New Zealand Wage Equations: 1988-2013

Matt Nolan

WORKING PAPER 05/2018
April 2018

Working Papers in Public Finance

VICTORIA
UNIVERSITY OF WELLINGTON
TE WHARE WĀNANGA
O TE ŪPOKO O TE IKA A MĀUI

Chair in Public Finance
Victoria Business School

The Working Papers in Public Finance series is published by the Victoria Business School to disseminate initial research on public finance topics, from economists, accountants, finance, law and tax specialists, to a wider audience. Any opinions and views expressed in these papers are those of the author(s). They should not be attributed to Victoria University of Wellington or the sponsors of the Chair in Public Finance.

Further enquiries to:
The Administrator
Chair in Public Finance
Victoria University of Wellington
PO Box 600
Wellington 6041
New Zealand

Phone: +64-4-463-9656
Email: cpf-info@vuw.ac.nz

Papers in the series can be downloaded from the following website:
http://www.victoria.ac.nz/cpf/working-papers

Working Papers in Public Finance

New Zealand Wage Equations: 1988-2013

Matt Nolan

April 9, 2018

Abstract

This paper reports estimates for wage equations for New Zealand wage and salary earners for four periods: 1987/88-1990/91, 1991/92-1997/98, 2000/01-2007/08, and 2008/09-2012/13. Utilizing data from the Household Economic Survey (HES) between 1987 and 2013 the appropriate data are pooled and separated into five demographic groups (coupled men, coupled women, single men, single women, and single parents), allowing the estimation of five wage equations for each year. Each of these 20 wage equations is tested for selection bias and estimation is adjusted using the Heckman correction where appropriate. Unlike prior estimates of the New Zealand wage equation, there is an allowance for variance in the selection equation.

The main purpose of such an exercise is to impute wage rates for those who are not employed, information that is necessary for future estimates of the subgroups preference between leisure and wage income. However, these estimates are also used to discuss the evolution of the wage equation for varying demographic groups through time. This allows for a quantitative description of how the return to observed characteristics changed for demographic subgroups between 1988 and 2013.

Statistics New Zealand disclaimer

Access to the data used in this study was provided by Statistics New Zealand under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. The results presented in this study are the work of the author, not Statistics NZ.

Acknowledgements

This paper is part of a larger project on 'Improving New Zealand's Tax Policy via International Tax Transfer Model Benchmarking', funded by an Endeavour Research Grant from the Ministry of Business, Innovation and Employment (MBIE) and awarded to the CPF (Professor Norman Gemmell). Access to the data used in this paper was provided by Statistics New Zealand in accordance with security and confidentiality provisions of the Statistics Act 1975. Facilities at the New Zealand Treasury were used when analysing this data. The results presented in this study is the work of the author, not Statistics New Zealand or the New Zealand Treasury. Special thanks goes to Professor John Creedy and Doctor Patrick Nolan who provided essential comments for early versions of this paper.

1 Introduction

This paper reports estimates of the wage equation for a number of demographic groups in New Zealand between 1987/88 and 2012/13. Using Household Economic Survey (HES) data four different periods of pooled data are created, and the wage equation for each demographic year is estimated for each pooled data period. The purpose of estimating these wage equations is to allow for the imputation of wage rates for individuals who are not employed, given that their wage rate is not observed in the data. Once wages for those who are not employed are imputed it is possible to ask *what if* questions about the income the individual would earn if they moved into work.

Thoughout the period between 1988 and 2013 the response of labour supply to changes in tax and transfer policies was seen as centrally important (eg Treasury 1996, Creedy and Mok 2015). As a result, any analysis of changes in incomes or the tax-transfer system during this period requires imputed wage rates for those who were not employed.

The method applied takes into account that there is sample selection bias in the data due to the fact that only the wages of those who are employed are observed (Cameron and Trivedi 2005). A standard way to correct for this is to make use of the Heckman correction, which involves incorporating the probability that an individual is employed based on their characteristics into the wage regression (Creedy et al. 2000).

Imputing wages in a way that accounts for this bias in the New Zealand context has been investigated for the HES92-HES01 period in Kalb and Scutella (2003*b*) and for the HES07-HES11 period in Mercante and Mok (2014*b*). This paper considers wage imputation over four subperiods: HES88-93, HES94-HES98, HES01-HES08, and finally HES09-HES13. The results for these subperiods are then compared, to give an indication of how the relationship may have changed over time.

By comparing the wage equations for these four year groupings, it is possible to discuss how returns to characteristics (education level, age, occupation etc) changed for each demographic subgroup between 1988 and 2013. The wage equations are not strictly comparable due to differences in the independent variables used for each model, and changes in the definitions used for some dummy variables (eg industry, ethnicity). Even so, given that the selected models are relatively similar across years changes in parameter estimates are a useful indication of change in the underlying

wage equation.

However, this paper does more than apply this existing method to a larger set of data. Heteroskedasticity is identified in the selection equation for most of the demographic groups over time, implying that the inverse mills ratio used in the wage equation is not consistently estimated. This paper estimates and corrects for the variance in the selection equations in order to improve estimates of the wage equation, and the corresponding imputed wages, for this period.

In future research the results of this wage imputation will be used to analyse changes in labour supply behaviour[1], and ultimately used to decompose changes in the observed income distribution in New Zealand into policy related and non-policy related effects (Bargain 2010). As a result, this process forms an important step in the analysis of the income distribution in New Zealand.

The paper is organised as follows: Section 2 outlines the data used in the paper. Section 3 introduces the method used to impute wages. Section 4 discusses the wage equations for each pooled year grouping. Section 5 reports the predicted wages from the method and Section 6 comes these wages to previous results from the literature. Section 7 concludes, and the full set of regression results are reported in tables in Section 8.

2 Household Economic Survey Data

The data used for this analysis comes from Statistics New Zealand's Household Economic Survey (HES). The HES is a household survey that has been run annually since the April 1974 year, except for a brief period where it was taken triennially between 1998 and 2006. The HES surveys 5,000 private households[2], with 3,000-4,000 responses of sufficient quality each year. HES years will be denoted using the year when the survey ended. For example HES95 refers to the household survey that took place between April 1994 until March 1995.

The survey collects data on income by source, hours of work, and a variety of household, family, and personal characteristics (eg ethnicity). A key advantage of the survey is that it links familial relationships. As a

[1] In a way consistent with the prior New Zealand research by both Kalb and Scutella (2003*a*) and Mercante and Mok (2014*a*).

[2] Private households exclude institutional households, such as resthomes and prisons.

result this survey allows the construction of economic family units. This allows the estimation of separate wage equations for individuals based on their familial role and to also include variables related to family status (eg partners income, child's age) into the estimation. An economic family unit refers to an adult, their partner[3], and their dependants[4]. Using these familial linkages five key groups are constructed: Coupled men, coupled women, single men (without children), single women (without children), and sole parents. Wage equations are then estimated for each of these groups.

Data regarding the characteristics of individuals was recoded into dummy variables for: Ethnicity, highest education level, region, industry and occupation of primary job, and age of youngest child.

Given these characteristics there are four separate pooled data set for which the wage equation for each demographic group is estimated: HES88-93, HES94-98, HES01-08, and HES09-13. The demographic groups of interest are coupled men, coupled women, single men without dependants, single women without dependants, and single parents.

Two sets of weights are available for the HES data. Household survey weights are provided that relate to the inverse probability of the household being selected for the survey. There are also calibrated survey weights which reweight the sample data to match a set of externally sourced population statistics Statistics New Zealand (2001). These weights are available from Statistics New Zealand from HES01 onwards. Treasury has calculated a set of calibrated weights for the entire time horizon of the survey in this paper which will be utilized here. These weights were originally used and recalibrated to consider counterfactual population distributions in Ball and Creedy (2015).

The estimation, and results provided, below focus on sample means so these weights are not used. This matches the analysis of Kalb and Scutella (2003*b*) and Mercante and Mok (2014*b*), but requires justification. Incorporating calibrated weights allows for inference with regards to the population of interest, and gives a better estimate of the population parameters in the wage equation. However, given the primary purpose of this paper is to imputing data for the sample non-weighted estimation is preferable.

The non-use of weights can be justified in the following way. For a question such as "what is the wage premium for living in Auckland" applying

[3] Where a defacto or married partner exists.

[4] Dependants are children aged below 15, or below 18 and in full-time education.

sample weights gives estimates for the population from the sample.

However, when imputing wage rates the question being asked is "what is available wage rate for individuals in the sample who are not employed" - a question about the sample not the population. By estimating wage equations with weights the imputed wage rates for these individuals uses the assumption that there is no sampling bias for those who are not employed. As a result, if the wage rate relationship for non employed respondents differed to non-respondents this would be a false assumption which would lead to biased results.

In the HES, individuals are asked about income and expenditure for the prior year. As a result the data contained in the entire HES95 survey refers to the period April 1993 until March 1995. The income data used refers to *total* income from all sources. However, the wage figure used in this paper is derived from *current* wage and salary earnings. Here *current* refers to a job that the individual is employed in at the end date of the survey. As a result, the wage figures used from the HES95 survey refer to wage and salary earnings that occurred between April 1994 and March 1995.

The **wage measure** used in this paper is the implied wage from wage and salary income in the HES data. This is equal to total current weekly wage and salary earnings from the individuals primary job divided by the number of hours worked per week in their primary job.[5]

On a few occasions inconsistent/unreliable wage data was removed. The key times this occurred were when the calculated wage rate was less than half the minimum wage, when the wage above $150 was reported, and when a very large income was earned on only one hour of work.

Excluded from the analysis were several groups who are not included in the corresponding labour supply analysis, and so who do not require estimated imputed wage rates. These are:

1. The self employed.
2. Those too young to be in the labour force (aged below 15).

[5] The earnings from prior jobs, self-employed work, and secondary jobs are ignored when calculating the wage. The primary job is defined as the wage and salary job the individual is currently employed in that offers the largest average weekly income at current hours of work.

3. Individuals who have reached retirement age.[6]
4. Individuals who are unable to work (those reporting being on the invalids benefit).
5. Those in full time study.

Data are pooled into four separate year groups, and the wage equation for each family type is estimated separately for each set of pooled years. The pooled years are HES88-93, HES94-98, HES01-08, HES09-13. The reason why these year groups were selected is discussed in the Section 4. The approximate sample size for each demographic group, in each pooled year period, is shown in Table 1.[7]

Table 1: HES sample sizes

	Coupled Men	Coupled Women	Single Men	Single Women	Sole Parent
			Sample size		
HES88-93	7,050	8,470	3,370	2,850	1,790
HES94-98	5,810	6,160	2,260	1,880	1,100
HES01-08	4,890	5,250	1,930	1,680	870
HES09-13	6,860	7,430	2,690	2,630	1,310

There are specific cases where the variables used for analysis differ between these pooled year groups. For the HES94-98 years reported ethnicity was restricted to only take two values – Maori and/or Pacific Islander and Other. As a result, the wage imputation for the HES94-98 period uses only a Maori-Pacific dummy, as compared to the Maori-Pacific only, Maori-Pacific full, and Other non-European dummies used for later periods.

For the estimates provided below there were a variety of different industry, occupation, and education categories used through time. To make the data comparable a correspondence was applied to all the years of analysis, to attempt to replicate the categories that held in the final year of analysis (2013). However, any such correspondence is imperfect and as Statistics New Zealand points out this provides significant comparability issues Statistics New Zealand (2006).

[6] Aged over 65 – therefore wages are still imputed for those who were at retirement age earlier in the time period of interest. A dummy variable is included in the analysis for people who are below 65 but are receiving National Superannuation.

[7] The sample sizes in this table have been randomly adjusted to a neighbouring 10 in line with Statistics New Zealand confidentially requirements.

The industry classification used refers to the ANZSIC06 classification codes [Level 1] Statistics New Zealand (2017*a*). The occupation classification refers to the ANZSCO06 classification codes [Level 1] Statistics New Zealand (2017*b*). The education levels follow from the NZQF qualification levels NZQA (2017).

The wage equations use wage data that is deflated by a wage index to make it comparable between years. As a result, the wage relationship that is discussed and shown in the coefficients indicate the relative importance of characteristics for setting wages - not their absolute impact on the wage level. Furthermore, due to this deflating the wage equations will be comparable between periods.

Outside of the HES data several other data sources from Statistics New Zealand are used. Quarterly Employment Survey (QES) and Earnings and Employment Survey (QEX) data provide an external estimate of average weekly earnings. The Consumer Price Index (CPI) is the measure of consumer prices, in this case we use CPI All Groups plus interest payments. Furthermore, the Household Labour Force Survey (HLFS) gives a measure of unemployment by sex.

3 Method of analysis

The method used to impute wages for those who are out of work is based on that applied by Kalb and Scutella (2003*b*) and Mercante and Mok (2014*b*).

When imputing wages for someone who is out of work, it is not sufficient to take average wage of individuals with the same observed characteristics who are employed as an estimate of the available wage rate. The fact that they are not in employment may suggest that they are different for unobserved reasons. If this is the case, and these unobserved factors are related to the wage a person receives, then the expected wage that person would earn if there were employed will vary from the average of those with the same observed characteristics.

The traditional way of dealing with this issue is to act as if there is an unobserved omitted variable which explains the difference in employment status between people with given observed characteristics.

Viewing the problem in this way, the process that determines whether someone is employed (the selection/employment equation) will be related

to the process that determines the wage available to the individual (the wage equation). As a result, estimating the imputed wage via an OLS regression on only those who are employed will give a biased estimate of the wage available to those who are not employed.

The solution to this is to apply the Heckman correction of Heckman (1979).

Both the employment status and the wage rate of an individual are a product of the individuals characteristics. In the case of employment, assume that there is some underlying propensity for the individual to be in paid employment which is a linear function of the observed characteristics of the individual (z). For the i^{th} individual this can be denoted as:

$$E_i^* = z_i^{'}\gamma + u_i \tag{1}$$

Where u_i is assumed to be independently distributed with $\mathrm{N}(0, 1)$.[8]

The employment index is neither the availability of jobs or individuals willingness to participate in work. Instead it is the probability that an individual receives a wage offer that is greater than their reservation wage - making it a hybrid of both factors. As a result, anything changing the employment index should not be interpreted as solely due to difference in the opportunity or willingness to work of individuals.

This underlying employment index then gives the probability that an individual will be observed in employment. Denoting observed employment status as E, where $E = 1$ is an index variable that represents that the individual is employed. Given this the possible outcomes are:

$$E_i = \begin{cases} 1 & \text{if } E_i^* > 0 \\ 0 & \text{if } E_i^* \leq 0 \end{cases} \tag{2}$$

Given the assumed distribution of u, the probability of being employed is $P(E_i = 1) = \Phi(z_i^{'}\gamma)$ where Φ is the standard normal distribution function. With employment status observed for everyone in the sample, the probability of being employed can be estimated via standard probit regression.

[8] There is no way of identifying the variance of this equation when all we observe is whether $E^* > 0$ or not and not the scale. However, as u_i is positive and γ scales with the variance this is independent of scale. As a result, u_i can be normalised (Cameron and Trivedi 2005)

For the observed data the wage rate (and therefore the wage conditional on being employed) for an individual takes the form:

$$W_i|_{E_i=1} = x_i^{'}\beta + \epsilon_i \tag{3}$$

Where x_i is a vector of characteristics[9], and $\epsilon_i \sim \mathrm{N}(0, \sigma_\epsilon)$.

Given the observable characteristics (x_i) of individuals out of employment, the OLS or censored Tobit estimate appear to be a natural way to impute the wage that is available to those who are not employed ($E = 0$). However, selection into employment and the wage available to the individual are not likely to be independent - namely if the covariance between ϵ and u is non-zero then the OLS estimate of the wage will be biased. This is shown by the conditional expectation of the wage rate given that the person works.

$$\begin{aligned} \mathrm{E}(w_i|E_i^* > 0) &= \mathrm{E}(x_i^{'}\beta + \epsilon_i|E_i > 0) \\ &= x_i^{'}\beta + \mathrm{E}(\epsilon_i|z_i^{'}\gamma + u_i > 0) \end{aligned} \tag{4}$$

As a result, estimates of the wage rate need to account for this bias by estimating $\mathrm{E}(\epsilon_i|z_i^{'}\gamma + u_i > 0)$. This is achieved by using the Heckman correction.

The Heckman correction involves fitting a probit model to the selection equation (equation (2)) to estimate the employment probability for each individual on the basis of their characteristics.

The procedure then incorporates the probability of the given individual being in employment in the wage equation. Given the assumption that the error term from the employment equation (u_i) and the wage equation (ϵ_i) are jointly normally distributed as $\mathrm{N}(0, 0, 1, \sigma_\epsilon, \rho)$ then $\mathrm{E}(\epsilon_i|z_i^{'}\gamma + u_i > 0) = \rho\sigma_\epsilon\lambda_i$.[10] Here ρ is the correlation between unobserved factors that influence the propensity to work and the unobserved factors that influence the wage rate. Furthermore, $\lambda_i(z_i^{'}\gamma) = \frac{\phi(z_i^{'}\gamma)}{\Phi(z_i^{'}\gamma)}$ where $\phi(z_i^{'}\gamma)$ is the standard normal density function evaluated at $z_i^{'}\gamma$ from the estimated probit model. λ is also called the inverse mills ratio (Maddala 1983).

[9] Where this vector of characteristics may in part overlap with the vector of characteristics from the employment equation, z_i.

[10] This comes from the definition of the truncated normal distribution, the proof is shown in Mercante and Mok (2014*b*).

In order to calculate wages β, ρ and σ_ϵ need to be estimated. By performing an OLS regression on the wage rate for those who are employed that includes the vector of λ_i's, λ, $\hat{\beta}$ and $\rho\hat{\sigma}_\epsilon$ can be estimated.[11] As a result, ρ can be recovered by dividing the estimated coefficient of λ by the estimate of $\hat{\sigma}_\epsilon$

The error term from this OLS regression ($\hat{v}$) does not give an unbiased estimate of σ_ϵ, as observed wages only provide a truncated distribution of wage rates. However, using the truncated variance formula of a left truncated normal distribution and where N is the size of the sample gives:

$$\hat{\sigma_\epsilon} = N^{-1} \sum_{i}^{N} [\hat{v_i}^2 + \widehat{\rho\sigma_\epsilon} \hat{\lambda_i} (x_i^{'} \hat{\beta} + \hat{\lambda_i})] \tag{5}$$

Give this error term, and with $\mathrm{E}(\epsilon_i | z_i^{'}\gamma + u_i < 0) = -\rho\sigma_\epsilon \frac{\phi(z_i^{'}\gamma)}{1-\Phi(z_i^{'}\gamma)}$ the conditional wage estimates can be written as:

$$\mathrm{E}(w_i | E_i = 1) = x_i^{'}\beta + \rho\sigma_\epsilon\lambda_i \tag{6}$$

$$\mathrm{E}(w_i | E_i = 0) = x_i^{'}\beta - \rho\sigma_\epsilon \frac{\phi(z_i^{'}\gamma)}{1 - \Phi(z_i^{'}\gamma)} \tag{7}$$

As a result, equation (7) can be used to impute the wage of individuals who are out of work.

This is a clear way of viewing the problem, and is also the general method used in this paper for imputing wages. However, it is not the method used in Kalb and Scutella (2003*a*) and Mercante and Mok (2014*a*), as it is more efficient to solve equations (1) and (3) simultaneously via maximum likelihood (Cameron and Trivedi 2005).

Woldridge (2010) points out that the key difference in assumptions between the two forms of estimation is that non-linear restrictions are imposed automatically on the structural parameters when the simultaneous estimation is used. Generally this is seen as a reasonable cost for the increase in estimation efficiency.

[11] This is the value of $\sigma_{\epsilon u}$, or the covariance between the wage and employment equations, by definition.

However, in this paper the two-stage Heckman method is used due to concerns about heteroskedasticity in the data and the clarity the two-stage method gives for dealing with this issue.

Given the similarity of this method to instrumental variable estimation, an exclusion restriction is usually recommended. The household income variable provides this restriction, as it is included in the selection equation but not the wage equation. For couples dummy variables related to children and partners are also only in the selection equation.

The estimates undertaken below are performed separately for different subsets of the data depending on sex and family type: Coupled Males (with or without children), Coupled Females (with or without children), Single Males (without children), Single Females (without children), and Sole Parents (Male or Female).

3.1 Industry and Occupation imputation for those out of work

Using the Heckman correction the wage estimates that are derived are a function of a number of observed characteristics. All of these characteristics are observed for people in employment, but some of the characteristics are not observed for those who are not working.

Specifically, the HES data collects information on the industry and occupation of an individual's job. However, if someone is not working no industry or occupation is collected. As a result, in order to impute a wage for someone who is out of work the industry and occupation of that individual also needs to be imputed.

The solution used in Kalb and Scutella (2003*b*) and Mercante and Mok (2014*b*) is to construct composite industry and occupation variables for the unemployed. These papers take HLFS survey data that captures the proportion of unemployed people who desire working in a given industry/occupation. This method of creating the composite industry via external data is recommended by Creedy et al. (2000).

In this paper the proportion of employed people of that demographic group who are employed in that industry/occupation is used as the industry/occupation variable for imputing wages. The HLFS data was not used as it was not available for the entire time period of interest, and also because its quality as an indicator of job take up is dubious. The vast

majority of respondents to the survey did not give a preferred industry or occupation, suggesting that the real distribution also differs from this external data.

With unemployment/non-participation tending to be higher among those who would receive or accept wage offers from lower waged industries/occupations this form of imputation biases the wage estimate for those out of work upwards slightly.

3.2 Correcting for heteroskedasticity in selection

Whether estimating the two-step method or the bivariate sample selection model, a probit model is used to estimate the probability of being employed given the individuals characteristics. These estimates could be inconsistent if there is heteroskedasticity in the error term of the employment equation. Woldridge (2010)[12]

In Creedy et al. (2000) the authors tested for heteroskedasticity in the selection equations for each group, and found that homoscedasticity was valid in each case. As a result, no adjustment was necessary. However, when performing specification tests on the New Zealand data heteroskedasticity was found in the majority of the selection equations.

To test for this potential source of bias a likelihood ratio test was performed for the difference between the parameters in the estimated employment model with heteroskedasticity adjusted standard errors and the model without this adjustment. If the parameter values differ in a statistically significant way, then the model likely suffers from heteroskedasticity.[13]

In order to make estimation of the selection equation consistent it is necessary to incorporate the form of the heteroskedasticity into the likelihood function.[14] No matter how large the dataset is, this is an issue for this modelling framework (Greene 2011).

[12] This differs from the fact the wage equation is heteroskedastic, as the wage equation is linear and is heteroskedastic by design

[13] The test performed is a chi^2 test comparing the hetprob estimate and the het estimate in Stata

[14] This can be done automatically with the hetprob function in Stata in a conceptually similar way to weighted least squares using the set of variables that are believed to be causing the heteroskedasticity (Wooldridge 2014)

To deal with the issue of heteroskedasticity the Heckman correction is used as defined above, but the initial probit equation estimated with using a heteroskedastic probit model as suggested by Gould (1998).

The hetprob command in Stata allows the estimation of a multiplicative form for the variance (Harvey 1976). Take a variance function $\sigma_v^2 = \{\exp(g_i\theta)\}^2$, where g_i is a vector of dependent variables and θ are the parameters. A multiplicative form of the variance assumes that the probability of employment can now be estimated as:

$$P(E_i = 1) = \Phi\left(\frac{z_i^{'}\gamma}{\sigma_v}\right) \tag{8}$$

These corrected estimates of the selection equation lead to new estimates of the inverse mills ratio, which is then fed into the linear estimator of the wage equation. On average this process led to a smaller inverse mills ratios than the estimated model without taking into account heteroskedasticity, but this was not always the case with single men in the HES94-98 data reporting a larger inverse mills ratio after adjustment.[15]

Estimated variance functions are given with the results.

4 Summary of wage equation results

The estimation of the above wage equations gives an estimate of how certain observed individual characteristics are associated with the wage rate an individual earns.

Care needs to be taken when interpreting the table of coefficients, as is discussed in the Section 8. Below is a brief non-technical description of how different variables are related to the wage rate in the data and how this evolved through time.

The reason why certain periods were selected for a given wage and employment equation will also be discussed. The general goal when selecting the year groupings was to create periods with a similar amount of data, similar data reporting/definitions, and corresponding to a similar point in the economic cycle.

[15] A priori we know that there is a bias due to heteroskedasticity in the error term of the selection equation, but we do not know what sign this bias would take.

4.1 HES88-93

4.1.1 Reason for year selection

The period from April 1987 to March 1993 was a period of economic turmoil in New Zealand - with falling household incomes, climbing unemployment, and rapid structural change.

Annual GDP was only 0.3% higher in the March 1993 year than it was in March 1988 while employment levels had fallen 6.1%. The unemployment rate climbed from 4.2% to nearly 11% during this period - and was starting a gradual decline in the year to March 1993. As a result, the economic and labour market conditions during this five year period were significantly different from those that were experience at other points in time in New Zealand.

4.1.2 Wage equation

For single people and couples the wage tended to rise with age (excluding any interaction with education level), albeit at a decreasing rate. The age premium peaks for those in their early 40's for the majority of the sample, consistent with Kalb and Scutella (2003*b*). The peak was relatively consistent across most demographic groups, occurring at between 42 and 43 for everyone except single parents - whose age wage premium peaked in their mid-60s. However, the age relationship for single parents in this sample is statistically insignificant so the general age relationship for this demographic group in this time period is uncertain.

In terms of ethnicity, there is a significant wage penalty non-Europeans during this time period for all demographic groups except single parents. Higher levels of education also led to a higher wage for all demographic groups.

Both industry and occupation, the correspondences used for this period matched poorly, so any results must be treated with caution. Given this warning, the results are reported below.

Conditional on occupation and other characteristics, industry wages vary in a relatively consistent way by demographic group. The three industries offering statistically significantly higher wages than the reference industry across all demographic group (agriculture) were the transport, communications, and financial services industry. Excluding sole parents, manu-

facturing, construction, wholesale trade, professional services, and public administration also offered a premium. However, the retail and accommodation industries offered negative to no wage premium across demographic types. The largest wage premiums in the HES88-93 period were found for the communications and financial service industries.

Relative to the reference occupation (labourers) managers and professionals offer the largest wage premium. However, every occupation except sales offered higher wages across demographic groups during this period.

The inverse mills ratio was negative for single parents during this period. A negative ratio implies that, for given observed characteristics, someone who is out of work has a higher available wage than someone who is working. Although this is possible, it is generally not seen as credible. In this instance the inverse mills ratio was not statistically different from zero and as a result this suggests that the existence of selection bias is rejected for single parents is rejected during this period.

4.1.3 Marginal effects

Following Kalb and Scutella (2003*b*) and Mercante and Mok (2014*b*), the wage premium for a number of characteristics for each subgroup is reported for a representative member of that group. Specifically, the premium for vocational education relative no formal education[16] and residing in the North of the North Island[17] relative to Auckland are reported for a 30 year old European. Finally the premium for various age groups are reported.

Having a vocational qualification[18] offered a significant wage premium for all demographic subgroups. Coupled Women, Single Men, and Single Women all had an estimated wage premium over 20% (25%, 23%, and 20%

[16] In previous papers the comparison was with postgraduate education. Relative to these papers the results of this paper showed a slightly higher wage premium for all groups, but the result that sole parents had a smaller premium remained. However, in HES94-98 the Single Male estimates grouped together all graduate and post-graduate education. As a result, vocational education has been selected as a consistent benchmark between the years.

[17] This includes Northland, Waikato, Bay of Plenty, and Gisborne.

[18] A vocational qualification is when an individual has a highest qualification that is between Level 4 and Level 6 in the NZQA Qualifications Framework

respectively). However, Coupled Men and Single Parent also reported double digit premiums (17% and 14% respectively).

All demographic subgroups except Single Parent reported a wage penalty for living outside of Auckland. Three of the groups had an estimated penalty that was statistically significant, Coupled Women (8.7%), Single Women (8.5%), and Coupled Men (6.8%).

Apart from Single Parents (whose age terms were statistically insignificant), the wage increase for a European with no formal education between the ages of 25 and 35 was substantial. Wage rates for Single Women and Men increase by over two-thirds (rising by 74% and 68% respectively) while Couple Men wages increase by a quarter. The estimated increase in Coupled Women wages was more modest (8.3%).

However, the wage effect of age is non-linear. By the age of 45 wages were higher for each demographic subgroup, but the rate of increase between 35 and 45 was significantly lower than between 25 and 35. The three largest increases were for Single Women, Single Men, and Coupled Men (up by 13%, 10%, and 6.5% respectively).

4.2 HES94-98

4.2.1 Reason for year selection

The second period selected refers to the data between HES94 and HES98. As a result, this includes survey data from April 1993 to March 1998.

This period was selected to coincide with the strong lift in economic activity in New Zealand from the March 1994 year. Over the March 1994 year, annual GDP rose by 6.4%. Activity increased by a further 16% (an average of 3.9%pa) between 1994 and 1998.

Employment rose alongside economic activity - increasing 12% between the March 1994 and March 1998 years. This saw the unemployment rate fall from 10% to 6.4% in 1997 - although this figure began to creep back up as the Asian Financial Crisis took hold in 1997/98.

4.2.2 Wage equation

The age relationship to wages remained similar to HES88-93 during this period, with the average premium rising slightly for Couples and Single Men. For Single Parents the age structure switched to being more closely in line with other groups, with the average peak earning age in the early 40s. The peak age among Single Women declined in this period. The peak wage age for all Coupled Men and Single Women shift out towards their mid-late 40s, while the peak stayed in the early 40s for Coupled Women and Single Men.

The ethnicity definition used in HES94-98 differed from that in HES88-93. Previously ethnicity was split into three categories: Maori/Pacific, European, and Other. In this data the split is only into Maori/Pacific and Other. As a result, the ethnicity figures are not strictly comparable. For this new split there was a clear wage penalty for the Maori/Pacific subgroup amongst all demographic groupings.

The returns to additional education remained positive in the HES94-98 period. Furthermore, estimated returns rose for single people while staying broadly unchanged for couples.

By industry communication and financial services remained the two industries with the largest wage premiums, followed by professional services and public administration.

Relative to agricultural work there remained a wage penalty for working in the retail and accommodation industries for coupled men and women. The education industry (introduced from the HES94-98 data set) also saw a wage penalty for all demographic types.

The large wage premium for managerial, professional, and personal service work remained in place in HES94-98. However, all occupations except machinery operating reported a wage premium above labouring during this period.

The inverse mills ratio for single men was substantially larger than estimates for other periods, or for other demographic groups, in this wage equation. This was due to a high estimate for the correlation between unobserved covariates in the wage and selection equations, suggesting that an underlying determinant of selection was also influencing the wage available to Single Men. As this was a period with historically very high employment growth for this group this temporary change in the wage re-

lationship is plausible.

4.2.3 Marginal effects

The wage premium associated with having a vocational qualification remained sizeable in HES94-98. Although only single men had an estimated premium over 20% (29%), Single Parents, Coupled Women, and Single Women were all close to this threshold (with premiums of 19%, 19%, and 17% respectively). The Coupled Men premium declined from HES88-93 but was still 14%.

There was a wage penalty for living outside Auckland remained for all demographic groups in HES94-98. For Coupled Women, Couple Men, and Single Women the penalty declined from HES88-93 (to 6.1%, 4.2%, and 3.9% respectively) while the Single Parent penalty (6.9%) was statistically insignificant. However, the Single Men penalty rose sharply to 12%.

A ten year increase in age from 25-35 lead to a double-digit percentage lift in wages for all demographic subgroups in HES94-98. Single Men reported by far the largest increase (84%), while the estimated increase Single Women, Coupled Men, and Single Parent was still over a quarter (38%, 35%, and 29% respectively). The wage increase was smallest for Coupled Women (14%) but was still higher than the estimated increase in HES88-93.

As in HES88-93, the wage increase between 35 and 45 was much lower than during the prior ten year age increase. However, the estimated wage premium for HES94-98 was slightly higher for all demographic subgroups than it had been in HES88-93.

4.3 HES01-08

4.3.1 Reason for year selection

The third period selected refers to the data between HES01 and HES08. As a result, this includes survey data from July 2000 to June 2008. Although the HES data was taken for a specific June year, there were not annual surveys during this period. As a result, the sample includes data from the HES01, HES04, HES07, and HES08 years.

This period was selected to coincide with the strong lift in economic activity in New Zealand following the Asian Financial Crisis. This is a longer time period than our other datasets, largely as a result of the sparse data available for this period. However, this also corresponds to an unusually long period of sustained economic growth - with GDP rising by at least 2.8% each year between March 2001 and March 2008.

Strong growth in economic activity was matched by an average 2.6%pa increase in employment during this period. This saw the unemployment rate decline from 5.9% in 2001 to 3.6% in 2008 - while labour force participation rate rose to historically high levels.

4.3.2 Wage equation

The age premiums generally rose during the HES01-08 period. The age effect on wage peaked all demographic groups in their mid-40s.

Once again the ethnicity definition changed from the previous estimates. For the HES01-08 period people could report multiple ethnicities. As a result, the categories are set in line with the analysis used by Mercante and Mok (2014*b*) to allow for comparability. The four groups that are defined are: Maori/Pacific (some)[19], Maori/Pacific (only)[20], European, Other. The reference group in this case is European.

For all demographic groups, those that responded with Maori/Pacific (only) or Other had lower wages than the reference group. However, the results for Maori/Pacific (some) are more mixed.

Returns to education remained positive in the HES01-08 year, but the level of the wage premium for couples declined slightly from its HES94-98 level.

Industry specific wage premiums remained large for financial services, professional services, and public administration during the HES01-08 period. However, among all groups the wage premium for working in the communications industry declined sharply. Wage premiums in the construction industry rose markedly during this period. The wage penalty for working in the retail and accommodation sectors held for both singles

[19] When an individual responds with a Maori/Pacific ethnicity and another ethnicity in their HES response

[20] When an individual responds with only Maori/Pacific ethnicities in their HES response

and couples during the HES01-08 period. However, the penalty for working in the education industry became statistically insignificant during this time.

Managerial and professional position continued to demand a significant wage premium in the HES01-08 period. The previous large premium for working in professional services largely disappeared for all demographic groups during this time. The wage premium (relative to working in a labouring job) for women in managerial and professional roles was significantly larger than for men during this period.

4.3.3 Marginal effects

The wage premium for vocational qualification dropped significantly for all groups except Coupled Men between the HES94-98 and HES01-08 estimates, with all other demographic subgroups reporting a premium between 11% and 13%.

Even the increase in the wage premium for Couple Men (to 18%) was conditional on age, due to the premium declining with the age of the individual. By around 50 years of age the wage premium for vocational work was equivalent to that reported in the HES94-98 results.

The wage penalty for living outside Auckland increased in HES01-08, with a double-digit estimated percentage penalty for all demographic subgroups. The largest estimated penalty was for Single Women at 14%. Single Men were the only group who reported a lower penalty for living outside Auckland than in HES94-98 at 10%.

In HES01-08, all demographic subgroups reported a sizeable wage premium for a ten year increase in age from 25 to 35. Although the Single Men premium declined from the 1990s it remained the largest premium at 45%. All subgroups had an estimated premium above 20%, with Coupled Men also above 40% (43%).

A smaller wage premium for a ten year increase in age from 35 to 45 persisted in HES01-08. Single Men, Single Parent, Coupled Men, and Coupled Women all had a higher estimated premium for this age increase than during the 1990s (17%, 17%, 16%, and 8.8% respectively). However, the Single Women premium declined to 7.5%.

4.4 HES09-13

4.4.1 Reason for year selection

The final period chosen consists of survey data between HES09 and HES13. As a result, this refers to survey results from between July 2008 and June 2013. The final period captures the Global Financial Crisis, Canterbury Earthquakes, and the start of the recovery – although the labour market stayed persistently weak according to survey and unemployment measures during this period.

GDP declined 1.1% in the March 2009 year as the combination of drought, tight monetary conditions, and the effects of the Global Financial Crisis took their tool. With global uncertainty persisting, and the 2010 and 2011 Canterbury Earthquakes taking place, the economy only grew at an average annual rate of 1.4%pa between March 2009 and March 2013.

The employment situation was even relatively worse. In March 2013 was only 0.3% above its March 2009 level - a similarly anaemic performance when compared to employment growth in our first period of interest in HES88-93. However, this did not come with quite the same sharp lift in the unemployment rate - with the annual unemployment rate peaking at 6.3% in the March 2013 year.[21] A mixture of net migration outflows, a reduction in labour force participation (from record high levels), and a relatively smaller decline in employment when the recession kicked off during the March 2009 year, had prevented a sharper lift in unemployment.

4.4.2 Wage equation

The age premium declined for all age groups in the HES09-13 period. The age where wages peaked continued to move out for most groups, varying between 45 and 50. The slight increase in the age where average wages peaked between 1988 and 2013 is consistent with the findings of Mercante and Mok (2014*b*). The wage penalty for Maori/Pacific (only) and Other ethnicity individuals persisted into the HES09-13 period.

The average wage premium for educational attainment remained at a similar level to prior years, but had become more strongly dependent on the

[21] Quarterly seasonally adjusted unemployment peaked at 6.7% in September 2012 - however, this figure appeared to be somewhat of an outlier and so annual averages are used to smooth out the result.

age of the individual than it had previously. In prior years an increasing educational wage premium with age had been insignificant, but in HES09-13 this relationship was strong and positive for all groups.

By industry, the wage premiums remained strongest in financial services and public administration. The wage premium for professional services remained similar to HES01-08, with a larger premium for coupled people than singles. Communication industry wage premiums fell for most groups, but remained high for coupled men. The retail and accommodation industries continued to report a wage penalty for all demographic groups in HES09-13.

In terms of occupations, managerial and professional roles continued to pay a significant premium. The reduction in the personal services wage premium continued into HES09-13, falling to near zero for all groups except Single Male. The wage premium (relative to working in a labouring job) for women in managerial and professional roles remained significantly larger than for men during this period.

The inverse mills ratio was negative for single parents during this period. A negative ratio implies that, for given observed characteristics, someone who is out of work has a higher available wage than someone who is working. Although this is possible, it is generally not seen as credible. In this instance the inverse mills ratio was not statistically different from zero and as a result this suggests that the existence of selection bias is rejected for Single Parent is rejected during this period.

4.4.3 Marginal effects

Apart from Single Parent (whose estimated premium fell below 10%), the wage premium for a vocational qualification rose for each demographic subgroup in the HES09-13 estimates. Although the premium rose to its highest level of any time period for Coupled Men (reaching 19%), the premium remained below its 1990s level for Coupled Women, Single Men, and Single Women (at 17%, 15%, and 12% respectively).

In HES09-13, the wage penalty for living outside Auckland declined from HES01-08 for all demographic subgroups. The penalty was relatively consistent across subgroups, with the largest penalty reported for Coupled Men (9.2%) and Single Men and Parent with the lowest estimated penalty (6.3%).

The age premium for then ten year increases from 25 and from 35 declined for all demographic subgroups in HES09-13.

The wage premium for Coupled Men and Women declined back to the level recorded in HES94-98. However, for single people the premium fell to its lowest recorded level. Single Women and Men had an estimated premium of 23% and 21% respectively for the ten year increase between 25 and 35, with a further 8.2% and 11% increase for the lift in age from 35 to 45. The wage premium for Single Parent was statistically insignificant (5.3% between 25 and 35).

5 Predicted wages

As well as providing an estimate of how certain observed individual characteristics are associated with the wage rate an individual earns the method described in Section 3 allows the imputation of wage rates for individuals who are not employed based upon their characteristics via equation (7). However, the conditional wage equations of (6) and (7) tend to give a distribution that is significantly narrower than the observed distribution of wages. As a result, a *random disturbance term* is added to reincorporate unobserved heterogeniety into individuals wage levels.

The addition of this random disturbance comes from the assumed relationship between the error terms from the employment equation (u_i) and from the wage equation (ϵ_i). Given estimated values for ρ and σ_ϵ take a random draw from this distribution and add it to the deterministic component of the wage equation calculated above. For someone in employment, with a random draw υ_i this gives the new wage equation as:

$$\mathrm{E}(w_i|E_i = 1) = x_i^{'}\beta + \rho\sigma_\epsilon\lambda_i + \upsilon_i \tag{9}$$

However, these error terms are also conditional. If someone is observed in employment, then a random draw of ϵ_i cannot be selected as υ_i if it implies that $u_i < z_i^{'}\gamma$ (or that $E_i < 0$). As a result, only error draws that continue to predict the observed employment status of the individual are accepted.

Even after adding this random disturbance back into the the wage prediction, the estimated wage distribution does remain narrower than the observed distribution for those in work.

Given this method, estimates for the average predicted wage condition on employment status can be made for each of the pooled years. The predicted wages are then deflated by the CPI to allow for comparison over time. These results are shown in Table 2.

Table 2: Average predicted wage (2008 prices)

	Coupled Men	Coupled Women	Single Men	Single Women	Sole Parent
Employed observed					
HES88-93	$21.81	$17.34	$16.43	$15.78	$17.52
HES94-98	$21.56	$17.66	$16.29	$15.76	$16.84
HES01-08	$24.25	$20.57	$18.67	$18.09	$18.37
HES09-13	$28.68	$24.28	$20.82	$20.76	$22.20
Employed predicted					
HES88-93	$21.62	$17.24	$16.21	$15.65	$17.34
HES94-98	$21.32	$17.45	$16.17	$15.63	$16.64
HES01-08	$23.83	$20.23	$18.22	$17.98	$18.24
HES09-13	$28.06	$23.83	$20.42	$20.40	$21.92
Not employed					
HES88-93	$15.58	$13.47	$13.71	$12.83	$15.01
HES94-98	$15.98	$13.34	$7.03	$13.77	$12.69
HES01-08	$15.38	$15.14	$14.36	$14.09	$10.69
HES09-13	$22.44	$17.08	$17.02	$17.98	$19.65

6 Comparison to previous estimates

Relative to Kalb and Scutella (2003*b*) and Mercante and Mok (2014*b*) all three sets of results reported above use the same underlying raw data and the same estimation techniques. As a result, we would expect our estimates to differ in these models only due to:

- The difference in the time periods used for estimating each individual model – this is especially important if the relationship genuinely varies over time.
- Differences in the model selected.
- Differences in how the data was deflated.
- Differences in the treatment of the raw data.
- Differences in the occupation and industry placement of those who are not employed.

The estimates reported in Kalb and Scutella (2003*b*) are deflated by an average wage index to the December 2001 quarter, while the estimates in Mercante and Mok (2014*b*) are deflated by an average wage index to the December 2011 quarter.[22] Similarly the estimates for each of the wage equations in this paper is deflated by a wage index to make the data within that set of pooled years comparable for estimation.

When comparing estimates from different time periods, the real wage figures in Table 2 are calculated by transforming the estimated wages back to nominal wages, then deflating the estimated nominal wage series by CPI to the June 2006 quarter.[23] Deflating the results in Kalb and Scutella (2003*b*) and Mercante and Mok (2014*b*) by CPI to June 2006 prices gives the following comparison table of average predicted wages:

Table 3: Average predicted wage (2008 prices)

	Coupled Men	Coupled Women	Single Men	Single Women	Sole Parent
Employed observed					
K&S (HES92-01)	\$20.78	\$17.43	\$15.52	\$13.77	\$15.93
Paper 92-01 avg	\$21.96	\$17.94	\$16.62	\$16.06	\$17.20
M&M (HES07-11)	\$28.29	\$23.55	\$20.63	\$20.85	\$21.82
Paper 07-11 avg	\$26.91	\$22.80	\$19.96	\$19.69	\$20.67
Employed					
K&S (HES92-01)	\$20.81	\$17.09	\$15.63	\$15.50	\$15.98
Paper 92-01 avg	\$21.71	\$17.75	\$16.44	\$15.93	\$17.02
M&M (HES07-11)	\$27.90	\$23.35	\$20.54	\$20.59	\$21.52
Paper 07-11 avg	\$26.37	\$22.39	\$19.54	\$19.43	\$20.45
Not employed					
K&S (HES92-01)	\$13.99	\$12.62	\$11.23	\$11.01	\$15.09
Paper 92-01 avg	\$15.81	\$13.60	\$9.62	\$13.58	\$13.02
M&M (HES07-11)	\$17.39	\$19.00	\$22.42	\$13.16	\$11.57
Paper 07-11 avg	\$19.62	\$16.30	\$15.96	\$16.42	\$16.07

When it comes to the observed data there were slight differences. Both prior studies used the same definition for calculating the wage rate using current labour income. However, the way the data has been deflated creates a difference between the historic series. In this paper the reported figures are taken from raw data at the date is is reported. It is this data that has been deflated by CPI to work out real wage figures in June 2006 prices.

[22] When estimating the models parameters, all three papers deflate wages by a wage index that is set to 1 at the quarters mentioned.

[23] In terms of the parameter values presented later, wages were deflated by average weekly earnings to the June 2006 quarter for all estimates - making the parameters comparable.

The figures this paper reports for the Mercante and Mok (2014*b*) result rely on averages from their paper, which were uprated by average weekly earnings to the December 2011 level. As earnings growth exceeded growth in consumer prices, this will push up the average value wages.[24] Similarly, the [11] data was uprated by average earnings growth to the December 2001 level, which will push down the average value of wages.[25]

However, even given this there is a significant gap between the observed single woman average wage used in this paper and the value from [11]. This gap is all the more surprising given that the imputed wages in both papers are fairly similar, and both papers work off the same underlying dataset.

There are also a smaller difference in the observed sole parent wage in the two papers. Here the definition of sole parent differed slightly in this paper to [11]. A sole parent in this analysis included anyone who did not report a partner in the same household - as a result, people saying they were in a defacto relationship with someone who was not in the household were coded as sole parents for this analysis. The reasoning for this is two-fold: Firstly, if the partner is not in the household it is less likely the individual will be supported in the standard way we would expect with an economic family unit. Secondly, if the partner is not in the household it is likely that the individual will be able to claim sole parent support benefits in the case where their work history allows it.

Relative to the Kalb and Scutella (2003*b*) results, the wage available to couples and single women who are out of work is slightly higher than can be explained by differences in the data. However, single men and single parent imputed wages are significantly lower.

The lower imputed wage for single men is predominantly due to a much higher inverse mills ratio in the HES94-98 period in this paper, due to a strong correlation between the unobserved elements of the selection and wage equations after the model is corrected for heteroskedasticity.

For most of the 1992-2001 period the estimates in this paper impute wages for single parents using a positive inverse mills ratio. However, in Kalb and Scutella (2003*b*) a negative inverse mills ratio was reported. This difference explains the fact that the imputed wage for single parents in this

[24] During this period the CPI rose 11.9% while average weekly earnings from the QEX rose 20.6%.

[25] During this period the CPI rose 12.4% while average weekly earnings from the QEX rose 17.7%.

paper is lower during this period.

Compared to the Mercante and Mok (2014*b*) estimates, couple men, single women, and sole parents have a higher imputed wage while coupled women and single men have a lower imputed wage. The results in this paper have a relatively more stable inverse mills ratio between groups than the Mercante and Mok (2014*b*), as a result the ratio of the wage available to employed vs not employed individuals is more stable.

Furthermore, even though a negative inverse mills ratio is recorded for sole parents during this period the available wage for this group is still lower given the characteristics of the individuals who are out of employment - as compared to the single male result in Mercante and Mok (Mercante and Mok) where the average wage rate is higher for those that are out of work.

If heteroskedasticity was not corrected for in these estimates then similar extreme swings in the inverse mills ratio were found for the data used in this paper. This implies that the relative stability of the estimates found in this paper come from correcting for heteroskedasticity in the selection equation which existed in the dataset.

7 Conclusion

In this paper wage equations for five demographic groups were estimated for four separate sets of years with the goal of inputing the wage rate available to people who were out of work between 1988 and 2013. These estimates corrected for selection bias and allow for heteroskedasticity in the selection equation thereby providing improved estimates of the imputed wage for this period.

Relative to prior estimates for New Zealand imputed wages were higher for coupled men and single women and lower for single men. The implied growth in wages available for not employed coupled women was softer than previously found, while the estimated decline in available wages for not employed single parents was reversed.

Additionally, this exercise outlines how the wage generating process changed between 1988 and 2013.

The age premium for all demographic groups rose and then fell for all demographic over the 1988-2013 period. The peak premium was roughly un-

changed for coupled individuals, but declined for single men and women. The age associated with the largest wage premium rose over time for all demographic groups, increasing from the early 40s to the late 40s.

Financial services, professional services and public administration paid a wage premium above agricultural work throughout the period of interest. However, the premium in the communication industry declined sharply. A wage penalty was found in the retail and accommodation industries across time.

Managerial and professional role had a consistent wage premium across demographic groups over time. However, couple and single women saw their wage premium for these roles rise during the 2000s. The large wage premium that existed for personal services at the start of this time period had all but disappeared by HES09-13.

8 Tables

8.1 Reading the tables

In the estimation process we solve for two variables - the natural logarithm of hourly wages and employment status. However, when evaluating the parameters it is important to keep in mind that there are interaction terms for age, qualification, and ethnicity. As a result, the parameters do not generally refer to the estimated marginal effect on the dependent variable on the independent variable.

When comparing the results, and any interpretation regarding the impact of a given variable (eg qualification) on wages it is important to keep in mind the importance of interaction terms.

For example, in both the following results and prior estimates performed by Kalb and Scutella (2003*b*) and Mercante and Mok (2014*b*) the coefficient on high levels of education (NCEA Level 7 and above) is often negative. However, this does not in of itself suggest that high levels of education lead to lower wages – as there are interaction terms which also capture the return to education in association with other variables (eg age) .

What does this mean? Take into account a qualification levels interaction with age. Generally, the return to education will rise as a person ages given that:

1. Age is a proxy for experience and their additional returns associated with the combination of education and experience.
2. People who have invested in education have sacrificed several years of work. Assuming that wages rise at a diminishing rate with years of experience this experience gap implies a similar pattern would be observed.

In this way, the combination of education + education*age represents the non-linearity inherent in the returns to education – and the parameters alone be interpreted sensibly individually.

Furthermore age is always positive, and so if a qualification level was positively associated with wages and everyone was the same age then we would have perfect multicollinearity. Although age varies, it is possible that the coefficient on a given qualification level will be poorly identified due to this relationship and so the individual parameters should not be used in isolation.

More generally, when considering what a variable represents it is important to remember the change in wages associated with it is conditional on all the other variables in the model. For example, the change in wages associated with a degree are conditional on the industry and occupation worked. However, in order to work in a given industry and occupation the individual may require that level of education - as a result we cannot always look at these parameter estimates in isolation.

Relative to previous literature this paper makes less use of interaction terms, as they tended to significantly increase standard errors of the parameters precision without significantly improving the fit of the model indicating a risk of overfitting.

Finally note that the R^2 of the wage equations and the pseudo-R^2[26] of the selection equations are relatively low, while the assumptions about the error terms required to causally interpret the model are likely to be unrealistic.[27]

Due to the limitations to interpretation, the main purpose of the estimation

[26] The McFadden pseudo-R^2 is used as a measure of goodness of fit for the probit selection model McFadden (1974). A common rule of thumb for such models is that a value between 0.2 and 0.4 suggests good model fit.

[27] In prior papers only a pseudo-R^2, as the wage and selection equations are estimated simultaneously. The process used in this paper is sequential, and so the R^2 refers to the wage equation while the pseudo-R^2 refers to only the selection equation - and so is not comparable to those from earlier New Zealand wage equation estimates.

exercise undertaken in this paper is to impute wages for those who are out of work, not to provide estimates of how certain factors determine the wage rate.

8.2 HES88-93

Table 4: Single people without children: HES88-93

	Wage Equation				Selection Equation			
	Single Men		Single Women		Single Men		Single Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.681	0.000	0.696	0.000	0.390	0.006	1.175	0.000
$Age^2 \div 100$	-0.082	0.000	-0.083	0.000	-0.058	0.004	-0.175	0.000
Highest Qualification attained (reference group less than NZCA level 1 completion)								
Level 1-3	0.079	0.000	0.128	0.000	0.431	0.000	0.596	0.000
Level 4-6	0.203	0.000	0.185	0.000	0.542	0.000	0.696	0.000
Level 7	0.316	0.000	0.286	0.000	0.622	0.000	0.518	0.000
Postgraduate	0.385	0.000	0.337	0.000	0.558	0.004	0.756	0.001
Other	0.249	0.000	0.185	0.000	0.384	0.071	0.352	0.078
Ethnicity (reference group European)								
Maori/Pacific	-0.034	0.174	-0.104	0.000	-0.500	0.000	-0.544	0.000
Other Ethnicity	-0.066	0.059	-0.086	0.036	-0.090	0.489	-0.355	0.026
Region (reference group is Auckland)								
North North Island	-0.031	0.159	-0.089	0.000	-0.269	0.000	-0.162	0.051
Central North Island	-0.092	0.000	-0.100	0.000	-0.202	0.012	-0.093	0.307
Wellington	0.061	0.004	0.026	0.135	0.194	0.041	0.235	0.011
Canterbury	-0.031	0.178	-0.069	0.003	0.148	0.104	0.000	1.000
Rest of South Island	-0.067	0.001	-0.090	0.000	-0.099	0.199	-0.190	0.033
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	-0.126	0.001	-0.111	0.044	_	_	_	_
Manufacturing	0.089	0.002	0.135	0.000	_	_	_	_
Construction	0.077	0.020	0.145	0.001	_	_	_	_
Wholesale Trade	0.108	0.010	0.178	0.000	_	_	_	_
Retail Trade	-0.015	0.660	0.008	0.820	_	_	_	_
Accommodation	-0.041	0.362	0.091	0.023	_	_	_	_
Transport	0.112	0.003	0.266	0.000	_	_	_	_
Communication	0.204	0.000	0.214	0.000	_	_	_	_
Financial Services	0.163	0.000	0.238	0.000	_	_	_	_
Professional Services	0.090	0.024	0.189	0.000	_	_	_	_
Public Administration	0.111	0.002	0.216	0.000	_	_	_	_
Education	_	_	_	_	_	_	_	_
Health	0.032	0.424	0.150	0.000	_	_	_	_
Occupation (reference group is Labourers								
Manager	0.197	0.000	0.183	0.000	_	_	_	_
Professional	0.206	0.000	0.242	0.000	_	_	_	_
Technical Worker	0.073	0.000	0.048	0.122	_	_	_	_
Personal Services	0.138	0.000	0.099	0.006	_	_	_	_
Administration	0.101	0.000	0.048	0.134	_	_	_	_
Sales	-0.011	0.723	0.036	0.403	_	_	_	_
Machine Operator	0.054	0.031	0.101	0.005	_	_	_	_
Other variables								
Year (trend)	-0.013	0.279	0.011	0.317	0.007	0.874	-0.056	0.213
Unemployment Rate	-0.013	0.149	-0.019	0.084	0.001	0.974	0.074	0.078
Constant	1.551	0.000	1.380	0.000	-0.156	0.550	-1.655	0.000
Hhld Non-labour income ($000)	_	_	_	_	-1.722	0.000	-1.365	0.017
Variance function								
Hhld Non-labour income ($000)	_	_	_	_	0.906	0.000	1.694	0.030
Model evaluation								
Mills Ratio	0.051	0.079	0.023	0.376	_	_	_	_
R-squared	0.419	_	0.450	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.076	_	0.122	_
LR test for Heteroskedasticity[28]	_	_	_	_	85.00	0.00	99.32	0.00

Table 5: Single parent: HES88-93

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Age÷10	0.013	0.945	1.199	0.000
Age^2 ÷ 100	-0.001	0.965	-0.151	0.000
Woman	-0.200	0.000	-0.167	0.119
Highest Qualification attained (reference - less than NCEA level 1)				
Level 1-3	0.088	0.056	0.433	0.000
Level 4-6	0.131	0.052	0.693	0.000
Level 7	0.149	0.076	0.709	0.000
Postgraduate	0.152	0.190	1.408	0.000
Other	0.101	0.317	0.756	0.005
Ethnicity (reference group European)				
Maori/Pacific	-0.010	0.842	-0.286	0.001
Other Ethnicity	-0.013	0.889	-0.101	0.632
Children				
Number of children	_	_	-0.009	0.826
Youngest Age = 0	_	_	-1.107	0.000
Youngest Age 1-3	_	_	-0.814	0.000
Youngest Age 4-5	_	_	-0.666	0.000
Youngest Age 6-9	_	_	-0.532	0.000
Youngest Age 10-12	_	_	-0.086	0.456
Industry (reference group is Agriculture or ANZSIC 2006 A)				
Mining	-0.021	0.838	_	_
Manufacturing	0.077	0.243	_	_
Construction	-0.017	0.858	_	_
Wholesale Trade	0.020	0.861	_	_
Retail Trade	-0.134	0.066	_	_
Accommodation	0.065	0.444	_	_
Transport	0.253	0.015	_	_
Communication	0.517	0.001	_	_
Financial Services	0.227	0.032	_	_
Professional Services	0.127	0.104	_	_
Public Administration	0.166	0.025	_	_
Education	_	_	_	_
Health	0.130	0.028	_	_

[28] chi^2 statistic reported. Calculated through hetprobit in Stata for unadjusted model.

Single parent: HES88-93 (continued)

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)				
Manager	0.428	0.000	_	_
Professional	0.430	0.000	_	_
Technical Worker	0.118	0.070	_	_
Personal Services	0.196	0.009	_	_
Administration	0.085	0.191	_	_
Sales	0.248	0.002	_	_
Machine Operator	0.123	0.110	_	_
Region (reference group is Auckland)				
North North Island	0.036	0.453	-0.128	0.205
Central North Island	-0.081	0.126	-0.224	0.042
Wellington	0.004	0.938	0.325	0.007
Canterbury	-0.065	0.203	0.003	0.980
Rest of South Island	-0.092	0.094	0.001	0.994
Other variables				
Year (trend)	-0.057	0.018	0.044	0.424
Unemployment Rate	0.027	0.220	0.009	0.854
Constant	3.435	0.000	-2.918	0.000
Hhld Non-labour income ($000)	_	_	0.268	0.098
Living with parents	_	_	-0.126	0.761
Model evaluation				
Mills Ratio	-0.086	0.276	_	_
R-squared	0.386	_	_	_
Pseudo R-squared	_	_	0.194	_
LR test for Heteroskedasticity	_	_	NA	NA

Table 6: Coupled people: HES88-93

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.381	0.000	0.206	0.000	1.230	0.000	1.494	0.000
$Age^2 \div 10$	-0.044	0.000	-0.025	0.000	-0.168	0.000	-0.220	0.000
Highest Qualification attained (reference group less than NCEA level 1 completion)								
Level 1-3	0.112	0.000	0.108	0.000	0.496	0.000	0.328	0.000
Level 4-6	0.192	0.000	0.225	0.000	0.616	0.000	0.641	0.000
Level 7	0.179	0.017	0.360	0.000	0.993	0.000	0.579	0.014
Postgraduate	0.034	0.779	0.461	0.000	1.589	0.000	1.085	0.000
Other	0.254	0.000	0.301	0.000	0.648	0.021	0.782	0.000
Postgrad × age ÷10	0.116	0.000	-	-	-	-	-	-
Level 7×age÷10	0.036	0.058	-	-	-	-	-	-
Level 4-6 × age ÷10	-0.012	0.275	-	-	-	-	-	-
Postgrad Maori/Pacific	-0.000	1.000	-	-	-	-	-	-
Level 7 Maori/Pacific	_	_	_	_	_	_	_	_
Level 4-6 Maori/Pacific	0.002	0.958	-	-	-	-	-	-
Postgrad Other Ethnicity	-0.061	0.272	-	-	-	-	-	-
Level 7 Other Ethnicity	_	_	_	_	_	_	_	_
Level 4-6 Other Ethnicity	0.025	0.725	-	-	-	-	-	-
Ethnicity (reference group European)								
Maori/Pacific	-0.096	0.019	-0.072	0.000	-0.341	0.000	-0.016	0.806
Other Ethnicity	-0.130	0.003	-0.095	0.004	-0.342	0.044	-0.408	0.001
Children								
Number of children	_	_	_	_	-0.082	0.023	-0.040	0.139
Youngest age = 0	_	_	_	_	-0.043	0.760	-2.041	0.000
Youngest age 1-3	_	_	_	_	-0.072	0.574	-1.453	0.000
Youngest age 4-5	_	_	_	_	0.026	0.867	-1.196	0.000
Youngest age 6-9	_	_	_	_	-0.141	0.321	-0.530	0.000
Youngest age 10-12	_	_	_	_	0.045	0.770	-0.132	0.169
Youngest age 13+	_	_	_	_	-0.108	0.406	0.173	0.055
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	-0.033	0.271	-0.073	0.055	_	_	_	_
Manufacturing	0.115	0.000	0.044	0.091	_	_	_	_
Construction	0.100	0.000	0.158	0.000	_	_	_	_
Wholesale Trade	0.086	0.007	0.187	0.000	_	_	_	_
Retail Trade	-0.103	0.000	-0.008	0.775	_	_	_	_
Accommodation	-0.275	0.000	-0.097	0.002	_	_	_	_
Transport	0.160	0.000	0.186	0.000	_	_	_	_
Communication	0.255	0.000	0.229	0.000	_	_	_	_
Financial Services	0.275	0.000	0.209	0.000	_	_	_	_
Professional Services	0.218	0.000	0.181	0.000	_	_	_	_
Public Administration	0.186	0.000	0.199	0.000	_	_	_	_
Education	_	_	_	_	_	_	_	_
Health	0.005	0.858	0.111	0.000	_	_	_	_

Coupled people: HES88-93 (continued)

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)								
Manager	0.235	0.000	0.234	0.000	_	_	_	_
Professional	0.300	0.000	0.244	0.000	_	_	_	_
Technical Worker	0.153	0.000	0.097	0.000	_	_	_	_
Personal Services	0.196	0.000	0.132	0.000	_	_	_	_
Administration	0.078	0.000	0.084	0.000	_	_	_	_
Sales	0.038	0.159	0.053	0.087	_	_	_	_
Machine Operator	0.095	0.000	0.110	0.000	_	_	_	_
Region (reference group is Auckland)								
North North Island	-0.070	0.000	-0.091	0.000	-0.324	0.001	-0.119	0.051
Central North Island	-0.090	0.000	-0.080	0.000	-0.304	0.003	-0.029	0.651
Wellington	0.044	0.006	0.012	0.480	0.080	0.455	0.120	0.073
Canterbury	-0.109	0.000	-0.066	0.000	0.058	0.595	-0.019	0.780
Rest of South Island	-0.079	0.000	-0.092	0.000	-0.048	0.648	-0.022	0.739
Other variables								
Year (trend)	-0.023	0.011	-0.000	1.000	-0.045	0.405	-0.075	0.027
Unemployment Rate	-0.005	0.405	-0.005	0.532	0.043	0.270	0.226	0.000
Constant	2.141	0.000	2.203	0.000	-1.495	0.012	-3.668	0.000
Hhld Non-labour income (thousands)	_	_	_	_	-0.139	0.006	-0.317	0.004
Partner participation	_	_	_	_	0.308	0.029	0.719	0.000
Partner income (thousands)	_	_	_	_	2.799	0.001	-0.166	0.000
Variance Function								
Age÷10	-	-	-	-	0.087	0.007	0.121	0.000
Hhld Non-labour income (thousands)	-	-	-	-	1.084	0.000	0.162	0.043
Partner participation	-	-	-	-	-	-	-0.229	0.002
Model evaluation								
Mills Ratio	0.085	0.002	0.066	0.000	_	_	_	_
R-squared	0.324	_	0.261	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.075	_	0.183	_
LR test for Heteroskedasticity	_	_	_	_	237.49	0.000	152.21	0.000

8.3 HES94-98

Table 7: Single people without children: HES94-98

	Wage Equation				Selection Equation			
	Single Men		Single Women		Single Men		Single Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.741	0.000	0.460	0.000	0.698	0.000	1.696	0.001
$Age^2 \div 100$	-0.089	0.000	-0.051	0.000	-0.109	0.000	-0.223	0.000
Highest Qualification attained (reference group less than NCEA level 1 completion)								
Level 1-3	0.187	0.000	0.094	0.000	0.634	0.000	0.993	0.000
Level 4-6	0.256	0.000	0.159	0.005	0.734	0.000	1.225	0.000
Level 7	0.395	0.000	-0.014	0.846	0.737	0.000	1.111	0.000
Postgraduate	-	-	-0.083	0.545	-	-	1.692	0.003
Other	0.157	0.000	0.162	0.016	0.416	0.083	0.745	0.022
Postgrad × age ÷10	-	-	0.118	0.000	_	_	-	-
Level 7×age÷10	-	-	0.082	0.000	_	_	-	-
Level 4-6 × age ÷10	-	-	0.000	1.000	_	_	-	-
Postgrad Maori/Pacific	-	-	-0.040	0.674	_	_	-	-
Level 4-6 Maori/Pacific	-	-	0.035	0.572	_	_	-	-
Postgraduate Other	-	-	0.131	0.373	_	_	-	-
Level 4-6 Other	-	-	-0.963	0.001	_	_	-	-
Ethnicity (reference group Non-Maori/Pacific)								
Maori/Pacific	-0.084	0.011	-0.037	0.233	-0.519	0.000	-0.609	0.001
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	0.062	0.159	0.018	0.768	_	_	_	_
Manufacturing	0.141	0.000	0.090	0.021	_	_	_	_
Construction	0.024	0.558	0.056	0.505	_	_	_	_
Wholesale Trade	0.125	0.018	0.101	0.076	_	_	_	_
Retail Trade	0.000	0.981	-0.039	0.330	_	_	_	_
Accommodation	-0.010	0.853	-0.020	0.634	_	_	_	_
Transport	0.175	0.000	0.042	0.419	_	_	_	_
Communication	0.245	0.000	0.229	0.000	_	_	_	_
Financial Services	0.319	0.000	0.184	0.000	_	_	_	_
Professional Services	0.124	0.006	0.118	0.003	_	_	_	_
Public Administration	0.219	0.000	0.134	0.002	_	_	_	_
Education	-0.032	0.682	-0.078	0.104	_	_	_	_
Health	0.031	0.543	0.051	0.145	_	_	_	_
Occupation (reference group is Labourer)								
Manager	0.284	0.000	0.307	0.000	_	_	_	_
Professional	0.239	0.000	0.281	0.000	_	_	_	_
Technical Worker	0.151	0.000	0.147	0.000	_	_	_	_
Personal Services	0.197	0.073	0.179	0.007	_	_	_	_
Administration	0.104	0.006	0.174	0.000	_	_	_	_
Sales	0.100	0.023	0.119	0.013	_	_	_	_
Machine Operator	0.037	0.218	0.007	0.891	_	_	_	_

Single people without children: HES94-98 (continued)

	Wage Equation				Selection Equation			
	Single Men		Single Women		Single Men		Single Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Region (reference group is Auckland)								
North North Island	-0.127	0.000	-0.040	0.110	-0.343	0.001	-0.210	0.214
Central North Island	-0.102	0.000	-0.134	0.000	-0.301	0.008	0.218	0.276
Wellington	-0.043	0.098	0.047	0.033	-0.084	0.477	0.275	0.122
Canterbury	-0.095	0.000	-0.082	0.001	0.018	0.884	0.212	0.272
Rest of South Island	-0.088	0.001	-0.103	0.000	-0.252	0.022	0.225	0.241
Other variables								
Year (trend)	0.008	0.317	0.002	0.739	0.104	0.006	-0.014	0.745
Unemployment Rate	-0.016	0.046	0.010	0.211	-0.006	0.864	-0.081	0.199
Constant	1.125	0.000	1.557	0.000	-0.448	0.347	-1.391	0.119
Hhld Non-labour income (thousands)	_	_	_	_	-0.888	0.032	-1.571	0.028
Living with parents	_	_	_	_	_	_	_	_
Variance Function								
Age÷10	-	-	-	-	-	-	0.109	0.065
Hhld Non-labour income (thousands)	-	-	-	-	0.911	0.056	0.840	0.246
Model evaluation								
Mills Ratio	0.430	0.000	0.017	0.395	_	_	_	_
R-squared	0.414	_	0.443	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.128	_	0.136	_
LR test for Heteroskedasticity	_	_	_	_	38.22	0.001	54.27	0.000

Table 8: Single Parent: HES94-98

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Age÷10	0.426	0.030	0.823	0.021
$Age^2 \div 10$	-0.050	0.046	-0.123	0.008
Woman	-0.096	0.112	-0.298	0.024
Highest Qualification attained (reference - less than NCEA level 1)				
Level 1-3	0.033	0.569	0.524	0.000
Level 4-6	0.178	0.010	0.641	0.000
Level 7	0.216	0.022	0.940	0.000
Postgraduate	0.435	0.003	0.545	0.090
Other	0.259	0.059	0.688	0.027
Ethnicity (reference group Non-Maori/Pacific)				
Maori/Pacific	-0.028	0.611	-0.157	0.116
Children				
Number of children	_	_	-0.190	0.000
Youngest Age = 0	_	_	-1.608	0.000
Youngest Age 1-3	_	_	-1.216	0.000
Youngest Age 4-5	_	_	-0.848	0.000
Youngest Age 6-9	_	_	-0.620	0.000
Youngest Age 10-12	_	_	-0.406	0.006
Industry (reference group is Agriculture or ANZSIC 2006 A)				
Mining	0.116	0.361	_	_
Manufacturing	0.053	0.569	_	_
Construction	0.067	0.644	_	_
Wholesale Trade	-0.034	0.829	_	_
Retail Trade	-0.051	0.624	_	_
Accommodation	-0.060	0.528	_	_
Transport	0.073	0.559	_	_
Communication	0.032	0.884	_	_
Financial Services	0.202	0.138	_	_
Professional Services	0.281	0.009	_	_
Public Administration	0.220	0.040	_	_
Education	-0.056	0.568	_	_
Health	-0.101	0.207	_	_

Single Parent: HES94-98 (continued)

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)				
Manager	0.183	0.087	_	_
Professional	0.397	0.000	_	_
Technical Worker	0.178	0.015	_	_
Personal Services	0.390	0.012	_	_
Administration	0.105	0.211	_	_
Sales	0.042	0.729	_	_
Machine Operator	0.147	0.154	_	_
Region (reference group is Auckland)				
North North Island	-0.072	0.222	-0.018	0.883
Central North Island	-0.064	0.302	0.166	0.202
Wellington	-0.023	0.724	0.139	0.341
Canterbury	0.016	0.817	0.298	0.050
Rest of South Island	-0.001	0.988	0.034	0.820
Other variables				
Year (trend)	0.022	0.196	0.048	0.170
Unemployment Rate	-0.004	0.849	-0.032	0.487
Constant	1.744	0.000	-0.420	0.615
Hhld Non-labour income (thousands)	_	_	-0.083	0.377
Living with parents	_	_	-1.184	0.140
Model evaluation				
Mills Ratio	0.036	0.665	_	_
R-squared	0.257	_	_	_
Pseudo R-squared	_	_	0.197	_
LR test for Heteroskedasticity	_	_	NA	NA

Table 9: Coupled people: HES94-98

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.425	0.000	0.279	0.000	0.786	0.000	1.399	0.000
$Age^2 \div 100$	-0.046	0.000	-0.033	0.000	-0.110	0.000	-0.191	0.000
Highest Qualification attained (reference group less than NCEA level 1 completion)								
Level 1-3	0.131	0.000	0.102	0.000	0.338	0.000	0.451	0.000
Level 4-6	0.134	0.000	0.170	0.000	0.493	0.000	0.634	0.000
Level 7	0.287	0.000	0.260	0.000	0.489	0.000	0.745	0.000
Postgraduate	0.416	0.000	0.377	0.000	0.663	0.000	1.370	0.000
Other	0.140	0.001	0.161	0.000	0.493	0.026	1.055	0.000
Ethnicity (reference group non-Maori/Pacific)								
Maori/Pacific	-0.069	0.008	-0.069	0.011	-0.177	0.127	-0.096	0.870
Children								
Number of children	_	_	_	_	-0.016	0.628	-0.130	0.001
Youngest age = 0	_	_	_	_	0.248	0.081	-1.850	0.000
Youngest age 1-3	_	_	_	_	-0.058	0.629	-1.216	0.000
Youngest age 4-5	_	_	_	_	-0.159	0.260	-0.825	0.000
Youngest age 6-9	_	_	_	_	-0.123	0.337	-0.352	0.00
Youngest age 10-12	_	_	_	_	-0.268	0.044	-0.103	0.406
Youngest age 13+	_	_	_	_	-0.132	0.259	0.054	0.607
Partner Characteristics								
Maori/Pacific	-0.040	0.139	-0.030	0.249	-0.232	0.044	0.229	0.045
Level 1-3	_	_	_	_	0.127	0.078	0.005	0.945
Level 4-6	_	_	_	_	0.228	0.010	0.018	0.791
Level 7	_	_	_	_	-0.009	0.949	-0.230	0.024
Postgraduate	_	_	_	_	-0.159	0.491	-0.570	0.001
Other	_	_	_	_	0.151	0.522	-0.017	0.924
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	0.031	0.376	0.125	0.001	_	_	_	_
Manufacturing	0.121	0.000	0.111	0.000	_	_	_	_
Construction	0.015	0.639	0.106	0.054	_	_	_	_
Wholesale Trade	0.111	0.004	0.087	0.034	_	_	_	_
Retail Trade	-0.072	0.024	-0.100	0.001	_	_	_	_
Accommodation	-0.287	0.000	-0.170	0.000	_	_	_	_
Transport	0.081	0.014	0.061	0.127	_	_	_	_
Communication	0.245	0.000	0.117	0.005	_	_	_	_
Financial Services	0.337	0.000	0.113	0.001	_	_	_	_
Professional Services	0.203	0.000	0.135	0.000	_	_	_	_
Public Administration	0.175	0.000	0.121	0.000	_	_	_	_
Education	-0.118	0.014	-0.113	0.001	_	_	_	_
Health	-0.031	0.376	-0.030	0.255	_	_	_	_

Coupled people: HES94-98 (continued)

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)								
Manager	0.324	0.000	0.313	0.000	_	_	_	_
Professional	0.287	0.000	0.374	0.000	_	_	_	_
Technical Worker	0.212	0.000	0.131	0.000	_	_	_	_
Personal Services	0.220	0.016	0.354	0.000	_	_	_	_
Administration	0.104	0.002	0.225	0.000	_	_	_	_
Sales	0.033	0.421	0.112	0.001	_	_	_	_
Machine Operator	0.099	0.000	-0.008	0.819	_	_	_	_
Region (reference group is Auckland)								
North North Island	-0.043	0.017	-0.063	0.000	-0.026	0.751	0.026	0.714
Central North Island	-0.106	0.000	-0.069	0.000	0.003	0.973	0.121	0.123
Wellington	0.061	0.001	0.024	0.182	0.229	0.026	0.302	0.001
Canterbury	-0.056	0.003	-0.084	0.000	0.013	0.888	0.044	0.582
Rest of South Island	-0.056	0.003	-0.082	0.000	-0.000	1.000	0.066	0.404
Other variables								
Year (trend)	0.005	0.405	0.006	0.230	0.074	0.008	-0.006	0.752
Unemployment Rate	-0.004	0.424	-0.006	0.317	-0.013	0.630	-0.059	0.042
Constant	1.758	0.000	2.037	0.000	-0.898	0.088	-1.736	0.004
Hhld Non-labour income (thousands)	_	_	_	_	-0.199	0.114	-0.254	0.013
Partner participation	_	_	_	_	0.358	0.002	0.932	0.000
Partner income (thousands)	_	_	_	_	2.284	0.000	-0.094	0.004
Variance Function								
Age÷10	-	-	-	-	-	-	0.122	0.002
Hhld Non-labour income (thousands)	-	-	-	-	0.365	0.003	0.202	0.050
Partner income (thousands)	_	_	_	_	0.892	0.000	-	-
Partner participation	-	-	-	-	-	-	-0.290	0.003
Model evaluation								
Mills Ratio	0.090	0.006	0.101	0.000	_	_	_	_
R-squared	0.287	_	0.256	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.107	_	0.177	_
LR test for Heteroskedasticity	_	_	_	_	87.59	0.000	93.59	0.000

8.4 HES01-08

Table 10: Single people without children: HES01-08

	Wage Equation				Selection Equation			
	Single Men		Single Women		Single Men		Single Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.497	0.000	0.400	0.000	1.131	0.000	1.297	0.000
$Age^2 \div 100$	-0.055	0.000	-0.046	0.000	-0.158	0.000	-0.165	0.000
Highest Qualification attained (reference group less than NCEA level 1 completion)								
Level 1/2	0.075	0.006	0.090	0.006	0.092	0.376	0.589	0.000
Level 3	0.130	0.001	0.077	0.336	0.392	0.026	1.024	0.000
Level 4-6	0.103	0.001	0.094	0.167	0.365	0.007	0.703	0.000
Level 7	0.285	0.000	-0.007	0.927	0.650	0.000	0.952	0.000
Postgraduate	0.382	0.000	-0.037	0.725	0.203	0.290	0.777	0.000
Other	0.180	0.000	0.082	0.261	0.149	0.484	0.516	0.011
Postgrad × age ÷10	-	-	0.078	0.001	-	-	-	-
Level 7×age÷10	-	-	0.049	0.004	-	-	-	-
Level 4-6 × age ÷10	-	-	0.006	0.644	-	-	-	-
Postgrad Maori/Pacific	-	-	0.077	0.222	-	-	-	-
Level 4-6 Maori/Pacific	-	-	-0.121	0.018	-	-	-	-
Postgraduate Other Non-European	-	-	0.120	0.029	-	-	-	-
Level 4-6 × Other Non-European	-	-	-0.020	0.758	-	-	-	-
Ethnicity (reference group European)								
Maori/Pacific	-0.036	0.380	0.001	0.980	-0.142	0.425	0.093	0.650
Maori/Pacific Only	-0.046	0.125	-0.091	0.009	-0.384	0.001	-0.580	0.000
Other Non-European	-0.094	0.001	-0.142	0.001	-0.549	0.000	-0.514	0.000
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	0.047	0.338	0.077	0.229	_	_	_	_
Manufacturing	0.096	0.008	0.037	0.330	_	_	_	_
Construction	0.050	0.200	0.052	0.488	_	_	_	_
Wholesale Trade	0.029	0.554	0.108	0.019	_	_	_	_
Retail Trade	-0.036	0.380	-0.120	0.001	_	_	_	_
Accommodation	-0.060	0.230	-0.109	0.005	_	_	_	_
Transport	0.041	0.373	0.057	0.264	_	_	_	_
Communication	0.161	0.007	-0.037	0.477	_	_	_	_
Financial Services	0.230	0.000	0.114	0.004	_	_	_	_
Professional Services	0.099	0.024	0.112	0.001	_	_	_	_
Public Administration	0.139	0.005	0.167	0.000	_	_	_	_
Education	0.084	0.148	-0.051	0.168	_	_	_	_
Health	0.143	0.011	-0.004	0.904	_	_	_	_

Single person without children: HES01-08 (continued)

	Wage Equation				Selection Equation			
	Single Men		Single Women		Single Men		Single Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)								
Manager	0.211	0.000	0.375	0.000	_	_	_	_
Professional	0.267	0.000	0.353	0.000	_	_	_	_
Technical Worker	0.099	0.000	0.094	0.016	_	_	_	_
Personal Services	0.031	0.535	0.080	0.057	_	_	_	_
Administration	0.067	0.111	0.147	0.000	_	_	_	_
Sales	0.017	0.678	0.116	0.006	_	_	_	_
Machine Operator	0.054	0.123	-0.021	0.943	_	_	_	_
Region (reference group is Auckland)								
North North Island	-0.109	0.000	-0.155	0.000	-0.151	0.216	-0.044	0.727
Central North Island	-0.123	0.000	-0.130	0.000	-0.217	0.098	0.246	0.113
Wellington	-0.048	0.065	-0.041	0.088	0.005	0.967	0.276	0.032
Canterbury	-0.127	0.000	-0.110	0.000	-0.033	0.790	0.214	0.118
Rest of South Island	-0.130	0.000	-0.107	0.000	-0.045	0.737	0.143	0.311
Other Variables								
Year (trend)	-0.001	0.868	-0.013	0.001	-0.001	0.972	-0.020	0.385
Unemployment Rate	0.002	0.912	-0.023	0.176	-0.027	0.736	-0.052	0.559
Constant	1.680	0.000	2.030	0.000	-0.516	0.379	-1.241	0.049
Hhld Non-labour income (thousands)	_	_	_	_	-1.128	0.012	-0.561	0.249
Variance Function								
Hhld Non-labour income (thousands)	-	-	-	-	1.249	0.000	1.153	0.066
Model evaluation								
Mills Ratio	0.099	0.001	0.060	0.453	_	_	_	_
R-squared	0.417	_	0.474	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.091	_	0.129	_
LR test for Heteroskedasticity	_	_	_	_	56.71	0.000	43.91	0.021

Table 11: Single parent: HES01-08

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Age÷10	0.389	0.039	1.282	0.000
$Age^2 \div 100$	-0.040	0.082	-0.163	0.000
Woman	-0.034	0.513	-0.108	0.368
Highest Qualification attained (reference - less than NCEA level 1)				
Level 1/2	0.065	0.237	0.524	0.000
Level 3	0.124	0.126	0.434	0.031
Level 4-6	0.121	0.055	0.563	0.000
Level 7	0.310	0.000	0.547	0.000
Postgraduate	0.460	0.000	0.731	0.000
Other	0.085	0.300	0.199	0.179
Ethnicity (reference group European)				
Maori/Pacific	-0.065	0.332	0.030	0.851
Maori/Pacific Only	-0.083	0.071	-0.152	0.144
Other Non-European	-0.086	0.159	-0.489	0.000
Children				
Number of children	_	_	0.022	0.617
Youngest Age = 0	_	_	-1.404	0.000
Youngest Age 1-3	_	_	-0.875	0.000
Youngest Age 4-5	_	_	-0.788	0.000
Youngest Age 6-9	_	_	-0.433	0.001
Youngest Age 10-12	_	_	-0.327	0.004
Industry (reference group is Agriculture or ANZSIC 2006 A)				
Mining	0.319	0.025	_	_
Manufacturing	0.250	0.001	_	_
Construction	0.284	0.003	_	_
Wholesale Trade	0.187	0.042	_	_
Retail Trade	0.031	0.675	_	_
Accommodation	0.001	0.990	_	_
Transport	0.207	0.029	_	_
Communication	0.002	0.987	_	_
Financial Services	0.216	0.022	_	_
Professional Services	0.207	0.007	_	_
Public Administration	0.180	0.034	_	_
Education	-0.014	0.842	_	_
Health	0.046	0.472	_	_

Single parent: HES01-08 (continued)

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)				
Manager	0.408	0.000	_	_
Professional	0.396	0.000	_	_
Technical Worker	0.079	0.210	_	_
Personal Services	0.135	0.072	_	_
Administration	0.228	0.000	_	_
Sales	0.127	0.126	_	_
Machine Operator	-0.056	0.551	_	_
Region (reference group is Auckland)				
North North Island	-0.138	0.008	-0.061	0.623
Central North Island	-0.055	0.290	0.193	0.165
Wellington	0.045	0.413	-0.057	0.654
Canterbury	0.006	0.913	0.419	0.002
Rest of South Island	-0.107	0.075	0.377	0.003
Other variables				
Year (trend)	-0.011	0.271	0.003	0.896
Unemployment Rate	-0.037	0.290	0.127	0.131
Constant	1.743	0.001	-2.618	0.001
Hhld Non-labour income (thousands)	_	_	-0.344	0.000
Variance Function				
Hhld Non-labour income (thousands)	-	-	-2.984	0.001
Model evaluation				
Mills Ratio	0.188	0.065	_	_
R-squared	0.400	_	_	_
Pseudo R-squared	_	_	0.200	_
LR test for Heteroskedasticity	_	_	8.41	0.004

Table 12: Coupled person: HES01-08

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.487	0.000	0.330	0.000	0.499	0.000	0.741	0.000
$Age^2 \div 100$	-0.054	0.000	-0.036	0.000	-0.078	0.000	-0.110	0.000
Highest Qualification attained (reference group less than NCEA level 1 completion)								
Level 1/2	0.086	0.000	0.031	0.103	0.106	0.141	0.222	0.000
Level 3	0.174	0.004	0.091	0.012	0.196	0.055	0.286	0.015
Level 4-6	0.201	0.001	0.111	0.000	0.337	0.000	0.326	0.000
Level 7	0.205	0.005	0.227	0.000	0.400	0.000	0.504	0.000
Postgraduate	-0.020	0.830	0.333	0.000	0.338	0.002	0.466	0.000
Other	0.277	0.000	0.127	0.000	0.146	0.301	0.222	0.018
Postgrad × age ÷10	0.104	0.000	-	-	-	-	-	-
Level 7×age÷10	0.030	0.078	-	-	-	-	-	-
Level 4-6 × age ÷10	-0.013	0.415	-	-	-	-	-	-
Postgrad Maori/Pacific	0.010	0.858	-	-	-	-	-	-
Level 4-6 Maori/Pacific	-0.031	0.304	-	-	-	-	-	-
Postgrad Other Non-European	-0.218	0.000	-	-	-	-	-	-
Level 4-6 Other Non-European	0.029	0.554	-	-	-	-	-	-
Ethnicity (reference group European)								
Maori/Pacific	0.024	0.480	-0.059	0.057	0.326	0.641	-0.115	0.201
Maori/Pacific Only	-0.175	0.000	-0.110	0.000	-0.691	0.000	0.050	0.397
Other Non-European	-0.153	0.033	-0.165	0.000	-0.800	0.000	-0.422	0.000
Children								
Number of children	_	_	_	_	0.034	0.331	-0.104	0.001
Youngest age = 0	_	_	_	_	-0.164	0.179	-1.233	0.000
Youngest age 1-3	_	_	_	_	-0.132	0.239	-0.723	0.000
Youngest age 4-5	_	_	_	_	-0.233	0.080	-0.505	0.000
Youngest age 6-9	_	_	_	_	-0.016	0.890	-0.251	0.008
Youngest age 10-12	_	_	_	_	-0.291	0.012	-0.081	0.363
Youngest age 13+	_	_	_	_	0.072	0.489	0.142	0.091
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	-0.022	0.518	0.019	0.659	_	_	_	_
Manufacturing	0.166	0.000	0.074	0.011	_	_	_	_
Construction	0.110	0.000	0.169	0.002	_	_	_	_
Wholesale Trade	0.153	0.000	0.075	0.043	_	_	_	_
Retail Trade	-0.024	0.439	-0.108	0.000	_	_	_	_
Accommodation	-0.246	0.000	-0.164	0.000	_	_	_	_
Transport	0.124	0.000	0.063	0.089	_	_	_	_
Communication	0.141	0.002	0.139	0.002	_	_	_	_
Financial Services	0.348	0.000	0.203	0.000	_	_	_	_
Professional Services	0.227	0.000	0.138	0.000	_	_	_	_
Public Administration	0.218	0.000	0.159	0.000	_	_	_	_
Education	-0.032	0.347	-0.057	0.035	_	_	_	_
Health	0.128	0.001	-0.012	0.631	_	_	_	_

Couple people: HES01-08 (continued)

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)								
Manager	0.320	0.000	0.387	0.000	_	_	_	_
Professional	0.270	0.000	0.405	0.000	_	_	_	_
Technical Worker	0.129	0.000	0.127	0.000	_	_	_	_
Personal Services	0.091	0.023	0.134	0.000	_	_	_	_
Administration	0.078	0.015	0.189	0.000	_	_	_	_
Sales	0.106	0.003	0.113	0.000	_	_	_	_
Machine Operator	0.004	0.873	0.028	0.495	_	_	_	_
Region (reference group is Auckland)								
North North Island	-0.120	0.000	-0.116	0.000	-0.071	0.337	0.030	0.579
Central North Island	-0.145	0.000	-0.144	0.000	0.023	0.798	0.050	0.420
Wellington	-0.039	0.030	-0.011	0.541	-0.023	0.774	0.142	0.014
Canterbury	-0.104	0.000	-0.090	0.000	0.099	0.255	0.159	0.009
Rest of South Island	-0.139	0.000	-0.097	0.000	-0.069	0.411	0.116	0.046
Other variables								
Year (trend)	0.001	0.803	-0.013	0.000	-0.008	0.657	-0.016	0.110
Unemployment Rate	0.004	0.739	-0.029	0.026	-0.034	0.521	-0.111	0.007
Constant	1.693	0.000	2.119	0.000	0.635	0.193	-0.108	0.773
Hhld Non-labour income (thousands)	_	_	_	_	-0.226	0.031	-0.110	0.012
Partner participation	_	_	_	_	0.471	0.000	0.469	0.000
Partner income (thousands)	_	_	_	_	-0.102	0.037	-0.028	0.020
Variance Function								
Maori/Pacific	-	-	-	-	0.356	0.374	-	-
Maori/Pacific Only	-	-	-	-	-0.452	0.034	-	-
Other Non-European	-	-	-	-	-0.320	0.108	-	-
Age÷10	-	-	-	-	-	-	0.045	0.199
Hhld Non-labour income (thousands)	-	-	-	-	0.386	0.001	0.102	0.089
Model evaluation								
Mills Ratio	0.193	0.000	0.103	0.000	_	_	_	_
R-squared	0.364	_	0.347	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.107	_	0.154	_
LR test for Heteroskedasticity	_	_	_	_	87.49	0.000	82.25	0.000

8.5 HES09-13

Table 13: Single people without children: HES09-13

	Wage Equation				Selection Equation			
	Single Men		Single Women		Single Men		Single Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.289	0.000	0.350	0.000	0.984	0.000	1.021	0.000
$Age^2 \div 100$	-0.029	0.000	-0.039	0.000	-0.108	0.000	-0.118	0.000
Highest Qualification attained (reference group less than NCEA level 1 completion)								
Level 1/2	0.029	0.227	0.067	0.005	0.286	0.007	0.219	0.029
Level 3	0.189	0.001	0.049	0.382	0.947	0.000	0.480	0.002
Level 4-6	0.160	0.008	0.112	0.042	0.788	0.002	0.286	0.021
Level 7	-0.082	0.300	-0.061	0.293	0.866	0.002	0.657	0.000
Postgraduate	0.013	0.162	-0.027	0.708	1.025	0.008	0.580	0.000
Other	0.191	0.004	0.161	0.007	1.000	0.001	0.482	0.006
Postgrad × age ÷10	0.082	0.000	0.055	0.001	-0.119	0.171	-	-
Level 7×age÷10	0.074	0.000	0.055	0.000	-0.68	0.331	-	-
Level 4-6 × age ÷10	-0.006	0.644	0.001	0.928	-0.127	0.013	-	-
Postgrad Maori/Pacific	-0.077	0.285	0.044	0.317	0.243	0.315	-	-
Level 7 Maori/Pacific	_	_	_	_	_	_	_	_
Level 4-6 Maori/Pacific	-0.103	0.028	0.029	0.479	0.030	0.871	-	-
Postgrad Other Non-European	-0.034	0.488	0.017	0.678	0.174	0.357	-	-
Level 7 Other Non-European	_	_	_	_	_	_	_	_
Level 4-6 Other Non-European	-0.086	0.105	0.008	0.875	0.203	0.371	-	-
Ethnicity (reference group European)								
Maori/Pacific	0.015	0.617	-0.045	0.121	-0.026	0.846	0.013	0.914
Maori/Pacific Only	-0.047	0.105	-0.111	0.000	-0.535	0.000	-0.504	0.000
Other Non-European	-0.066	0.052	-0.058	0.046	-0.523	0.000	-0.441	0.000
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	0.002	0.962	-0.007	0.922	_	_	_	_
Manufacturing	0.085	0.003	0.038	0.250	_	_	_	_
Construction	0.053	0.077	0.019	0.725	_	_	_	_
Wholesale Trade	0.017	0.646	0.071	0.083	_	_	_	_
Retail Trade	-0.078	0.022	-0.122	0.000	_	_	_	_
Accommodation	-0.103	0.004	-0.091	0.002	_	_	_	_
Transport	0.034	0.383	0.062	0.121	_	_	_	_
Communication	0.081	0.078	-0.012	0.785	_	_	_	_
Financial Services	0.195	0.000	0.153	0.000	_	_	_	_
Professional Services	0.103	0.005	0.035	0.274	_	_	_	_
Public Administration	0.141	0.000	0.139	0.000	_	_	_	_
Education	-0.002	0.961	-0.014	0.617	_	_	_	_
Health	0.038	0.409	0.031	0.233	_	_	_	_

Single people without children: HES09-13 (continued)

	Wage Equation				Selection Equation			
	Single Men		Single Women		Single Men		Single Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)								
Manager	0.233	0.000	0.323	0.000	_	_	_	_
Professional	0.271	0.000	0.362	0.000	_	_	_	_
Technical Worker	0.120	0.000	0.077	0.032	_	_	_	_
Personal Services	0.097	0.004	0.044	0.116	_	_	_	_
Administration	0.090	0.008	0.170	0.000	_	_	_	_
Sales	0.058	0.070	0.092	0.003	_	_	_	_
Machine Operator	0.052	0.073	0.002	0.969	_	_	_	_
Region (reference group is Auckland)								
North North Island	-0.065	0.001	-0.074	0.000	0.020	0.835	0.021	0.827
Central North Island	-0.088	0.002	-0.108	0.000	0.214	0.061	0.216	0.058
Wellington	0.005	0.275	-0.018	0.344	0.250	0.006	0.242	0.008
Canterbury	-0.050	0.023	-0.081	0.000	0.334	0.001	0.332	0.001
Rest of South Island	-0.076	0.002	-0.063	0.003	0.083	0.407	0.081	0.418
Other variables								
Year (trend)	-0.011	0.023	-0.009	0.199	0.031	0.348	0.031	0.348
Unemployment Rate	-0.003	0.785	0.010	0.363	-0.045	0.378	-0.047	0.357
Constant	2.106	0.000	1.967	0.000	-1.145	0.003	-1.089	0.004
Hhld Non-labour income (thousands)	_	_	_	_	-1.153	0.000	-1.145	0.000
Variance Function								
Hhld Non-labour income (thousands)	_	_	_	_	0.703	0.000	0.696	0.000
Model evaluation								
Mills Ratio	0.046	0.238	0.012	0.689	_	_	_	_
R-squared	0.369	_	0.486	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.087	_	0.103	_
LR test for Heteroskedasticity	_	_	_	_	83.85	0.000	63.59	0.000

Table 14: Single parent: HES09-13

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Age÷10	0.108	0.624	1.035	0.001
$Age^2 \div 100$	-0.011	0.432	-0.102	0.002
Woman	-0.086	0.027	-0.074	0.572
Highest Qualification attained (reference group less than NCEA level 1 completion				
Level 1/2	0.057	0.18	0.481	0.000
Level 3	0.062	0.268	0.567	0.000
Level 4-6	0.089	0.112	0.768	0.190
Level 7	0.238	0.000	0.785	0.000
Postgraduate	0.356	0.000	1.183	0.000
Other	0.150	0.025	0.205	0.340
Ethnicity (reference group European)				
Maori/Pacific	-0.068	0.089	-0.026	0.837
Maori/Pacific Only	-0.008	0.829	-0.210	0.046
Other Non-European	-0.108	0.016	-0.153	0.314
Children				
Number of children	_	_	-0.122	0.011
Youngest Age = 0	_	_	-0.797	0.001
Youngest Age 1-3	_	_	-0.633	0.000
Youngest Age 4-5	_	_	-0.589	0.001
Youngest Age 6-9	_	_	-0.355	0.012
Youngest Age 10-12	_	_	-0.356	0.012
Industry (reference group is Agriculture or ANZSIC 2006 A)				
Mining	0.096	0.337	_	_
Manufacturing	0.118	0.014	_	_
Construction	0.183	0.029	_	_
Wholesale Trade	0.118	0.121	_	_
Retail Trade	-0.105	0.117	_	_
Accommodation	-0.040	0.551	_	_
Transport	0.122	0.142	_	_
Communication	0.292	0.004	_	_
Financial Services	0.135	0.057	_	_
Professional Services	0.224	0.001	_	_
Public Administration	0.213	0.001	_	_
Education	0.038	0.497	_	_
Health	0.064	0.218	_	_

Single parent: HES09-13 (continued)

	Wage Equation Single Parent		Selection Equation Single Parent	
	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)				
Manager	0.441	0.000	_	_
Professional	0.253	0.000	_	_
Technical Worker	0.121	0.055	_	_
Personal Services	-0.060	0.267	_	_
Administration	0.151	0.003	_	_
Sales	0.081	0.227	_	_
Machine Operator	-0.059	0.482	_	_
Region (reference group is Auckland)				
North North Island	-0.065	0.104	0.215	0.078
Central North Island	-0.074	0.093	0.307	0.028
Wellington	0.070	0.080	0.267	0.043
Canterbury	-0.065	0.122	0.405	0.004
Rest of South Island	-0.014	0.756	0.432	0.003
Other variables				
Year (trend)	-0.028	0.020	0.005	0.910
Unemployment Rate	0.014	0.484	-0.049	0.471
Constant	2.534	0.000	-1.899	0.007
Hhld Non-labour income (thousands)	_	_	6.360	0.000
Variance equation				
Hhld Non-labour income (thousands)	_	_	3.130	0.000
Model evaluation				
Mills Ratio	-0.084	0.148	_	_
R-squared	0.423	_	_	_
Pseudo R-squared	_	_	0.189	_
LR test for Heteroskedasticity	_	_	83.00	0.000

Table 15: Coupled people: HES09-13

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Age÷10	0.427	0.000	0.250	0.000	2.409	0.000	0.982	0.000
$Age^2 \div 100$	-0.047	0.000	-0.027	0.000	-0.293	0.000	-0.127	0.000
Highest Qualification attained (reference group less than NCEA level 1 completion								
Level 1/2	0.055	0.002	0.081	0.000	0.443	0.006	0.344	0.000
Level 3	0.093	0.068	0.054	0.299	0.728	0.000	0.429	0.000
Level 4-6	0.165	0.002	0.099	0.057	0.736	0.001	0.530	0.000
Level 7	-0.037	0.602	0.200	0.001	0.701	0.001	0.775	0.000
Postgraduate	0.186	0.017	0.106	0.114	0.651	0.003	0.681	0.000
Other	0.199	0.000	0.051	0.354	0.894	0.002	0.477	0.000
Postgrad × age ÷10	0.044	0.006	0.051	0.000	-	-	-	-
Level 7×age÷10	0.084	0.000	0.013	0.279	-	-	-	-
Level 4-6 × age ÷10	0.003	0.764	0.018	0.72	-	-	-	-
Postgrad Maori/Pacific	-0.062	0.206	0.018	0.617	-	-	-	-
Level 7 Maori/Pacific	_	_	_	_	_	_	_	_
Level 4-6 Maori/Pacific	-0.055	0.076	0.027	0.413	-	-	-	-
Postgrad Other Non-European	-0.142	0.000	-0.032	0.317	-	-	-	-
Level 7 Other Non-European	_	_	_	_	_	_	_	_
Level 4-6 Other Non-European	0.008	0.833	0.017	0.637	-	-	-	-
Ethnicity (reference group European)								
Maori/Pacific	0.007	0.795	-0.065	0.009	-0.279	0.184	0.116	0.172
Maori/Pacific Only	-0.123	0.000	-0.132	0.000	-0.645	0.000	-0.091	0.155
Other Non-European	-0.166	0.000	-0.137	0.000	-0.833	0.000	-0.543	0.000
Children								
Number of children	_	_	_	_	0.220	0.140	-0.150	0.000
Youngest age = 0	_	_	_	_	0.171	0.465	-1.189	0.000
Youngest age 1-3	_	_	_	_	-0.173	0.421	-0.839	0.000
Youngest age 4-5	_	_	_	_	-0.013	0.963	-0.468	0.000
Youngest age 6-9	_	_	_	_	0.071	0.782	-0.243	0.008
Youngest age 10-12	_	_	_	_	0.378	0.206	0.149	0.128
Youngest age 13+	_	_	_	_	0.181	0.435	0.192	0.016
Industry (reference group is Agriculture or ANZSIC 2006 A)								
Mining	0.027	0.352	-0.038	0.291	_	_	_	_
Manufacturing	0.172	0.000	0.056	0.015	_	_	_	_
Construction	0.096	0.000	0.116	0.001	_	_	_	_
Wholesale Trade	0.128	0.000	0.124	0.000	_	_	_	_
Retail Trade	-0.061	0.024	-0.135	0.000	_	_	_	_
Accommodation	-0.141	0.000	-0.103	0.000	_	_	_	_
Transport	0.114	0.000	0.089	0.007	_	_	_	_
Communication	0.225	0.000	0.084	0.020	_	_	_	_
Financial Services	0.414	0.000	0.172	0.000	_	_	_	_
Professional Services	0.281	0.000	0.193	0.000	_	_	_	_
Public Administration	0.258	0.000	0.211	0.000	_	_	_	_
Education	0.011	0.705	-0.043	0.041	_	_	_	_
Health	0.153	0.000	0.042	0.036	_	_	_	_

Coupled people: HES09-13 (continued)

	Wage Equation				Selection Equation			
	Coupled Men		Coupled Women		Coupled Men		Coupled Women	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Occupation (reference group is Labourer)								
Manager	0.277	0.000	0.392	0.000	_	_	_	_
Professional	0.279	0.000	0.371	0.000	_	_	_	_
Technical Worker	0.099	0.000	0.109	0.000	_	_	_	_
Personal Services	0.007	0.809	0.001	0.967	_	_	_	_
Administration	0.091	0.001	0.157	0.000	_	_	_	_
Sales	0.064	0.027	0.082	0.002	_	_	_	_
Machine Operator	-0.003	0.892	0.001	0.979	_	_	_	_
Region (reference group is Auckland)								
North North Island	-0.097	0.000	-0.090	0.000	0.219	0.158	0.075	0.211
Central North Island	-0.089	0.000	-0.074	0.000	0.381	0.040	0.076	0.250
Wellington	0.018	0.230	0.000	1.000	0.214	0.170	0.264	0.000
Canterbury	-0.092	0.000	-0.050	0.001	0.751	0.000	0.130	0.028
Rest of South Island	-0.082	0.000	-0.073	0.000	0.255	0.122	0.160	0.008
Other variables								
Year (trend)	-0.010	0.012	-0.006	0.453	0.040	0.293	-0.022	0.271
Unemployment Rate	0.00	0.901	-0.000	1.000	-0.139	0.086	0.021	0.498
Constant	1.934	0.000	2.136	0.000	-2.803	0.003	-1.278	0.000
Hhld Non-labour income (thousands)	_	_	_	_	-0.547	0.001	-0.400	0.000
Partner participation	_	_	_	_	1.886	0.008	0.392	0.000
Partner income (thousands)	_	_	_	_	1.064	0.111	-0.126	0.000
Variance Function								
Age÷10	-	-	-	-	0.063	0.020	-	-
Number of kids	-	-	-	-	0.092	0.013	-	-
Partner participation	-	-	-	-	0.562	0.001	-	-
Partner income (thousands)	-	-	-	-	0.301	0.013	0.052	0.038
Hhld Non-labour income (thousands)	-	-	-	-	0.295	0.000	0.241	0.001
Model evaluation								
Mills Ratio	0.041	0.023	0.142	0.000	_	_	_	_
R-squared	0.387	_	0.403	_	_	_	_	_
Pseudo R-squared	_	_	_	_	0.087	_	0.135	_
LR test for Heteroskedasticity	_	_	_	_	156.62	0.000	94.89	0.000

References

[1] Ball, C. and Creedy, J. [2015], 'Inequality in New Zealand 1983/84 to 2013/14', *Treasury Working Papers* (6), 1–33.

[2] Bargain, O. [2010], 'Back to the future: Decomposition analysis of distributive policies using behavioural simulations', *Institute for the Study of Labour: Discussion Paper Series* (5226), 1–18.

[3] Cameron, A. C. and Trivedi, P. K. [2005], *Microeconometrics: Methods and Applications*, Cambridge University Press.

[4] Creedy, J., Duncan, A., Scutella, R. and Harris, M. [2000], 'Wage functions for demographic groups in Australia', *Australian Journal of Labour Economics* **4**(4), 296–316.

[5] Creedy, J. and Mok, P. [2015], 'Labour supply in New Zealand and the 2010 tax and transfer changes', *New Zealand Treasury Working Paper* (13).

[6] Gould, W. [1998], 'Hetprob: Stata module to estimate heteroskedastic probit model'.

[7] Greene, W. H. [2011], *Econometric Analysis*, Pearson.

[8] Harvey, A. [1976], 'Estimating regression models with multiplicative heteroscedasticity', *Econometrica* **44**(3), 461–65.

[9] Heckman, J. [1979], 'Sample selection bias as a specification error', *Econometrica* **47**(1), 153–161.

[10] Kalb, G. and Scutella, R. [2003*a*], *New Zealand Treasury Working Paper* (23).

[11] Kalb, G. and Scutella, R. [2003*b*], 'Wage and Employment Rates in New Zealand from 1991 to 2001', *New Zealand Treasury Working Paper* (13).

[12] Maddala, G. S. [1983], *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

[13] McFadden, D. L. [1974], *Frontiers in Econometrics*, Academic Press, New York, chapter Chapter 4: Conditional logit analysis of qualitative choice behavior, pp. 105–142.

[14] Mercante, J. and Mok, P. [2014*a*], 'Estimation of labour supply in New Zealand', *New Zealand Treasury Working Paper* (8).

[15] Mercante, J. and Mok, P. [2014*b*], 'Estimation of wage equations for New Zealand', *New Zealand Treasury Working Paper* (9).

[16] NZQA [2017], 'New Zealand qualifications framework'.

[17] Statistics New Zealand [2001], 'The introduction of integrated weighting to the 2000/2001 Household Economic Survey'.

[18] Statistics New Zealand [2006], 'Business demographic statistics review'.

[19] Statistics New Zealand [2017*a*], 'ANZSIC06 industry classifications'.

[20] Statistics New Zealand [2017*b*], 'ANZSIC06 occupation classifications'.

[21] Treasury [1996], 'Briefing to the incoming government 1996', *Briefing to the Incoming Government 1996* .

[22] Woldridge, J. M. [2010], *Econometric Analysis of Cross Section and Panel Data*, 2 edn, MIT Press.

[23] Wooldridge, J. [2014], 'Test for heteroskedasticity in logit/probit models'.

About the Author

Matt Nolan is a PhD Student at the School of Economics and Finance at Victoria University of Wellington, New Zealand.
Email: Matt.Nolan@vuw.ac.nz

VICTORIA UNIVERSITY OF WELLINGTON
TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

Chair in Public Finance
Victoria Business School

Working Papers in Public Finance