3.8 per firm). The number of annual job observations, annual job FTE, number of workers and number of firms increased 14, 17, 16 and 10 percent, respectively, between the first and last years.

We focus here on the FTE-weighted statistics. The average job-year FTE employment is 0.78, and 45 percent of job-year observations involve less than full-time employment during the year.[12] The employment weighted average age of workers was 38 years (and increased 1.4 years or 4 percent over the period), and 46 percent were female. There was a large relative (18 percent) decline in the fraction of workers who also received (working-age) non-work payments (largely welfare benefits), from 13 to 11 percent over the period. The employment-weighted average firm annual employment and FTE employment are 300 and 163 workers, respectively.[13] All earnings and incomes have been

[11] Data are available for months following March 2005 but are less complete due to lags in filing EMS returns.
[12] A job-year observation is classed as less than full-time employment here if it is less than full-time in any month the job exists during the year, and does not include less than full-year employment – i.e. jobs that don't exist for the full 12 months.
[13] That is, the average employment (FTE) unit works in a firm that employs 300 workers during the year, with average FTE employment of 163. In contrast, the (unweighted) average firm employs 15 workers during the year, with average FTE employment of 6.6.

adjusted using the Consumers Price Index (CPI) and expressed in constant (December quarter 2005) dollar values. Average annual job earnings are \$36,500, average annual worker earnings are \$40,500 (on average, 91 percent earned from their main, highest-paying, job), average annual earnings/worker in firms is \$25,600, and the average FTE job annual earnings rate is \$44,100.

3. Statistical Model Specification and Estimation Issues

In this section we outline the statistical framework adopted to analyse the contributions to the job-level earnings rate of a limited set of observable worker characteristics, and time-invariant (unobservable) worker and firm effects. Letting *i*, *j* and *t* index workers, firms, and time (year), and defining job-*ij* as the employment relationship between worker-*i* and firm-*j*, the unit of observation for this analysis is a unique (*ijt)* job-year combination. We estimate models of the following form:

$$y_{ijt} = \theta_i + \psi_j + x_{ijt}\beta + \tau_t + \varepsilon_{ijt} \qquad (6)$$

where y_{ijt} is the log(annual FTE earnings rate) of the job-*ij* in year-*t*, which is associated with worker-*i* ($i = 1, \dots N$) employed in firm-*j* ($j = 1, \dots J$); θ_i is the time-invariant effect associated with worker-*i*, which represents their earnings premium across the firms they work for; similarly, ψ_j is the time-invariant effect associated with firm-*j*, which represents the earnings premium it pays to all its workers; x_{ijt} is a vector of observable worker and firm-level characteristics that affect earnings, and β is the associated parameter vector; τ_t are time effects; and ε_{ijt} is a residual that captures idiosyncratic job-match effects, measurement errors, etc.

Our approach to the estimation and analysis is to control for earnings variation across worker age and sex characteristics, and aggregate year effects, and then consider how the alternative components of earnings vary across various subgroups defined by worker and firm characteristics. Specifically, we allow for unrestricted sex-year specific age profiles. Also, we use information on all employment jobs observed in the LEED,[14] and weight each job-year observation in the estimation by its estimated FTE employment. Using annual earnings tends to smooth noisy monthly earnings patterns, and also

[14] This is in contrast to most previous analyses that typically focus either on full-time workers and/or a worker's main job in each year (e.g. see Abowd, Kramarz and Margolis, 1999). We provide some subsequent analysis on various restricted samples to assess the robustness of the results.

lessens the impact of seasonal earnings patterns, both of which are prevalent in the LEED. That full-time jobs and/or those that last longer are likely to provide a less noisy signal of the underlying worker and firm effects also supports using FTE employment weights in equation 6.

Given the large number of person and firm fixed effects parameters in the model, it is not feasible to follow the standard estimation approach of direct least squares estimation, which would involve inversion of a very large sparse covariate matrix. Instead, we use a weighted variant of the *exact* solution for estimation of this model, as described in Abowd, Creecy and Kramarz (2002) (ACK). We adopt their approach of using a preconditioned conjugate-gradient algorithm developed by Dongarra, Duff, Sorensen and Van der Vorst (1991), and implemented in Fortran for this application by ACK.[15]

3.1. Identification of Age, Year and Person Effects

In the presence of worker fixed effects, time and age are perfectly collinear in a balanced panel, so these effects are not identified in equation 6.[16] We have an unbalanced panel, and also measure age on an employment-weighted basis within each year which means age is not perfectly synchronised over time. Nonetheless, we believe the resulting identification associated with estimating equation 6 is tenuous at best. Appendix 2 contains a discussion of preliminary results and problems associated with these from estimating equation 6.

---

[15] Shewchuk (1994) provides an introduction to the conjugate gradient method. ACK's Fortran programs are available for download from http://instruct1.cit.cornell.edu/~jma7/fortran_code.zip. We adapted the programs to allow for weighted estimation.

[16] This is essentially an example of the well-known problem of identification of age, cohort and time effects in relationships of interest (e.g. Hall, Mairesse and Turner, 2005). To see this, consider the simple case with year dummy variables (D_{ts}=1(t=s), a dummy variable for year-s), a linear age trend ($\delta.Age_{it}$) and individual fixed effects (α_i):

$$y_{it} = \alpha_0 + \sum_{s=1}^{T-1} \gamma_s D_{ts} + \delta.Age_{it} + \alpha_i + \varepsilon_{it}, \quad i = 1,...,N; t = 0,...,T-1.$$

Noting that $Age_{it} = Age_{i0} + t$, implies this equation can be expressed as

$$y_{it} = \alpha_0 + \sum_{s=1}^{T-1} (\gamma_s + s.\delta) D_{ts} + \delta.Age_{i0} + \alpha_i + \varepsilon_{it}, \quad i = 1,...,N; t = 0,...,T-1.$$

Applying fixed effects estimation to this equation, the time constant Age_{i0} variable drops out, and the time variation in the year dummy variables is required to identify both γ_s and δ, which is not possible.

Because of this identification problem, we adopt a two-step procedure where, in the first stage, we estimate unrestricted sex-age earnings profiles for each year and, in the second stage, use the residuals from this exercise to estimate the (unobserved) worker and firm effects. That is, in the first stage, we regress job FTE annual earnings rate on a full set of worker sex-age dummies, allowing the coefficients to vary by year. The regression-adjusted earnings are then projected onto full sets of worker, firm and time dummy variables. The estimating equations are:

$$\begin{aligned} y_{ijt} &= \beta_{gAt} + \varepsilon_{ijt} \\ \hat{\varepsilon}_{ijt} &= \theta_i + \psi_j + \tau_t + u_{ijt} \end{aligned} \tag{7}$$

where β_{gAt} is a vector of coefficients on a full set of sex * age * year dummy variables, and $\hat{\varepsilon}_{ijt}$ is the residual from the first-stage regression. This approach identifies the combined second stage dependent variable (i.e. the combined worker, firm, and idiosyncratic job-year effects) as orthogonal to the unrestricted sex-year age profiles estimated in the first-stage. The year dummy variables in the second stage regression (τ_t) are used to control for compositional changes over the period.

3.2. Grouping and the Uniqueness of Person and Firm Fixed Effects

The first step for estimation and identification is to allocate job-year observations into distinct 'connected' groups of firms and workers. ACK (p. 3) summarise the essence of this connectedness:

> "Connecting persons and firms requires that some of the individuals in the sample be employed in multiple employers. When a group of persons and firms is connected, the group contains all the workers who ever worked for any of the firms in the group and all the firms at which any of the workers were ever employed. In contrast, when a group of persons and firms is not connected to a second group, no firm in the first group has ever employed a person in the second group, nor has any person in the first group ever been employed by a firm in the second group."

ACK consider the identification of worker and firm effects that arise with the simultaneous estimation of worker and firm fixed effects models such as that shown in the second line of Equation 7. As with standard fixed effects models, restrictions are required in order to identify the relevant effects of interest. First, suppose there are G distinct non-overlapping groups of connected workers and firms. Within a group g containing N_g persons and J_g firms, it is possible to identify the group mean, $N_g - 1$ worker effects and $J_g - 1$ firm effects, yielding $N_g + J_g - 1$ identified effects. Across all G groups, there are $N + J - G$ estimable effects.

Second, the estimated effects are not unique, and an explicit identification procedure must be imposed. The non-uniqueness of estimates arises because, within each group, it is arbitrary which effect is omitted – the group mean, one of the worker effects or one of the firm effects. To obtain unique estimates, we restrict the overall mean firm effect to be zero, and the mean worker effect within each group to be zero. Given these restrictions, we can identify the overall mean of the dependent variable, and $N + J - G - 1$ worker and firm fixed effects.

We apply the ACK grouping algorithm to data on all *job-years* observed during the six years of our data. The results of this grouping are shown in Table 2, together with summaries of Abowd et al's (2004) separate findings for France, the US State of Washington and seven US states. This algorithm generates a set of 12,200 non-overlapping groups of workers and firms, and the largest group contains the vast majority (99.7%) of job-year observations. The proportion in the largest group is slightly higher than in the French or US data, possibly partly due to the use of information on all jobs rather than main job only. To see this, we have also applied the grouping algorithm on the subset of each worker's highest earnings 'main job' in each year, and the results are presented in the second panel of Table 2. There are two points of interest to note from this panel. First, as well as losing approximately one-third of job-year observations associated with multiple job holdings, selecting workers' main jobs also eliminates some 11 percent of firms from the sample. Second, the degree of connectedness drops: e.g. the fractions of observations, workers and firms in the largest group are all lower (albeit still accounting for the vast majority of total observations), while the total number of groups also roughly doubles. Thus, the inclusion of a broader set of jobs serves to increase the degree of connectedness between workers and firms.[17]

[17] Over the entire sample, 45,655 job-year observations accounting for 29,421 FTE-years are for single-firm groups. Single-firm groups account for around 0.24 percent of observations and 0.37 percent of FTE. There are 11,461 firms (3.6% of firms) and 14,512 workers (0.5% of workers) in single-firm groups.

3.3. Results – Contributions to Earnings Variation across Jobs

Based on the grouping and the two-stage estimation approach described above, Table 3 contains a summary of the main estimation results from FTE-weighted regressions of job-year observations. In the first panel, we present the means, and standard deviations in parentheses, of the raw job-year log(FTE earnings), together with the estimated contributions of the first-stage observables (i.e. the sex, age, year effects), the second-stage worker and firm effects, and the residual earnings component. The first row contains the main FTE-employment weighted estimation results of job-year FTE earnings. The logarithm of the FTE annual job earnings rate has a mean of 10.54 (equivalent to \$37,800 per year), with a standard deviation of 0.34.[18]

By construction, the first stage residuals have zero mean. The earnings component associated with the "first stage covariates" has a standard deviation of 0.17, while the worker and firm effects have standard deviations of 0.24 and 0.10, respectively. The variability in each of these components reflects the degree of heterogeneity across workers and firms, as measured across intensity-weighted job-years.[19] These results suggest there is greater systematic variability in job earnings within observable age-sex worker demographic subgroups than there is across the groups – i.e. the standard deviation of worker effects (0.24) exceeds the variability of earnings accounted for by the sex-age profiles (0.17).

Overall, the model accounts for 90.3 percent of the variation in FTE annual job earnings rates. To gauge the importance of the various components to job earnings we measure how much each contributes to the total variation in job earnings. For each component, we have calculated both the "simple R^2" from the regression of log(job earnings rate) on that component and also the "marginal R^2", which is the increase in R-squared associated with adding that component to the model that already includes all other components. The simple R^2s associated with the observable variation across

[18] The standard deviations are calculated as $\Sigma w_i(y_i - E[y_i])^2 /(N-1)$, and can thus be interpreted as the standard deviation across jobs with $w_i = 1$ (i.e. full-time full-year jobs).

[19] Alternatively, if we give equal weight to each person observed in the sample, the mean worker effect is -0.11 with a variance of 0.15 (sd=0.38). Similarly, the variance of firm fixed effects weighted equally across all observed firms has a mean of -0.10 and variance of 0.08 (sd=0.29). These negative unweighted means indicate that workers and firms with higher employment intensity over the period tend to have higher estimated effects. Persons and firms that receive higher weight also tend to have less extreme estimated effects, partly due to the accentuated variability associated with smaller firms with low connectedness, as described in section 3.5.

sex, age and/or years, worker effects and firm effects, respectively, are 0.26, 0.49 and 0.25, while the marginal R^2s of each of these is 0.24, 0.49 and 0.09. Based on these estimates, the worker effects make the largest contribution, followed by the worker demographics. Given that we do not observe education levels (and other common worker characteristics) in our data, the worker effect component absorbs the impact of human capital variation that is included as part of observable interpersonal variation in some other studies.[20] Also, the lower contribution of the firm effects reflects, in part, the greater clustering of jobs across firms, with an average of 58 job-year observations per firm, compared with only 7 job-year observations per person. As we will see subsequently, the similarity of the simple and marginal R^2s for the worker effects and observable effects, and the substantial drop in marginal versus simple R^2 for the firm effects, is explained by the finding that these two effects are negatively correlated, while each is positively correlated with the firm effects.

The second row of Table 3 contains analogous results based on unweighted estimation – i.e. treating each job with equal weight.[21] The average FTE annual earnings rate of job-years is about 25 percent lower,[22] and the variability in earnings higher, than the weighted estimates, implying part-time and/or part-year jobs have lower and more variable earnings rates than full-time full-year jobs. The overall explanatory power of the unweighted specification is lower (R^2 = 0.84), however the relative sizes of person and firm contributions are similar to the weighted estimation.

3.4. Results – Correlations across Earnings Components

In the second panel of Table 3 we present the estimated correlations between the various components described in panel A across job-year observations,[23] again both FTE employment weighted and unweighted. Perhaps the main result of interest in this panel is the correlation between the estimated

[20] Strictly speaking, the worker effect absorbs only the impacts of such factors that are orthogonal to the sex-age profiles estimated in the first stage. That the simple and marginal R^2s for the worker effects are the same, implies such factors are largely orthogonal to the observable covariates.

[21] The unweighted estimates are derived by giving equal weight to each job-year observation, regardless of employment intensity. Both the first and second-stage regressions are re-estimated to produce these figures.

[22] Here and subsequently, we interpret log-point differences as percentages. This approximation is very close for small differences, but deteriorates for larger differences – e.g. 0.01, 0.1, 0.20 and 0.3 log differences correspond to 1.01, 10.5, 22.1 and 35.0 percent differences.

[23] That is, the log(FTE Earnings) of each job-year observation is decomposed into its stage-1 observable effects, and stage-2 time, worker and firm effects, and residual. The correlations between these various components are presented in the table.

worker and firm effects. The correlation between the weighted estimates is 0.12 and between the unweighted estimates is 0.18. These correlations are greater than the 0.08 estimated by Abowd, Kramarz and Pérez-Duarte (2003) for seven US states, and -0.03 and -0.28 for Washington state and France, respectively, estimated by ACK, suggesting possibly greater positive assortative matching of workers and firms in New Zealand.

In addition, the table shows that the worker effects are negatively correlated with the observed sex-age profiles, suggesting that (lower earning) females and non-prime-aged workers, on average, have positive effects, and males and prime-aged workers have negative effects.[24] Counterbalancing this effect, there is a positive correlation between firm effects and worker demographics, which suggests that higher-paying firms predominantly employ males and prime-aged workers. We return to these and related issues in section 4 when we discuss subgroup patterns.

3.5. Strength of Identification

Although a single worker leaving a firm is sufficient to identify the firm fixed effect, such identification is tenuous and has the effect of introducing a negative correlation between worker and firm fixed effects. To understand this, consider the case of two firms with identical earnings distributions, and suppose a relatively low-paid worker leaves firm A and secures a highly paid position in firm B. This change would identify a difference between firm A and firm B fixed effects (ψ_B - ψ_A) equal to the earnings gain of the worker. In addition, the remaining workers in firm A would appear to have relatively high person fixed effects (θ_i), and the workers in firm B would have relatively low person fixed effects. The greater the number of 'contrasts' (i.e. firm changes for workers within a connected group) used to identify the firm fixed effects, the smaller will be the impact of such induced negative correlation.

To investigate the possible impact of such potentially tenuous identification, we first measure the degree of each firm's connectedness, and then consider the relationship between this level of

[24] The raw age-log(earnings) profiles for males and females shown in appendix Figure A1(a) provide an accurate picture of the relative earnings of workers of different age and sex.

connectedness and the correlation between the estimated worker and firm effects.[25] In particular, let J_i be the number of distinct firms that worker-*i* has worked in. Our index of connectedness (χ_j) for a given firm is the number of workers who have been employed both in firm-*j* and in at least one other firm. That is,

$$\chi_j = \sum_{i \in j} 1(J_i > 1) \tag{8}$$

where the summation is over all workers ever employed in firm-*j*. We then estimate the correlation between the estimated worker and firm effects for jobs within firms with differing levels of connectedness. In Figure 1 we graph the estimated correlation of worker and firm fixed effects, stratified by the index of connectedness (χ_j). Where the number of workers with links to other firms is low, there is a negative or low correlation between worker and firm effects, consistent with tenuous identification.[26] The figure also shows the cumulative proportion of FTE employment accounted for by jobs where connectedness is low. Although not shown in this figure, the estimated firm and worker effects associated with low-connected firms also tend to be relatively extreme values. Around 20 percent of FTE employment is in firms where the number of 'outside links' is fewer than 25.

In Table 3 we have also summarised the results when we trim firms with low connectedness (χ_j<25).[27] The results in panel A show that, on average, job FTE earnings of this subsample are about 3 percent higher than in the full sample, which is attributed to about 1 percent higher worker effects and 2 percent higher firm effects. As expected, the resulting correlation between worker and firm effects (0.17), reported in panel B, is about 50 percent higher for this sample.

3.6. Comparison with One-way Fixed Effects Estimates

Abowd and Kramarz (1999) clearly document the biases that can arise in the presence of correlated firm and worker effects, when estimation is based on one-way fixed effects only. The degree of bias

[25] See also Abowd, Kramarz, Lengermann and Pérez-Duarte (2004) for a discussion of this issue.
[26] Although we interpret the patterns in Figure 1as a consequence of tenuous identification arising from low connectedness, it may be that the negative correlations are genuine. Also, low connectedness is most likely among small firms, and a negative correlation of worker and firm effects for such firms would generate the observed patterns.
[27] For comparison, we present the full set of estimates based on this subsample in appendix Tables A2 and A3. The results, in general, are similar.

is, however, an empirical question. In our data, we find a positive correlation between worker and firm fixed effects, which implies that one-way estimates will lead to an overstatement of the degree of worker and firm heterogeneity, and of the strength of positive assortative matching.

To illustrate this, we use the residuals from the first-stage regression in equation 7 to estimate, sequentially, first one-way worker then one-way firm fixed effects and, alternatively, first one-way firm then worker fixed effects. When we estimate the worker (firm) effects first, we refer to these as the "one-way" worker (firm) effects, and when we estimate the worker (firm) effects second, we refer to these as the "order dependent" worker (firm) effects. Table 4 presents the correlations between these various estimates and also the two-way estimates described above. The respective estimated worker effects are quite highly correlated: 0.81 between the one-way and order-dependent estimates, 0.90 between the order-dependent and two-way estimates, and 0.94 between the one-way and two-way estimates. The corresponding firm effects are less highly correlated, ranging from 0.50 between the one-way and order-dependent estimates to 0.76 between the one-way and two-way estimates.

Compared with the correlation of 0.12 between the jointly estimated two-way worker and firm fixed effects estimates, the correlation between the one-way worker and firm effects is substantially higher (0.60), while the correlation between the order-dependent person and firm effects is almost zero (0.01). These results derive from the fact that the one-way person fixed effects estimates reflect the mean firm effect for the firms in which a person has worked, with a similar effect for one-way firm effects, while the order-dependent estimates have each had these common components stripped out.

Alternatively, because the one-way person effects capture the mean firm effect of the firms in which a worker has been employed (similarly, firm effects capture mean worker effects), and the positive correlation between the two-way effects, the one-way estimates leads to an overstatement of worker (and firm) heterogeneity. That is, the standard deviation of person fixed effects is overstated by 12 percent (0.27/0.24), and of firm fixed effects by 66 percent (0.17/0.10). The implied correlation between worker and firm effects is also exaggerated by the one-way estimates – by a factor of 5 (0.60/0.12). Also, as expected, the standard deviation of the order-dependent estimates is lower, particularly for firm fixed effects.

4. Analyses of Heterogeneity, Compositional Change and Matching

We now turn our attention to interpreting the patterns across various worker and firm-level dimensions. For this analysis, we rely exclusively on two-way fixed effects estimates. We focus on three separate but related issues: first, describing the extent of heterogeneity in job earnings rates across different dimensions; second, compositional effects associated with the strong increase in employment observed over the sample period; and third, the degree and variation in matching between workers and firms across different dimensions.

4.1. Heterogeneity across Worker and Firm Subgroups

With minor variations, a similar degree of heterogeneity is evident for both workers and firms and across different subgroups. Differences in observed earnings may reflect differences in worker and/or firm factors. We begin by describing the components of earnings variation of the observable and fixed-effects factors for various subgroups defined by alternative worker and firm characteristics. This description of the results across various worker and firm subgroups is summarised in Table 5. We first describe the patterns across worker sex and age subgroups. The average job earnings rate of males is about 27 percent higher than for females. Conditional on the respective estimated sex-age profiles, on average, males work in jobs with 2 percent higher firm effects than average. A consequence of this and of the identification restrictions,[28] females work in jobs with 2 percent lower firm effects, and the average worker effects of males and females are -2 and +2 percent, respectively. The average job earnings rate across age groups describes a concave age earnings profile, with younger workers earning substantially less than, and older workers also earning less than, prime-aged workers. In addition, both young and old workers, on average, have jobs with low firm effects (7 percent lower than average for those aged under 20 years, and 2 percent lower for those aged 60–69 years).

The principal firm characteristic that we observe in the LEED is the industry the firm operates in, and we describe the pattern of results across 1-digit industries. Unsurprisingly, there is substantial cross-industry variation in job earnings rates. For example, the average industry-level log(FTE annual job-

[28] Literally, the symmetry is exact if we have equally balanced male and female subsamples.

earnings rate) varies 0.78 across industries from a low of 10.16 ($25,800) in the Accommodation, Cafes and Restaurants industry to a high of 10.94 ($56,400) in the Electricity, Gas and Water Supply industry. Furthermore, conditional on observable worker demographics, the average worker effect varies from a low of about 13 percent below the overall average in Agriculture, Forestry and Fishing to a high of 14 percent above the average in Finance and Insurance, while the average firm effect ranges from a low of 14 percent below the overall average in Accommodation, Cafes and Restaurants to a high of 19 percent above the average in Mining.

As well as across industry variations in these earnings components, there is also substantial variation in the degree of variability within industries. For example, the standard deviation of firm effects ranges from a low of 0.05 in Accommodation, Cafes and Restaurants, in Government Administration and Defence, and in Education to a high of 0.15 in Mining. This suggests that, as well as there being relatively high-paying firms in Mining, there is a larger degree of firm heterogeneity in Mining compared with other industries. Similarly, as well as Finance and Insurance having the highest worker effects, on average, this industry also has the highest variability in worker effects (sd=0.33), suggesting both high-earning workers and also substantial heterogeneity. In contrast, Agriculture, Forestry and Fishing has both the lowest average and variability in worker effects.

Figure 2 describes the patterns across industries of raw earnings, together with the differentials attributed to worker demographics, worker effects and firm effects. Industries are ordered from left to right along the horizontal axis in terms of increasing average log(job earnings rates). First, this figure shows that workers in industries with high raw earnings also tend to have demographics associated with higher than average earnings, although the cross-industry demographic relationship is weaker than the raw earnings pattern. Second, high earnings industries tend to have higher average worker effects and higher positive firm effects. That is, although there is quite a lot of variation, the raw industry earnings differentials generally reflect all three of the demographic, worker and firm effects.

The next subgroup dimension we consider in Table 5 is based on the geographic location of the firms. Consistent with other research (e.g. Lewis and Stillman, 2005) this shows that jobs located in Auckland and Wellington have 8–10 percent higher earnings rates than the overall average. The first-stage estimates of the worker demographic effects on earnings are almost the same across geographic subgroups, and our estimates attribute the differences almost entirely to (unobserved) worker and firm effects. In Auckland, both the average worker and the average firm effects are about 4 percent higher than their respective overall averages. In Wellington, the average worker and firm effect differences are 6 percent and 3 percent, respectively. Job earnings in Auckland and Wellington are also more variable than overall, and this higher variability is attributed to greater variability in worker effects than elsewhere.

The final subgroup description we present in Table 5 is of the stability of employment of workers and firms. For this purpose, in each year, we stratify workers by whether they worked full year and/or full-time in every month they worked; and somewhat analogously, we stratify firms by whether the annual employment in a firm consisted predominantly of full-time and/or full-year workers.[29]

We first focus on the characterisation of workers' employment stability. It appears that the part-time characterisation is the dimension along which job earnings rates primarily differ. Part-time workers earn on the order of 40 percent (strictly, 40 log-points) lower job earnings than the overall worker average. The lower earnings are due to 9–12 percent lower earnings associated with observable demographics, 18–19 percent lower worker effects, and 5–6 percent lower firm effects. However, given that our measurement of effective employment is biased upwards (and FTE job earnings biased downwards) for part-time workers, these results should be interpreted with some caution.

A similar, though more muted, pattern applies to firms' employment stability. For example, the job earnings in firms with predominantly "part-time" employment are 4–10 percent lower than the overall average, which is due to 0–3 percent lower earnings associated with worker demographics, 0–3 percent lower worker effects, and 4 percent lower firm effects. The latter suggests that firms that use a relatively large fraction of part-time employment pay lower earnings rates than other firms (who employ the same workers). These are probably the two most important between-group differences.

[29] More specifically, we classify a firm's employment in a year as "full-time" if the number of observed worker-months in LEED is at least 75 percent of the potential number of months given the number of workers employed by the firm during the year, and we classify the firm as "full-time" if the level of FTE employment in the firm is at least 95 percent of the number of worker-months. See Hyslop and Maré (2006) for a more detailed description and discussion of these employment stability measures.

4.2. Employment Composition Effects over the Period

As described in the data section above, there was strong growth both in the number of workers and effective annual employment over the sample period (16–17 percent between 1999/2000 and 2004/05). In addition, the number of active firms increased by nearly 10 percent over the period. Against this backdrop, we describe the compositional changes associated with the workforce and firms over the sample period, summarised in Table 6. The first row of Table 6 reports the means and variability of various components of earnings variation for the full sample of job-years, repeated from Table 3. The next panel describes the annual composition effects for each year over the period. The first two columns show that the average FTE annual earnings rate increased about 6 percent (0.06 log-points) over the period, and this increase is attributed to the sex-age profiles and aggregate time dummies in the first-stage regression. Most of this change appears to be due to aggregate time effects. For example, allowing unrestricted age earnings profiles by age and across males and females, but restricting these to be parallel in different years, we estimate the aggregate time effect to be 6 percent (0.06 log-points) between the first and last years.[30]

The estimates of the second-stage components in the subsequent columns characterise the composition changes over the period, conditional on the observed year-specific sex-age profiles. Although the first-stage regression ensures that the mean of the dependent variable in the second stage (first-stage residual, ε_{ijt}) is zero for each year, the composition of workers and firms is changing, as reflected in their (average) estimated effects, and the second-stage time effects will balance these changes out. Compared with the average estimated worker effect for all workers observed during the 6-year sample period, the averages based on the samples of workers in each year declines from 2 percent higher for workers in the first year (1999/2000) to 3 percent lower for workers in the last year (2004/05). There is a smaller decline in the average of firm effects by year from about 1 percent above average in 1999/2000 to marginally less than average in 2004/05. As discussed, the second-stage estimated time effects act to balance out the impact of these declining average worker and firm effects over the sample, and show an increase from -3 percent in 1999/2000 to 3 percent in 2004/05.

[30] Relaxing these specification to allow separate sex-year dummy variables, we estimate male aggregate (time) earnings growth of 4.4 percent, and female earnings growth of 7.9 percent.

The pattern of changing time effects reflects the dynamics of worker and firm inflows and outflows over the period. To shed further light on this issue, we first characterise both firms and workers by their respective entry and/or exit patterns over the sample period, and then compare the effects across mutually exclusive subgroups.[31] For this purpose, we have stratified workers (firms) into the following four groups: "continuers" who are observed working (employing workers) in each of the 6 sample years; "entrants" who are not observed in the first year but are observed in the last year, and make a single transition into work (employing) over the sample; symmetrically, "exiters" who are observed in the first year but not in the last year, and make a single transition out of work (employing) over the sample; and "other" workers (firms) who work (employ) intermittently over the six years. The results from this exercise are presented in the third panel of Table 6.

A rough characterisation of these groups is as follows. First, continuing firms account for the large majority (80 percent) of the job-year observations over the sample, while continuing workers account for about 60 percent. The average job earnings associated with each of these groups is slightly higher than the overall average, and each of the other groups have lower average job earnings. Single-exit firms and workers have lower average observable covariate effects than the continuers, but have similar unobservable 'effects'. The single-entrant and "other" groups also have lower observable effects than continuers, and also have much lower worker (about -6 percent for firms and -10 percent for workers) and firm (about -4 percent for firms and -3 percent for workers) unobserved effects.

A more detailed description shows that the average job earnings rate in the continuing firm sample is 2 percent higher than the full sample average, and this difference is attributed equally to observable worker demographics and (unobserved) worker effects, with the average firm effects of continuing firms being the same as the full sample. The job earnings rates of the three other subgroups of firms are lower than the overall average. The earnings rates in single-entrant firms are 9 percent lower: 1 percent lower due to observable worker demographics, 5 percent lower due to each of worker and firm effects, and 2 percent higher due to time effects (i.e. entering firms, on average, appear later in the period). Similarly, the job earnings in exiting firms are about 6 percent lower than the overall

[31] This is analogous to comparing productivity patterns across "continuing", "entering" and "exiting" firms – e.g. see Law and McLellan (2005).

average: 2 percent lower due to worker observables, 2 percent lower worker and 1 percent higher firm effects, and 2 percent lower time effects (again reflecting that exiting firms appear earlier in the period). For all other firms that appear during the sample, the average job earnings rate is 14 percent lower than the overall job average: 3 percent lower due to observable worker demographics, 7 percent lower average worker effects, and 4 percent lower average firm effects.

A similar description applies to the subgroups of workers. First, continuing workers have, on average, 5 percent higher job earnings rates than the overall sample: 3 percent higher due to observable demographics, 2 percent higher due to unobservable worker effects, and 1 percent higher attributed to the firms they work for. Entering workers have substantially (22 log-points or, roughly, 22 percent) lower average job earnings than the overall sample average: 9 percent lower due to observable demographics (in particular, entering workers are, on average, younger), 10 percent lower worker effects, 4 percent lower due to the firms they work for, and 2 percent higher due to the time effects. Exiting workers' average job earnings are only 2 percent lower than the overall sample average: 3 percent lower due to demographics, 2 percent higher worker effects, and 2 percent lower due to time effects. Finally, other workers (with intermittent worker patterns over the period) also have substantially lower job earnings rates (about 18 percent lower than the overall sample average): 5 percent lower due to demographics, 11 percent lower worker effects, and 3 percent lower due to their employing firms.

These results are consistent with a simple hypothesis that, during a business cycle upswing, the composition of the workforce changes as it expands, and lower productivity workers and firms are drawn into employment. This hypothesis potentially explains both the time pattern of worker effects for workers observed in different years shown in the second panel and also the pattern of effects across the "continuing", "entry" and "other" subgroups of workers shown in the third panel. However, there are some caveats around this issue, including possible bias associated with the estimated effective employment measure of workers who receive non-employment income (e.g. benefit receipt) and/or those who work part time. If the groups of "entrants" or "other" workers have a greater incidence of part-time work and/or receipt of non-employment income, then their comparative earnings rates are likely to be downwards biased (e.g. “other” workers likely have lower attachment to the labour market, as demonstrated by their patterns of intermittent employment). We intend to return to this topic in more detail in a subsequent paper, and defer attention to these issues until then.

4.3. Matching of Workers and Firms

Our final focus in this section is on the degree of, and variation in, matching between workers and firms across different dimensions. For this purpose, in Table 7, we present the correlations between the various estimated components of job earnings for the full sample of job-year observations in the first row (repeated from Table 3), and also for the stratified subsamples of jobs discussed above and described in Tables 5 and 6.

In this analysis we focus primarily on the correlation between the estimated worker and firm effects of job earnings conditional on the observable worker demographics. As discussed above, the full sample correlation between worker and firm effects is 0.12, suggesting there is positive sorting of workers across jobs along this dimension. In considering the subsequent subgroup correlations it is important to realise that these are “within-group” correlations (i.e. the correlations are calculated relative to the subgroup worker and firm effect means), and exclude any between-group correlation effect that is included in the overall correlation estimate.

The estimated correlations for male and female workers are similar to the overall correlation. Over the age profile, the correlation is strong for prime-age workers (e.g. for 30–39 year olds, the correlation is 0.19), and weaker for young and old workers (correlations or 0.04 and 0.05 for workers aged under 20 years and 60–69 years, respectively). For young workers, the low correlation between worker and firm effects may be due to early labour market “job shopping”, which results in more “random” matching between workers and firms, and also associated with relatively more higher-ability workers spending time in non-career jobs (e.g. students) than later in life. For older workers, the lower correlation may be due to greater non-wage compensation associated with older cohorts and/or reflect that the earnings rate is a less important measure of the attractiveness of a job towards the later in the working life. Alternatively, it may be that matching contributes to the shape of the age-earnings profile – i.e. some individuals' peak earnings during their life cycle may be the result of better job matches.

There is substantial variation in the worker and firm effect correlations across industries. For example, the correlation is negative (about -0.05) in Agriculture, Forestry and Fishing, and in Construction, and quite low in several other industries, while there is a strong correlation (0.29) in Communication Services, and also quite high correlation in some other industries.

In terms of possible compositional changes over the sample period, the correlation between worker and firm effects across jobs is roughly constant over the period, although falls somewhat over the final two years. The correlation is lower for the "single-entrant" (0.05) and "other" (0.04) firm subgroups, and higher for the subgroup of "single-exit" (0.17) workers. Across regions, the worker-firm effect correlation is higher in Auckland (0.13) and Wellington (0.20) and relatively low in Christchurch (0.05) and "Other" (0.04) areas.

Finally, across the subgroups characterised by "employment stability", the correlation between worker and firm effects is negative for the subgroups of part-time workers (-0.01 and -0.04 for part-year and full-year subgroups, respectively). It is also relatively low for the subgroups of full-time workers (0.04 and 0.02 for the part-year and full-year subgroups). Thus, it appears, that much of the overall correlation between worker and firm effects is associated with the full-time / part-time dimension of employment stability, and that there's little evidence of matching within these groups. The correlation for the subgroups of "full-year" firms are also relatively low (0.05 for the "part-time" and 0.00 for the "full-time" firms).

5. Concluding Discussion

In this paper we have documented the joint estimation of worker and firm effects, together with observable worker demographic effects, associated with the FTE annual earnings rate of jobs in Statistics New Zealand's LEED. The analysis of these factors has focused on three broad themes. First, we examine how much of the variation in job earnings rates is attributable to observable worker demographic factors (age and sex), unobserved worker effects and unobserved firm effects, and document the degree of heterogeneity in each factor across various dimensions. We find that the unobserved worker fixed effects account for about one-half of the variance in job earnings, while worker demographics account for one-quarter, and the firm fixed effects account for 10–25 percent of the variance.

Second, we explore the compositional changes in workers and firms over the period. Based on changes in the annual average of worker and firm effects over the period associated with such compositional changes, we estimate that worker and firm effects declined about 5 and 1 percent, respectively, over the period. The declining average worker effect is consistent with the hypothesis that there are compositional changes in employment and the labour force over the business cycle that lowers average worker productivity during booms compared to recessions, and suggests there has been about a 1 percent annual decline over the sample period. Our results also show that males and prime-age workers, on average, work for higher-paying firms than females and young or old workers, which is due to the changing composition of employment over the period. We intend to return to this topic with a more extensive analysis in a subsequent research project.

Third, we document patterns of sorting of workers and firms across jobs over the sample period. The correlation between worker and firm effects is 0.12, which implies there is a tendency for high-earning workers to work for high-paying firms, suggesting positive complementarities between workers and firms. The 0.12 correlation between worker and firm effects is relatively higher than international estimates, and suggests that relying on either one-way worker or firm effects estimation will be more problematic in the New Zealand context. We have also found quite strong between-group sorting along the full-time / part-time dimension of employment.

We have examined the robustness of these results in a variety of subsample and estimation sensitivity tests. These sensitivity tests include unweighted estimation across all job-year observations, and based on subsamples using only workers' main jobs held in each year, using only jobs of workers estimated to be employed full-time during the year, and using only jobs in firms that have a reasonable degree of connectivity to other firms. Broadly speaking, the results are qualitatively robust across these sensitivity tests. However, the estimated correlation between worker and firm effects varies significantly across the samples considered: it is higher based on unweighted jobs (0.18) or more connected firms (0.17), but lower based on main job (0.06) or full-time worker (-0.01) subsamples. Further analysis is required to understand the importance of these dimensions.

After flexibly controlling for earnings variation across the life-cycle, we find evidence of worker sorting across firms along several dimensions. For example, sorting appears to be stronger in Auckland and Wellington than other regions. Perhaps unsurprisingly, there is also substantial variation across industries. We find generally positive associations between industries' job earnings rates and each of the average earnings based on worker demographics, average worker effects and average firm effects, although there is also heterogeneity between different industries.

References

Abowd, John M., Robert H. Creecy, and Francis Kramarz (2002). "Computing person and firm effects using linked longitudinal employer- employee data", *LEHD Technical Paper* 2002-06, available from www.lehd.census.gov.

Abowd, John M., and Francis Kramarz (1999a). "The Analysis of Labor Markets Using Matched Employer-Employee Data", in Orley Ashenfelter and David Card (eds) *Handbook of Labor Economics*, Vol. 3B, Chapter 40.

Abowd, John M., Francis Kramarz, and David N. Margolis (1999). "High Wage Workers and High Wage Firms", *Econometrica*, Vol. 67, No. 2, pp. 251-333.

Abowd, John M., Francis Kramarz, and Sébastien Pérez-Duarte (2003). "Sorting Workers Between and Within Industries" mimeo, February 2003, available at http://cep.lse.ac.uk/seminarpapers/28-02-03-KRA1.pdf.

Abowd, John M., Francis Kramarz, Paul Lengermann, and Sébastien Pérez-Duarte (2004). "Are Good Workers Employed by Good Firms? A test of a simple assortative matching model for France and the United States" mimeo, January 2004, available from http://instruct1.cit.cornell.edu/~jma7/aklp01v12.pdf.

Acemoglu, Daron (1997). "Matching, heterogeneity and the evolution of income distribution", *Journal of Economic Growth* 2, pp. 61-92.

Andersson, Fredrick, Simon Burgess, and Julia I Lane (2005). "Cities, Matching and the Productivity gains of agglomeration" mimeo (http://www.ecn.bris.ac.uk/www/ecsb/papers/cities.pdf; previous (2004) version available as CEP Discussion paper 6048 http://cep.lse.ac.uk/pubs/download/dp0648.pdf).

Becker, Gary S (1973). "A theory of marriage: Part I", *Journal of Political Economy* 81, pp. 813-46.

Burgess, Simon, Julia Lane, and Kevin L McKinney (2004). "Matching, reallocation and changes in earnings dispersion".

Crichton, Sarah, Steve Stillman, and Dean Hyslop (2005). "Returning to Work from Injury: Longitudinal Evidence on Employment and Earnings", mimeo, Statistics New Zealand, available at http://www.stats.govt.nz/NR/rdonlyres/A76660DF-92D9-47FE-828C-65B2BE9EB763/0/LEEDReturntoworkfromInjuryV20805.pdf#search=%22returning%20to%20work%20from%20injury%20LEED%20%22.

Davis, Steve J. & Haltiwanger, John (1991). "Wage Dispersion Between and Within U.S. Manufacturing Plants", *Brookings Papers on Economic Activity* pp. 115-200.

Dongarra, Jack J., Iain S. Duff, Danny C. Sorensen and Henk A. Van der Vorst (1991). *Solving Linear Systems on Vector and Shared Memory Computers*. SIAM, Philadelphia.

Groshen, Erica (1991a). "Five Reasons Why Wages Vary Among Employers", *Industrial Relations*, pp. 350-81.

Groshen, Erica (1991b). "Sources of Intra-Industry Wage Dispersion: How Much do Employers Matter?", *Quarterly Journal of Economics*, pp. 869-84.

Hall, Bronwyn H., Jacques Mairesse and Laure Turner (2005). "Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists", mimeo.

Hyslop, Dean R., and David C. Maré (2006). "Patterns of Labour Market Fragmentation, and Worker and Firm Employment and Earnings Change, 1999-2005", mimeo, Statistics New Zealand.

Kelly, Neil (2003). "Prototype Outputs Using Linked Employer-Employee Data", Statistics New Zealand report, available at http://www.stats.govt.nz/domino/external/pasfull/pasfull.nsf/e17092639373f80b4c2567ed0009410e/4c2567ef00247c6acc256dbf0015303f?OpenDocument.

Lane, Julia I, Salmon, Laurie A. & Spletzer, James R (2001). "Establishment Wage Differentials", *Urban Institute Research Paper*, available at http://www.urban.org/publications/410505.html.

Law, David and Nathan McLellan (2005). "The Contributions from Firm Entry, Exit and Continuation to Labour Productivity in New Zealand", *New Zealand Treasury* Working Paper, WP 05/01.

Lewis, Geoff and Steven Stillman (2005). "Regional Economic Performance in New Zealand: How Does Auckland Compare?", *New Zealand Treasury* Working Paper, WP 05/08.

Seyb, Allyson (2003). "The Longitudinal Business Frame" mimeo, Statistics New Zealand, available from www.stats.govt.nz.

Shewchuk, Jonathan Richard (1994). "An Introduction to the Conjugate Gradient Method without the agonizing pain – Edition 1 ¼", mimeo, School of Computer Science, Carnegie Mellon University.

Shimer R. and L. Smith (2000). "Assortative Matching and Search", *Econometrica* 68, pp. 343-69.

Solon, Gary, Robert Barsky, and Jonathan A. Parker (1994). "Measuring the Cyclicality of Real Wages: How Important is Composition Bias?", *Quarterly Journal of Economics*, Vol. 109, No. 1, pp. 1-25.

Statistics New Zealand (2003). *Linked Employer-Employee Feasibility Report*, available at http://www.stats.govt.nz/domino/external/pasfull/pasfull.nsf/e17092639373f80b4c2567ed0009410e/4c2567ef00247c6acc256dbf0015303f?OpenDocument.

Appendix 1 FTE Calculation -comparison with Household Labour Force Survey

In order to give a sense of the reliability and possible bias in our employment intensity measure, we have compared the estimated average employment intensity and the fraction estimated to be full-time with analogous estimates using Household Labour Force Survey (HLFS) data for workers over the sample period. The results are summarised in appendix Table A1. First, we estimated an analogous employment intensity measure for wage and salary workers from the June Quarter HLFS Income Supplement (HLFS-IS) over the sample period, using reported weekly earnings and non-employment incomes together with the relevant minimum wage rate.[32] Both the average employment intensity and the fraction employed full-time estimated in the LEED are 2–3 percent lower than their HLFS-IS counterparts: average employment intensity is 0.87 compared with 0.89, and the fraction estimated to be full-time is 0.73 compared with 0.76. Average employment intensity for males is about 10 percent higher than for females (0.91 versus 0.82 in the LEED), while the fraction employed full-time is roughly 20 percent higher (0.82 versus 0.64 in the LEED).

Second, we compare these earnings-based measures of employment intensity, with a direct hours-based measure of employment intensity using reported hours worked in the main HLFS quarterly surveys. Using workers' (usual) weekly hours worked, we first censored hours above at 40 hours, and constructed their employment intensity as the ratio of hours-worked to 40, and a full-time indicator for those working at least 40 hours. The estimated average employment intensity is 0.85 (0.92 for males and 0.77 for females), and the fraction working full-time is 0.66 (0.83 for males and 0.48 for females). The estimates for males are reasonably close to their LEED and HLFS-IS earnings-based counterparts, but the estimates for females are both somewhat lower (particularly the fraction employed full-time). Next, we repeated this exercise using 30 hours as the full-time threshold, which is the standard survey definition of full-time work. The results from this exercise are remarkably similar to the HLFS-IS earnings-based estimates, especially for females: average employment intensity is 0.89 (0.94 for males, 0.84 for females), and the fraction working full-time is 0.78 (0.89 for

[32] The HLFS-IS reference period varies by income source and, for wage and salary earners, by payment type (hourly wage versus salary), and is reported on a weekly basis in the data extract.

males, 0.66 for females). In comparison with our LEED estimates, the 40-hours based measure appears a better match for males, while the 30-hours based measure is closer for females.

In summary, we believe these results provide some assurance that, first, the LEED employment intensity construct has similar properties to analogous survey estimates and, second, in the absence of any direct hours measure, it provides a useful first-order adjustment for estimating differing levels of employment intensity across workers. Furthermore, a closer look at the reported hours distribution in the HLFS shows that as well as there being a substantial fraction of part-time employment, a large fraction of workers also work more than the standard full-time level. In fact, about one-third of workers report usual hours of less than 40 hours per week, one-third report 40 hours, and one-third report more than 40 hours in the HLFS over the sample period. Thus, using a single level of full-time employment will also bias downwards the employment level of those working long hours, and bias upwards their (hourly) earnings rate.

Appendix 2 Identification of Age, Year and Person Effects

Assuming age and year changes are perfectly synchronised, it is not possible to identify age, worker and year effects simultaneously in a balanced panel. In fact, with worker fixed effects, year and age effects are not separately identified. Identification is possible when factors that perturb the deterministic link between within-worker variation of age and time.

For example, in preliminary estimation, with each worker's age measured as their employment-weighted average over the year, we were able to obtain estimates of all the parameters in the model. However, the estimates are extremely sensitive to variations in weighting, sample selection and variable specification – e.g. using cumulative experience rather than age, top and bottom censoring of age, etc. Time effects are identified by changes in the composition of worker and firm effects over time. An apparent consequence of such fragile identification is that the estimates obtained are generally characterised by exaggerated age and time profiles, together with large changes in mean worker effects over time.

For example, appendix Figure A1 shows the male and female raw age-log(earnings) profiles, together with estimated profiles from a specification that allows sex specific quartiles in age, together with year dummy variables. The estimated patterns described in Figure A1 imply, first, a massive increase in aggregate earnings (30 percent) over the period compared with the raw average increase of 6 percent and, second, a similarly massive decrease in cohort quality – e.g. Figure A1(a) implies that, after controlling for life-cycle effects, the relative quality of 65–year-olds is on the order of 200 log-points compared with 15–year-olds. We believe these are simply incredible and essentially reflect poor identification.

Table 1

Sample Characteristics

	Unweighted				FTE Employment Weighted			
	Pooled Years	1999/00	2004/05	2000-05 Change	Pooled Years	1999/00	2004/05	2000-05 Change
Job Characteristics								
No. Job-year Observations	18,676,324	2,920,760	3,334,000	14.1%	18,676,324	2,920,760	3,334,000	14.1%
No. FTE Job-year Observations	8,018,349	1,239,560	1,452,851	17.2%	8,018,349	1,239,560	1,452,851	17.2%
FTE Employment	0.429	0.424	0.436	2.7%	0.782	0.781	0.784	0.5%
Fraction < FT Employment	0.690	0.694	0.693	-0.1%	0.451	0.459	0.457	-0.4%
Annual Earnings	$18,923	$18,538	$19,819	6.9%	$36,501	$36,074	$37,806	4.8%
FTE Earnings	$36,608	$36,487	$37,837	3.7%	$44,077	$43,680	$45,481	4.1%
Worker Characteristics								
No. Worker-year Observations	11,716,402	1,825,411	2,110,895	15.6%	11,716,402	1,825,411	2,110,895	15.6%
Age	36.4	35.8	37.2	4.0%	38.0	37.2	38.6	3.9%
Female	0.492	0.490	0.493	0.7%	0.462	0.462	0.461	-0.1%
Rec. Working-age Non-earnings	0.201	0.221	0.174	-21.3%	0.123	0.132	0.108	-18.0%
Received NZS	0.024	0.022	0.027	20.1%	0.015	0.013	0.019	47.0%
Fraction of year with:								
Earnings	0.796	0.789	0.801	1.6%	0.936	0.933	0.938	0.5%
Working-age Non-earnings	0.120	0.139	0.099	-28.9%	0.055	0.063	0.045	-28.5%
NZS Income	0.021	0.021	0.024	14.0%	0.013	0.012	0.016	37.7%
FTE Employment	0.684	0.679	0.688	1.4%	0.878	0.875	0.879	0.5%
Fraction < FT Employment	0.607	0.618	0.611	-1.2%	0.426	0.437	0.434	-0.7%
No. Jobs / year	1.59	1.60	1.58	-1.3%	1.57	1.58	1.55	-2.0%
Annual Earnings	$30,165	$29,662	$31,303	5.5%	$40,455	$39,962	$41,776	4.5%
Fraction from Main-job	0.908	0.909	0.908	-0.2%	0.914	0.916	0.914	-0.1%
FTE Earnings	$38,576	$38,302	$39,980	4.4%	$44,077	$43,680	$45,481	4.1%
Firm Characteristics								
No. of Firm-year Observations	1,211,202	195,039	213,760	9.6%	1,211,202	195,039	213,760	9.6%
Fraction of Year Observed	0.829	0.821	0.835	1.6%	0.984	0.982	0.985	0.3%
Annual Employment	15.4	15.0	15.6	4.2%	300.4	283.0	309.5	9.4%
FTE Employment	6.6	6.4	6.8	6.9%	162.6	152.7	170.6	11.7%
Average Earnings / Worker	$16,596	$15,890	$17,709	11.4%	$25,607	$25,122	$26,615	5.9%
FTE Earnings / Worker	$35,788	$35,062	$37,510	7.0%	$44,077	$43,680	$45,481	4.1%

Notes: A total of 2,776,361 workers, 322,713 firms and 9,729,904 jobs (worker-firm combinations) were observed over the period. Years are April-March – e.g. 2000 refers to April 1999-March 2000. All income values are in December quarter 2005 $ values, adjusted using the Consumers Price Index (CPI).
Symbols: … not applicable

Table 2

Results of Grouping Algorithm

	Largest Group	Fraction in Largest Group (%)	Second Largest Group	Average of all Other Groups	Total of all Groups
New Zealand (PBN All jobs)					
Observations	18,624,844	99.7	201	4	18,676,324
FTE Employment	7,985,444	99.6	103	3	8,018,349
Persons	2,760,560	99.4	76	1	2,776,361
Firms	309,713	96.0	12	1	322,713
Groups	1	0.0	1	12,198	12,200
Estimable effects	3,070,272	99.5	75	…	3,086,874
New Zealand (PBN Main jobs)					
Observations	11,610,712	99.1	378	5	11,716,402
FTE Employment	7,232,463	99.1	167	3	7,299,365
Persons	2,744,214	98.8	86	1	2,776,361
Firms	263,358	91.6	5	1	287,480
Groups	1	0.0	1	22,692	22,694
Estimable effects	3,007,571	98.9	90	…	3,041,147
France					
Observations	4,682,420	88.3	51	4	5,305,108
Persons	974,985	83.6	31	1	1,166,305
Firms	334,637	64.2	1	1	521,180
Groups	1	0.0	1	141,550	141,552
Estimable effects	1,309,621	84.7	31	…	1,545,933
Washington State					
Observations	3,999,598	99.1	276	15	4,036,171
Persons	292,945	98.7	33	2	296,801
Firms	81,107	94.5	3	2	85,864
Groups	1	0.0	1	2,426	2,428
Estimable effects	374,051	98.4	35	…	380,237
Seven US States					
Observations	285,402,315	99.4	90	4	287,241,891
Persons	64,441,382	94.3	38	9	68,329,212
Firms	3,200,067	87.4	8	1	3,662,974
Groups	1	0.0	1	430,529	430,531
Estimable effects	67,641,448	94.5	45	…	71,561,655[a]

Notes: France and Washington State information from Abowd et al (2002, p. 6). 'Seven US States' information covers the states of California, Florida, Illinois, Maryland, Minnesota, North Carolina and Texas (see Abowd et al, 2003, table 4).
[a] The number of estimable effects (N+J-G) differs from the N+J-1 reported (incorrectly) in Abowd et al (2003).
Symbols: … not applicable

Table 3

A: Summary of Earnings Components

	Log(Earning Rate) (y_{ijt})	First-stage Covariates (β_{gAt})	Time Effects (τ_t)	Worker Effects (θ_i)	Firm Effects (ψ_j)	Residual (u_{ijt})
1. FTE Weighted	10.54	10.54	0	0.00	0.00	0.00
(Overall R^2=0.903)	(0.34)	(0.17)	(0.01)	(0.24)	(0.10)	(0.10)
Simple R^2	…	0.26	…	0.49	0.25	…
Marginal R^2	…	0.24	…	0.49	0.09	…
2. Unweighted	10.32	10.32	0	0.00	0.00	0.00
	(0.55)	(0.30)	(0.02)	(0.36)	(0.14)	(0.22)
3. FTE Weighted	10.57	10.54	0	0.01	0.02	0.00
(χ_j >25)	(0.34)	(0.17)	(0.01)	(0.24)	(0.09)	(0.10)

B: Correlation Between Earnings Components

Description	(y_{ijt},τ_t)	(y_{ijt},β_{gAt})	(y_{ijt},θ_i)	(y_{ijt},ψ_j)	(y_{ijt},u_{ijt})	(β_{gAt},θ_i)	(β_{gAt},ψ_j)	(θ_i,ψ_j)
1. FTE Weighted	0.046	0.507	0.697	0.495	0.305	-0.088	0.202	0.119
	[0.000]	[0.029]	[0.055]	[0.017]	[0.011]	[-0.004]	[0.004]	[0.003]
2. Unweighted	0.053	0.538	0.645	0.487	0.400	-0.087	0.214	0.175
	[0.001]	[0.088]	[0.127]	[0.038]	[0.049]	[-0.009]	[0.009]	[0.009]
3. FTE Weighted (χ_j >25)	0.044	0.519	0.714	0.518	0.311	-0.073	0.264	0.168
	[0.000]	[0.029]	[0.057]	[0.016]	[0.011]	[-0.003]	[0.004]	[0.004]
Comparison with International Results								
Abowd et al (2002):								
• Washington	…	0.304	0.511	0.518	0.306	-0.530	0.143	-0.025
• France	…	0.141	0.704	0.201	0.169	-0.068	0.023	-0.283
Abowd et al (2003)	…	0.224	0.468	0.484	0.402	-0.553	0.095	0.080

Notes: In the first panel, the entries in parentheses are standard deviations; in the second panel, the entries in square brackets are covariances.
Symbols: … not applicable

Table 4

Comparison of One-way and Two-way Fixed Effects Estimates

		Standard Deviation	Correlations					
			One way		Order dependent		Two way	
			θ_i	ψ_j	θ_i	ψ_j	θ_i	ψ_j
One way	θ_i	0.266	1.000	0.597	0.807	0.220	0.938	0.419
	ψ_j	0.173		1.000	0.078	0.502	0.397	0.762
Order dependent	θ_i	0.201			1.000	0.008	0.898	0.017
	ψ_j	0.029				1.000	0.052	0.721
Two way	θ_i	0.237					1.000	0.119
	ψ_j	0.104						1.000

Notes: Estimates are based on regressions weighted by FTE employment. The order-dependent worker effect is estimated conditional on the estimation of the one-way firm effect, and the order-dependent firm effect is estimated conditional on the estimation of the one-way worker effect.

Table 5

Summary of Earnings Components

Mean (No. Obs)	Log(Earning Rate) (y_{ijt})	First-stage Covariates (β_{gAt})	Time Effects (τ_t)	Worker Effects (θ_i)	Firm Effects (ψ_j)	Residual (u_{ijt})
Full sample	10.54	10.54	0	0.00	0.00	0.00
(n=18.7m)	(0.34)	(0.17)	(0.01)	(0.24)	(0.10)	(0.10)
Worker Demographics						
Male	10.67	10.67	0.00	-0.02	0.02	0.00
	(0.36)	(0.17)	(0.01	(0.26)	(0.11)	(0.11)
Female	10.40	10.40	-0.00	0.02	-0.02	0.00
	(0.29)	(0.11)	(0.01)	(0.21)	(0.09)	(0.10)
Aged (in years):						
Under 20	9.94	9.94	-0.00	0.06	-0.07	0.01
(n=2.7m)	(0.19)	(0.07)	(0.01)	(0.14)	(0.07)	(0.08)
20–29	10.42	10.42	-0.00	0.00	-0.00	-0.00
(n=5.0m)	(0.23)	(0.09)	(0.01)	(0.16)	(0.09)	(0.10)
30–39	10.65	10.65	-0.00	-0.02	0.02	-0.00
(n=4.4m)	(0.33)	(0.11)	(0.01)	(0.26)	(0.11)	(0.11)
40–49	10.66	10.66	0.00	-0.01	0.01	0.00
(n=3.7m)	(0.38)	(0.15)	(0.02)	(0.29)	(0.11)	(0.11)
50–59	10.64	10.64	0.00	0.00	-0.00	-0.00
(n=2.3m)	(0.39)	(0.14)	(0.02)	(0.31)	(0.12)	(0.11)
60–69	10.54	10.54	0.00	0.04	-0.02	-0.02
(n=0.6m)	(0.36)	(0.11)	(0.01)	(0.29)	(0.12)	(0.13)
Industry						
A Agriculture, Forestry	10.29	10.51	0.00	-0.13	-0.09	0.00
and Fishing	(0.19)	(0.14)	(0.01)	(0.14)	(0.08)	(0.08)
B Mining	10.90	10.70	0.00	0.00	0.19	0.00
	(0.37)	(0.17)	(0.02)	(0.26)	(0.15)	(0.12)
C Manufacturing	10.62	10.61	-0.00	-0.06	0.07	0.00
	(0.33)	(0.19)	(0.02)	(0.24)	(0.10)	(0.11)
D Electricity, Gas and	10.94	10.66	-0.00	0.11	0.18	0.00
Water Supply	(0.41)	(0.17))	(0.02)	(0.30)	(0.10)	(0.14)
E Construction	10.58	10.62	0.00	-0.06	0.01	0.00
	(0.29)	(0.19)	(0.02)	(0.21)	(0.11)	(0.11)
F Wholesale Trade	10.66	10.60	-0.00	0.03	0.04	0.00
	(0.35)	(0.18)	(0.02)	(0.30)	(0.11)	(0.12)
G Retail Trade	10.25	10.44	-0.00	-0.08	-0.11	0.00
	(0.29)	(0.20)	(0.01)	(0.19)	(0.08)	(0.09)
H Accommodation,	10.16	10.40	0.00	-0.10	-0.14	0.00
Cafes and Restaurants	(0.21)	(0.15)	(0.01)	(0.15)	(0.05)	(0.08)
I Transport and Storage	10.65	10.63	0.00	-0.03	0.05	0.00
	(0.34)	(0.17)	(0.01)	(0.24)	(0.11)	(0.11)
J Communication	10.70	10.55	-0.00	0.04	0.11	0.00
Services	(0.40)	(0.16)	(0.01)	(0.27)	(0.11)	(0.12)
K Finance and Insurance	10.82	10.54	-0.00	0.14	0.15	0.00
	(0.45)	(0.16)	(0.02)	(0.33)	(0.11)	(0.14)
L Property and Business	10.65	10.55	0.00	0.07	0.03	0.00
Services	(0.36)	(0.15)	(0.01)	(0.26)	(0.11)	(0.11)
M Government Admin	10.75	10.57	-0.00	0.10	0.08	0.00
and Defence	(0.33)	(0.17)	(0.02)	(0.26)	(0.05)	(0.11)
N Education	10.62	10.53	0.00	0.12	-0.03	0.00
	(0.31)	(0.13)	(0.01)	(0.26)	(0.05)	(0.11)
O Health & Community	10.46	10.48	-0.00	0.03	-0.04	0.00
Services	(0.34)	(0.12)	(0.01)	(0.27)	(0.09)	(0.10)
P Cultural and	10.51	10.52	0.00	0.02	-0.03	0.00
Recreational Services	(0.32)	(0.16)	(0.01)	(0.22)	(0.09)	(0.10)
Q Personal and Other	10.51	10.56	0.00	-0.03	-0.02	0.00
Services	(0.33)	(0.18)	(0.01)	(0.21)	(0.13)	(0.10)

Table 5 (continued)

Summary of Earnings Components

Mean (No. Obs)	Log(Earning Rate) (y_{ijt})	First-stage Covariates (β_{gAt})	Time Effects (τ_t)	Worker Effects (θ_i)	Firm Effects (ψ_j)	Residual (u_{ijt})
Metro versus Non-metro						
Auckland	10.62	10.55	0.00	0.04	0.04	0.00
	(0.36)	(0.17)	(0.01)	(0.26)	(0.10)	(0.11)
Wellington	10.64	10.55	-0.00	0.06	0.03	0.00
	(0.38)	(0.17)	(0.01)	(0.27)	(0.11)	(0.11)
Christchurch	10.49	10.55	0.00	-0.03	-0.03	0.00
	(0.31)	(0.17)	(0.01)	(0.22)	(0.10)	(0.10)
Other	10.46	10.54	0.00	-0.04	-0.03	0.00
	(0.30)	(0.17)	(0.01)	(0.21)	(0.10)	(0.10)
Employment Stability						
Firms:						
Part year, Part-time	10.44	10.51	-0.00	-0.03	-0.04	-0.00
	(0.29)	(0.16)	(0.01)	(0.21)	(0.08)	(0.10)
Full year, Part-time	10.50	10.54	0.00	-0.00	-0.04	-0.00
	(0.37)	(0.18)	(0.02)	(0.28)	(0.13)	(0.11)
Part year, Full-time	10.82	10.61	-0.00	0.08	0.12	0.00
	(0.39)	(0.18)	(0.02)	(0.30)	(0.11)	(0.13)
Full year, Full-time	10.81	10.63	0.00	0.06	0.11	0.00
	(0.43)	(0.20)	(0.02)	(0.33)	(0.16)	(0.12)
Workers:						
Part year, Part-time	10.14	10.42	-0.00	-0.19	-0.05	-0.04
	(0.17)	(0.12)	(0.01)	(0.12)	(0.06)	(0.09)
Full year, Part-time	10.18	10.45	0.00	-0.18	-0.06	-0.02
	(0.24)	(0.17)	(0.01)	(0.18)	(0.10)	(0.10)
Part year, Full-time	10.77	10.57	-0.00	0.11	0.04	0.05
	(0.34)	(0.15)	(0.01)	(0.26)	(0.11)	(0.15)
Full year, Full-time	10.77	10.61	0.00	0.11	0.04	0.01
	(0.37)	(0.19)	(0.02)	(0.31)	(0.13)	(0.11)

Notes: Estimates are based on regressions weighted by FTE employment.

Table 6

Summary of Earnings Components

Mean (No. Obs)	Log(Earning Rate) (y_{ijt})	First-stage Covariates (β_{gAt})	Time Effects (τ_t)	Worker Effects (θ_i)	Firm Effects (ψ_j)	Residual (u_{ijt})
Full sample	10.54	10.54	0	0.00	0.00	0.00
(n=18.7m)	(0.34)	(0.17)	(0.01)	(0.24)	(0.10)	(0.10)
Annual Composition Effects						
1999/2000	10.52	10.52	-0.03	0.02	0.01	0.00
(n=2.9m)	(0.35)	(0.19)		(0.25)	(0.11)	(0.12)
2000/01	10.52	10.52	-0.02	0.02	0.00	0.00
(n=3.0m)	(0.34)	(0.18)		(0.24)	(0.10)	(0.10)
2001/02	10.53	10.53	-0.01	0.01	0.00	0.00
(n=3.1m)	(0.33)	(0.17)		(0.24)	(0.10)	(0.10)
2002/03	10.54	10.54	0.00	-0.00	-0.00	0.00
(n=3.1m)	(0.33)	(0.16)		(0.24)	(0.10)	(0.10)
2003/04	10.57	10.57	0.02	-0.01	-0.00	0.00
(n=3.2m)	(0.33)	(0.16)		(0.23)	(0.10)	(0.10)
2004/05	10.58	10.58	0.03	-0.03	-0.00	0.00
(n=3.3m)	(0.33)	(0.16)		(0.23)	(0.11)	(0.11)
Panel Transitions						
Firms:						
Continuers	10.56	10.55	-0.00	0.01	0.00	-0.00
(n=15.1m)	(0.34)	(0.17)	(0.01)	(0.25)	(0.10)	(0.11)
Single entrants	10.45	10.53	0.02	-0.05	-0.05	-0.00
(n=1.8m)	(0.28)	(0.15)	(0.01)	(0.20)	(0.11)	(0.10)
Single exiters	10.48	10.52	-0.02	-0.02	0.01	0.00
(n=1.4m)	(0.31)	(0.16)	(0.01)	(0.21)	(0.12)	(0.11)
Other	10.40	10.51	-0.00	-0.07	-0.04	-0.00
(n=0.5m)	(0.25)	(0.13)	(0.01)	(0.17)	(0.11)	(0.09)
Workers:						
Continuers	10.59	10.57	-0.00	0.02	0.01	0.00
(n=11.8m)	(0.35)	(0.17)	(0.01)	(0.25)	(0.11)	(0.11)
Single entrants	10.32	10.45	0.02	-0.10	-0.04	0.00
(n=3.0m)	(0.27)	(0.17)	(0.01)	(0.18)	(0.09)	(0.08)
Single exiters	10.52	10.51	-0.02	0.02	0.00	0.00
(n=1.7m)	(0.34)	(0.16)	(0.01)	(0.25)	(0.10)	(0.11)
Other	10.36	10.49	0.00	-0.11	-0.03	0.00
(n=2.2m)	(0.26)	(0.14)	(0.01)	(0.18)	(0.08)	(0.09)

Notes: Estimates are based on regressions weighted by FTE employment.

Table 7

Correlation Between Earnings Components

Description	(y_{ijt},τ_t)	(y_{ijt},β_{gAt})	(y_{ijt},θ_i)	(y_{ijt},ψ_j)	(y_{ijt},u_{ijt})	(β_{gAt},θ_i)	(β_{gAt},ψ_j)	(θ_i,ψ_j)
Full sample	0.046	0.507	0.697	0.495	0.305	-0.088	0.202	0.119
	[0.000]	[0.029]	[0.055]	[0.017]	[0.011]	[-0.004]	[0.004]	[0.003]
Worker Demographics								
Male	0.035	0.487	0.721	0.484	0.305	-0.070	0.166	0.126
Female	0.064	0.380	0.768	0.479	0.335	-0.067	0.150	0.131
Aged (in years):								
Under 20	0.233	0.354	0.679	0.437	0.512	-0.162	0.133	0.036
20–29	0.014	0.399	0.686	0.524	0.386	-0.073	0.134	0.121
30–39	0.032	0.320	0.809	0.506	0.339	-0.057	0.012	0.186
40–49	0.038	0.386	0.791	0.485	0.291	-0.076	0.189	0.136
50–59	0.041	0.373	0.794	0.460	0.299	-0.076	0.186	0.097
60–69	0.044	0.316	0.780	0.426	0.346	-0.096	0.155	0.050
Industry								
A Agric, Forest and Fish[(5)]	0.102	0.471	0.424	0.388	0.399	-0.385	0.036	-0.049
B Mining	0.104	0.367	0.698	0.528	0.339	-0.158	0.067	0.126
C Manufacturing	0.045	0.488	0.652	0.431	0.327	-0.163	0.160	0.037
D Electricity, Gas &Water	0.049	0.493	0.786	0.382	0.356	0.053	0.123	0.109
E Construction	0.038	0.526	0.528	0.415	0.333	-0.236	0.107	-0.044
F Wholesale Trade	0.028	0.464	0.751	0.427	0.309	-0.056	0.083	0.135
G Retail Trade	0.085	0.624	0.537	0.375	0.333	-0.173	0.141	0.011
H Accomm, Cafes & Restau	0.101	0.552	0.502	0.320	0.383	-0.285	0.072	0.038
I Transport and Storage	0.035	0.423	0.687	0.475	0.330	-0.175	0.093	0.139
J Communication Services	0.008	0.504	0.777	0.564	0.337	0.029	0.218	0.289
K Finance and Insurance	0.034	0.541	0.833	0.301	0.389	0.216	0.057	0.035
L Property and Bus Serv	0.025	0.484	0.777	0.485	0.296	0.031	0.137	0.170
M Govt Admin and Defence	0.055	0.469	0.771	0.202	0.339	-0.051	0.044	0.025
N Education	0.046	0.370	0.818	0.181	0.355	-0.074	0.046	0.014
O Health & Commun Serv	0.058	0.390	0.828	0.462	0.298	-0.009	0.147	0.197
P Cultural and Recr Serv	0.046	0.495	0.734	0.499	0.310	-0.047	0.135	0.197
Q Personal & Other Serv	0.040	0.555	0.612	0.591	0.272	-0.122	0.267	0.108
Compositional Change								
1999/2000	…	0.534	0.664	0.502	0.291	-0.088	0.218	0.123
2000/01	…	0.531	0.685	0.501	0.281	-0.090	0.218	0.125
2001/02	…	0.501	0.712	0.501	0.309	-0.092	0.211	0.127
2002/03	…	0.490	0.720	0.500	0.322	-0.089	0.203	0.125
2003/04	…	0.488	0.720	0.492	0.319	-0.080	0.196	0.113
2004/05	…	0.484	0.706	0.489	0.318	-0.068	0.188	0.099
Firms:								
Continuers	0.055	0.510	0.705	0.494	0.304	-0.085	0.221	0.126
Single Entrants	0.036	0.479	0.646	0.489	0.310	-0.112	0.103	0.052
Single Exiters	0.014	0.499	0.646	0.529	0.325	-0.119	0.168	0.110
Other	0.011	0.451	0.599	0.519	0.334	-0.166	0.081	0.043
Workers:								
Continuers	0.095	0.489	0.701	0.483	0.310	-0.097	0.197	0.101
Single Entrants	0.101	0.581	0.602	0.497	0.289	-0.154	0.209	0.088
Single Exiters	0.034	0.433	0.736	0.504	0.307	-0.114	0.165	0.165
Other	0.034	0.462	0.659	0.502	0.357	-0.156	0.144	0.135

Table 7 (continued)

Correlation Between Earnings Components

Description	(y_{ijt},τ_t)	(y_{ijt},β_{gAt})	(y_{ijt},θ_i)	(y_{ijt},ψ_j)	(y_{ijt},u_{ijt})	(β_{gAt},θ_i)	(β_{gAt},ψ_j)	(θ_i,ψ_j)
Metro versus Non-metro								
Auckland	0.041	0.503	0.734	0.470	0.306	-0.042	0.187	0.129
Wellington	0.028	0.496	0.760	0.518	0.295	-0.010	0.198	0.195
Christchurch	0.057	0.530	0.663	0.455	0.311	-0.109	0.217	0.045
Other	0.057	0.524	0.624	0.483	0.317	-0.159	0.217	0.044
Employment Stability								
Firms:								
Part year, Part-time	0.046	0.506	0.676	0.446	0.334	-0.133	0.191	0.110
Full year, Part-time	0.063	0.453	0.721	0.422	0.296	-0.099	0.091	0.046
Part year, Full-time	0.014	0.424	0.762	0.390	0.322	-0.067	0.051	0.119
Full year, Full-time	0.056	0.433	0.735	0.405	0.275	-0.069	0.068	0.001
Workers:								
Part year, Part-time	0.096	0.482	0.338	0.442	0.410	-0.040	0.129	-0.012
Full year, Part-time	0.081	0.486	0.427	0.457	0.348	-0.380	0.138	-0.040
Part year, Full-time	0.022	0.349	0.692	0.403	0.402	-0.143	0.073	0.037
Full year, Full-time	0.047	0.382	0.732	0.401	0.177	-0.179	0.114	0.018

Notes: Estimates are based on regressions weighted by FTE employment.
Symbols: … not applicable

Figure 1

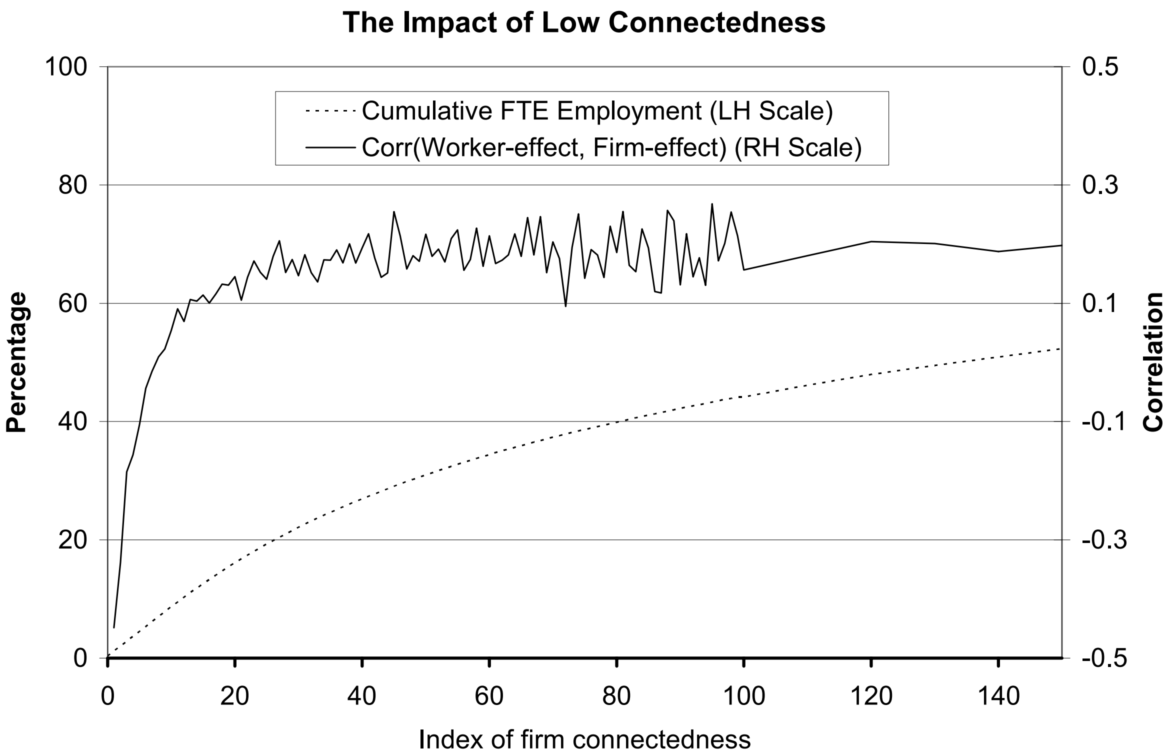
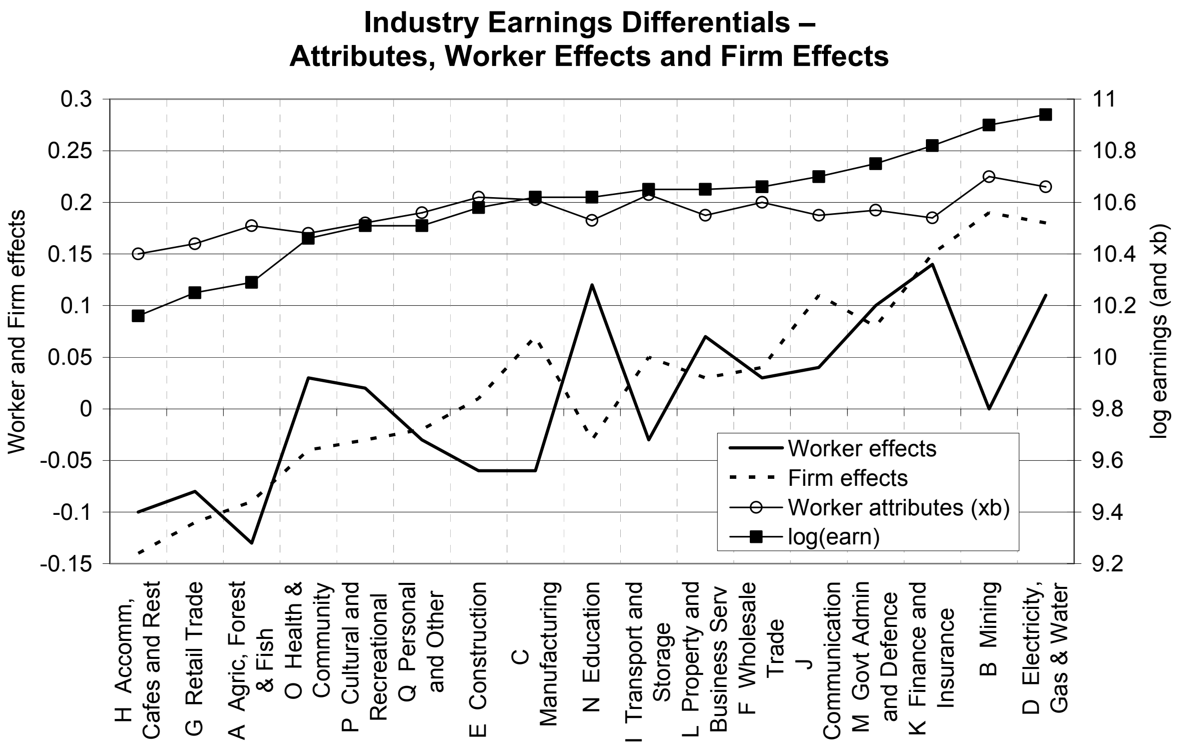

The Impact of Low Connectedness
Cumulative FTE Employment (LH Scale)
Corr(Worker-effect, Firm-effect) (RH Scale)
Percentage
Correlation
100
80
60
40
20
0
0.5
0.3
0.1
-0.1
-0.3
-0.5
0
20
40
60
80
100
120
140
Index of firm connectedness

Figure 2



Industry Earnings Differentials –
Attributes, Worker Effects and Firm Effects
Worker and Firm effects
log earnings (and xb)
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
11
10.8
10.6
10.4
10.2
10
9.8
9.6
9.4
9.2
Worker effects
Firm effects
Worker attributes (xb)
log(earn)
H Accomm, Cafes and Rest
G Retail Trade
A Agric, Forest & Fish
O Health & Community
P Cultural and Recreational
Q Personal and Other
E Construction
C Manufacturing
N Education
I Transport and Storage
L Property and Business Serv
F Wholesale Trade
J Communication
M Govt Admin and Defence
K Finance and Insurance
B Mining
D Electricity, Gas & Water

Appendix Table A1

Comparison of LEED and HLFS-based Measures of Employment Intensity

	All	Males	Females
Average Monthly Employment Intensity			
LEED (Earnings)	0.865	0.909	0.818
HLFS-IS (Earnings)	0.888	0.936	0.839
HLFS (40 hours)	0.845	0.921	0.766
HLFS (30 hours)	0.890	0.942	0.837
Fraction Employed Full-time			
LEED (Earnings)	0.733	0.819	0.643
HLFS-IS (Earnings)	0.764	0.868	0.656
HLFS (40 hours)	0.660	0.829	0.484
HLFS (30 hours)	0.778	0.894	0.658

Notes: All estimates are based on workers aged 15 and over. The LEED estimates are based on PAYE employees, and the HLFS and HLFS-IS are based on wage and salary workers. The LEED and HLFS-IS employment intensity is measured as the lesser ratio of employment earnings to total income or full-time (40 hours per week) minimum wage earnings. The HLFS employment intensity is measured as reported “usual hours” worked, censored at 40 (or 30) hours, as a fraction of 40 (or 30).

Appendix Table A2

Summary of Earnings Components, Excluding Low-connected Firms (χ_j <25)

Mean (No. Obs)	Log(Earning Rate) (y_{ijt})	First-stage Covariates (β_{gAt})	Time Effects (τ_t)	Worker Effects (θ_i)	Firm Effects (ψ_j)	Residual (u_{ijt})
Full sample	10.57	10.54	0.00	0.01	0.02	0.00
	(0.34)	(0.17)	(0.01)	(0.24)	(0.09)	(0.10)
Worker Demographics						
Male	10.70	10.67	-0.00	-0.01	0.04	0.00
	(0.36)	(0.17)	(0.01)	(0.26)	(0.09)	(0.11)
Female	10.42	10.40	-0.00	0.03	-0.01	-0.00
	(0.29)	(0.11)	(0.01)	(0.22)	(0.08)	(0.10)
Aged (years):						
Under 20	9.93	9.94	-0.00	0.05	-0.07	0.01
	(0.19)	(0.07)	(0.01)	(0.13)	(0.06)	(0.08)
20–29	10.44	10.42	-0.00	0.01	0.01	-0.00
	(0.23)	(0.09)	(0.01)	(0.15)	(0.08)	(0.10)
30–39	10.67	10.65	-0.00	-0.01	0.04	-0.00
	(0.33)	(0.11)	(0.01)	(0.26)	(0.09)	(0.11)
40–49	10.69	10.66	0.00	0.01	0.03	0.00
	(0.38)	(0.15)	(0.02)	(0.30)	(0.10)	(0.11)
50–59	10.67	10.64	0.00	0.02	0.01	0.00
	(0.39)	(0.15)	(0.02)	(0.31)	(0.10)	(0.12)
60–69	10.58	10.54	0.00	0.06	0.00	-0.02
	(0.36)	(0.11)	(0.01)	(0.29)	(0.09)	(0.13)
Industry						
A Agriculture, Forestry	10.31	10.52	0.01	-0.13	-0.08	0.00
and Fishing	(0.18)	(0.13)	(0.01)	(0.13)	(0.06)	(0.08)
B Mining	10.94	10.71	0.00	0.01	0.23	0.00
	(0.37)	(0.16)	(0.02)	(0.27)	(0.14)	(0.13)
C Manufacturing	10.65	10.62	-0.00	-0.05	0.09	0.00
	(0.33)	(0.18)	(0.02)	(0.25)	(0.09)	(0.11)
D Electricity, Gas and	10.95	10.66	-0.00	0.11	0.19	0.00
Water Supply	(0.42)	(0.17))	(0.02)	(0.31)	(0.10)	(0.14)
E Construction	10.66	10.65	0.00	-0.05	0.05	-0.00
	(0.29)	(0.18)	(0.02)	(0.21)	(0.08)	(0.11)
F Wholesale Trade	10.69	10.59	-0.00	0.04	0.06	0.00
	(0.38)	(0.18)	(0.02)	(0.28)	(0.09)	(0.12)
G Retail Trade	10.24	10.41	-0.00	-0.07	-0.11	-0.00
	(0.29)	(0.20)	(0.01)	(0.19)	(0.06)	(0.09)
H Accommodation,	10.17	10.40	-0.00	-0.09	-0.13	-0.00
Cafes and Restaurants	(0.21)	(0.15)	(0.01)	(0.15)	(0.04)	(0.08)
I Transport and Storage	10.68	10.64	-0.00	-0.02	0.07	-0.00
	(0.32)	(0.17)	(0.01)	(0.24)	(0.09)	(0.11)
J Communication	10.72	10.55	-0.00	0.05	0.12	0.00
Services	(0.40)	(0.16)	(0.01)	(0.27)	(0.10)	(0.13)
K Finance and	10.86	10.54	-0.00	0.14	0.17	0.00
Insurance	(0.45)	(0.16)	(0.02)	(0.33)	(0.07)	(0.14)
L Property and Business	10.67	10.55	0.00	0.07	0.05	0.00
Services	(0.35)	(0.15)	(0.01)	(0.25)	(0.09)	(0.11)
M Government Admin	10.75	10.57	-0.00	0.10	0.08	0.00
and Defence	(0.33)	(0.17)	(0.02)	(0.26)	(0.05)	(0.11)
N Education	10.63	10.53	0.00	0.13	-0.03	-0.00
	(0.31)	(0.13)	(0.01)	(0.26)	(0.04)	(0.11)
O Health & Community	10.48	10.48	-0.00	0.03	-0.03	0.00
Services	(0.35)	(0.12)	(0.01)	(0.27)	(0.08)	(0.10)
P Cultural and	10.54	10.52	0.00	0.03	-0.01	0.00
Recreational Services	(0.33)	(0.16)	(0.01)	(0.23)	(0.08)	(0.10)
Q Personal and Other	10.63	10.59	0.00	-0.00	0.04	0.00
Services	(0.33)	(0.18)	(0.01)	(0.21)	(0.10)	(0.10)

(Continued)

Summary of Earnings Components, Excluding Low-connected Firms (χ_j <25)

Mean (No. Obs)	Log(Earning Rate) (y_{ijt})	First-stage Covariates (β_{gAt})	Time Effects (τ_t)	Worker Effects (θ_i)	Firm Effects (ψ_j)	Residual (u_{ijt})
Annual Composition						
1999/2000	10.55	10.52	-0.03	0.04	0.02	-0.00
	(0.35)	(0.19)		(0.25)	(0.09)	(0.12)
2000/01	10.54	10.52	-0.02	0.03	0.02	-0.00
	(0.34)	(0.18)		(0.24)	(0.09)	(0.10)
2001/02	10.56	10.53	-0.01	0.02	0.02	-0.00
	(0.33)	(0.16)		(0.24)	(0.09)	(0.10)
2002/03	10.56	10.54	0.00	0.01	0.01	0.00
	(0.33)	(0.16)		(0.24)	(0.09)	(0.10)
2003/04	10.59	10.57	0.02	-0.00	0.01	0.00
	(0.33)	(0.16)		(0.23)	(0.09)	(0.10)
2004/05	10.61	10.58	0.03	-0.02	0.01	0.00
	(0.33)	(0.16)		(0.23)	(0.09)	(0.11)
Panel Transitions						
Firms:						
Continuers	10.58	10.55	-0.00	0.01	0.02	-0.00
	(0.34)	(0.17)		(0.24)	(0.09)	(0.11)
Single entrants	10.47	10.52	0.01	-0.04	-0.02	-0.00
	(0.28)	(0.14)	(0.01)	(0.18)	(0.08)	(0.09)
Single exiters	10.55	10.53	-0.02	0.00	0.04	-0.00
	(0.32)	(0.16)	(0.01)	(0.21)	(0.09)	(0.11)
Other	10.47	10.51	-0.00	-0.04	0.00	-0.00
	(0.25)	(0.12)	(0.01)	(0.16)	(0.08)	(0.09)
Workers:						
Continuers	10.62	10.57	-0.00	0.03	0.02	0.00
	(0.35)	(0.17)	(0.01)	(0.25)	(0.10)	(0.11)
Single entrants	10.33	10.44	0.02	-0.10	-0.02	-0.00
	(0.27)	(0.16)	(0.01)	(0.18)	(0.07)	(0.08)
Single exiters	10.55	10.51	-0.02	0.04	0.02	0.00
	(0.35)	(0.16)	(0.01)	(0.25)	(0.09)	(0.11)
Other	10.37	10.48	0.00	-0.10	-0.01	-0.00
	(0.26)	(0.13)	(0.01)	(0.18)	(0.07)	(0.09)
Metro versus Non-metro						
Auckland	10.64	10.54	-0.00	0.05	0.05	-0.00
	(0.36)	(0.17)	(0.01)	(0.26)	(0.09)	(0.11)
Wellington	10.67	10.55	-0.00	0.08	0.05	0.00
	(0.38)	(0.17)	(0.01)	(0.27)	(0.10)	(0.11)
Christchurch	10.51	10.55	-0.00	-0.02	-0.01	0.00
	(0.32)	(0.17)	(0.01)	(0.22)	(0.08)	(0.10)
Other	10.49	10.54	0.00	-0.03	-0.01	-0.00
	(0.30)	(0.17)	(0.01)	(0.21)	(0.09)	(0.10)
Employment Stability						
Firms:						
Part year, Part-time	10.46	10.51	-0.00	-0.02	-0.03	-0.00
	(0.29)	(0.16)	(0.01)	(0.21)	(0.07)	(0.10)
Full year, Part-time	10.60	10.55	0.00	0.03	0.01	-0.00
	(0.38)	(0.18)	(0.02)	(0.29)	(0.09)	(0.11)
Part year, Full-time	10.87	10.62	-0.00	0.10	0.15	0.00
	(0.39)	(0.18)	(0.02)	(0.30)	(0.09)	(0.13)
Full year, Full-time	10.88	10.65	0.00	0.08	0.15	0.00
	(0.41)	(0.19)	(0.02)	(0.33)	(0.10)	(0.13)
Workers:						
Part year, Part-time	10.15	10.42	-0.00	-0.18	0.04	-0.04
	(0.17)	(0.12)	(0.01)	(0.12)	(0.05)	(0.09)
Full year, Part-time	10.20	10.44	0.00	-0.18	-0.05	-0.02
	(0.24)	(0.17)	(0.01)	(0.17)	(0.08)	(0.10)
Part year, Full-time	10.81	10.57	-0.00	0.13	0.05	0.05
	(0.34)	(0.15)	(0.01)	(0.26)	(0.09)	(0.16)
Full year, Full-time	10.80	10.61	-0.00	0.12	0.05	0.01
	(0.37)	(0.19)	(0.02)	(0.31)	(0.11)	(0.11)

Appendix Table A3

Correlation Between Earnings Components, Excluding Low-connected Firms (χ_j <25)

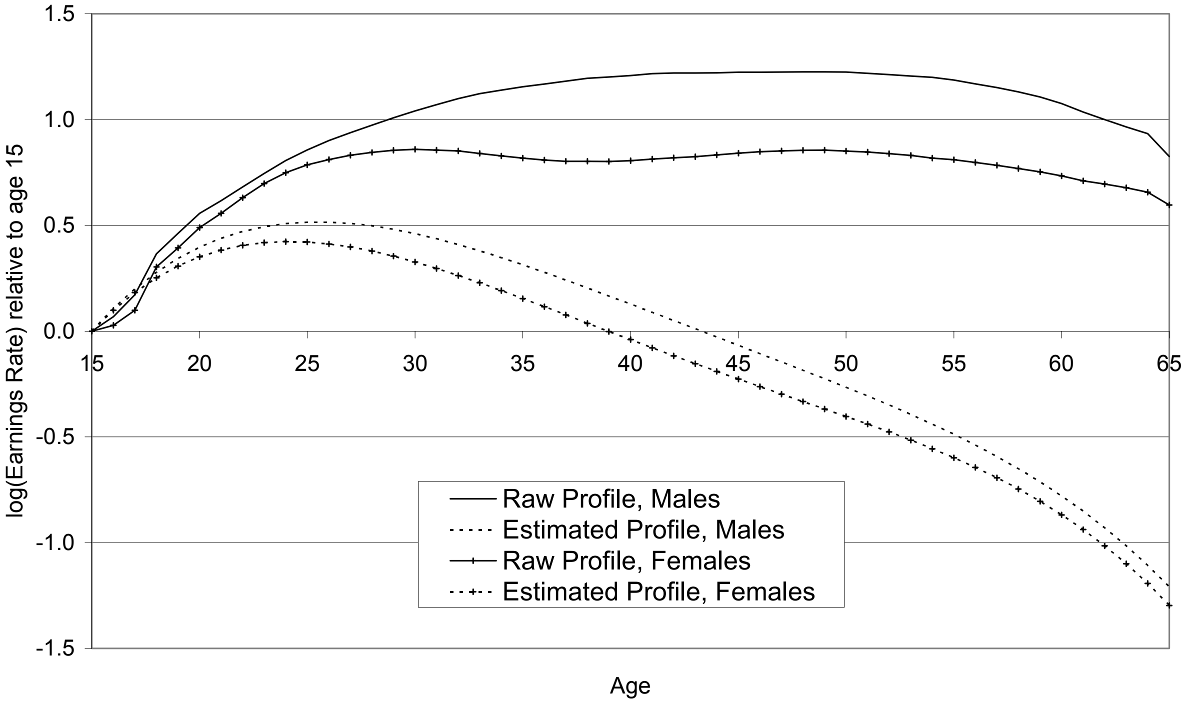
Description	(y_{ijt},τ_t)	(y_{ijt},β_{gAt})	(y_{ijt},θ_i)	(y_{ijt},ψ_j)	(y_{ijt},u_{ijt})	(β_{gAt},θ_i)	(β_{gAt},ψ_j)	(θ_i,ψ_j)
Full sample	0.044	0.519	0.714	0.518	0.311	-0.073	0.264	0.168
Composition Effects								
1999/2000	…	0.547	0.680	0.519	0.292	-0.067	0.277	0.162
2000/01	…	0.544	0.701	0.520	0.286	-0.072	0.276	0.165
2001/02	…	0.513	0.727	0.520	0.314	-0.077	0.268	0.169
2002/03	…	0.502	0.736	0.522	0.328	-0.074	0.264	0.173
2003/04	…	0.501	0.738	0.517	0.327	-0.066	0.261	0.169
2004/05	…	0.497	0.726	0.516	0.327	-0.057	0.258	0.163
Firms:								
Continuers	0.053	0.519	0.717	0.510	0.308	-0.072	0.267	0.159
Single entrants	0.039	0.513	0.684	0.592	0.323	-0.101	0.234	0.267
Single exiters	-0.009	0.521	0.688	0.552	0.342	-0.076	0.259	0.203
Other	-0.015	0.494	0.667	0.576	0.364	-0.114	0.212	0.238
Workers:								
Continuers	0.094	0.497	0.716	0.494	0.315	-0.087	0.249	0.134
Single entrants	0.108	0.598	0.620	0.560	0.294	-0.141	0.297	0.181
Single exiters	0.035	0.450	0.757	0.529	0.317	-0.085	0.224	0.232
Other	0.026	0.481	0.682	0.554	0.367	-0.137	0.221	0.240
Worker Demographics								
Male	0.034	0.498	0.745	0.497	0.311	-0.046	0.214	0.179
Female	0.062	0.394	0.784	0.494	0.340	-0.051	0.189	0.186
Aged (years):								
Under 20	0.240	0.364	0.695	0.486	0.529	-0.159	0.162	0.147
20–29	0.008	0.409	0.703	0.542	0.392	-0.062	0.170	0.181
30–39	0.031	0.325	0.822	0.513	0.346	-0.060	0.161	0.232
40–49	0.037	0.396	0.803	0.494	0.296	-0.074	0.251	0.174
50–59	0.037	0.380	0.808	0.468	0.305	-0.077	0.250	0.141
60–69	0.039	0.314	0.798	0.439	0.361	-0.104	0.215	0.120
Industry								
A Agriculture, Forest and Fishing	0.083	0.474	0.475	0.433	0.404	-0.369	0.074	0.099
B Mining	0.095	0.387	0.712	0.498	0.350	-0.028	0.070	0.071
C Manufacturing	0.045	0.489	0.669	0.425	0.335	-0.155	0.183	0.061
D Electricity, Gas and Water Supply	0.052	0.496	0.789	0.376	0.359	0.055	0.126	0.114
E Construction	0.038	0.501	0.593	0.390	0.354	-0.229	0.112	0.061
F Wholesale Trade	0.030	0.473	0.773	0.460	0.316	-0.042	0.122	0.233
G Retail Trade	0.086	0.655	0.559	0.422	0.345	-0.133	0.235	0.093
H Accomm, Cafes and Restaurants	0.101	0.576	0.509	0.327	0.386	-0.260	0.122	0.046
I Transport and Storage	0.032	0.435	0.700	0.469	0.336	-0.162	0.111	0.172
J Communication Services	0.008	0.515	0.783	0.561	0.344	0.042	0.248	0.301
K Finance and Insurance	0.037	0.551	0.856	0.281	0.352	0.230	0.068	0.121
L Property and Business Services	0.022	0.489	0.801	0.548	0.300	0.038	0.186	0.312
M Government Admin and Defence	0.056	0.472	0.772	0.205	0.340	-0.050	0.045	0.043
N Education	0.046	0.368	0.821	0.135	0.360	-0.074	0.032	-0.012
O Health & Community Services	0.054	0.388	0.841	0.514	0.298	-0.009	0.180	0.291
P Cultural and Recr Services	0.043	0.513	0.750	0.513	0.318	-0.021	0.171	0.249
Q Personal and Other Services	0.032	0.569	0.643	0.624	0.292	-0.107	0.361	0.201
Metro versus Non-metro								
Auckland	0.043	0.518	0.748	0.499	0.312	-0.029	0.256	0.191
Wellington	0.027	0.508	0.775	0.546	0.301	0.005	0.255	0.258
Christchurch	0.053	0.543	0.680	0.474	0.316	-0.094	0.287	0.083
Other	0.052	0.535	0.644	0.502	0.322	-0.142	0.282	0.078
Employment Stability								
Firms:								
Part year, Part-time	0.044	0.513	0.688	0.454	0.336	-0.119	0.225	0.127
Full year, Part-time	0.060	0.457	0.773	0.397	0.299	-0.069	0.155	0.124
Part year, Full-time	0.015	0.418	0.785	0.349	0.338	-0.056	0.028	0.147
Full year, Full-time	0.053	0.435	0.790	0.345	0.302	-0.054	0.084	0.096
Workers:								
Part year, Part-time	0.094	0.495	0.352	0.482	0.421	-0.397	0.199	0.037
Full year, Part-time	0.075	0.506	0.442	0.500	0.357	-0.374	0.218	0.028
Part year, Full-time	0.022	0.365	0.715	0.414	0.417	-0.126	0.138	0.118
Full year, Full-time	0.045	0.388	0.749	0.392	0.176	-0.171	0.154	0.039

Notes: Estimates are based on regressions weighted by FTE employment.

Symbols: … not applicable

Appendix Figure A1

(a) Age Profiles



(b) Time Effects