New Zealand Journal of Employment Relations, 44(1): 1-20

An empirical examination of the gender pay gap in New Zealand

GAIL PACHECO*, CHAO LI** and BILL COCHRANE***

Abstract

New Zealand has often been described as a leader in the field of gender equality. Yet, while women have achieved substantial gains in a range of outcomes (education and labour force participation for example), the gender pay gap has changed very little. This study uses confidentialised microdata from Statistics New Zealand to examine the gap in a multitude of ways. We begin by applying the traditional Oaxaca decomposition technique, before accounting for selection, distributional differences and matching. We find that the gap is largely unexplained (83 per cent). Importantly, we correct for selection bias for both men and women – which produces counterbalancing effects such that the net result is broadly similar to that prior to the correction. We also employ propensity score matching, as a further check of robustness of results, and find only minor movements in the unexplained gap. Finally, distributional analysis illustrates evidence in favour of the glass-ceiling hypothesis.

Keywords: Gender, wage gap, selection, quantile, matching
JEL classification: J16; J3

1. Introduction

New Zealand led the world as the first country where women achieved the right to vote, and there have been substantial gains in recent years for women in a range of outcomes, such as education; labour force participation; and health. However, the gender pay gap has not diminished in the last decade – particularly if we compare the gap between the last time there was substantive analysis in this space for the New Zealand labour market (Dixon, 2003) and the end of the sample timeframe of this study 2015. Over this period of 12 years, the gender pay gap has hovered at the 12 per cent mark. The main aim of this research

---

* Director, New Zealand Work Research Institute, Auckland University of Technology. Corresponding author: gail.pacheco@aut.ac.nz
** Research Fellow, University of Waikato
*** Senior Lecturer, School of Social Sciences, University of Waikato

Acknowledgements:
We would like to thank Sylvia Dixon for her peer review feedback and suggestions that she provided on an earlier version of this study. We are also grateful to a number of others for providing us with helpful comments at various conferences and forums – including the team at the Ministry for Women; Statistics New Zealand's microdata team; and the New Zealand Association of Economists. We also wish to acknowledge and thank the Ministry for Women for sponsoring this research. Any errors or omissions remain the responsibility of the authors.

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 authors, not Statistics New Zealand.

is, therefore, to explore whether the factors that contribute/explain the gap have changed over time (given the transformation in other aspects of the labour market by gender), and what proportion of the gap can now be explained by observable information about the individual and their job. While the raw gender pay gap is regularly reported in the mainstream media (based on average or median earnings for males and females), it is meaningless without controlling for differences in characteristics. These include individual, household, occupation, industry and other job characteristics of the person. Of particular note is that there has been no gender pay gap analysis, controlling for relevant observable information post-2003.

This study relies on the use of unit record Income Survey data from Statistics New Zealand to estimate the gender pay gap while controlling for a wide range of observable characteristics. We employ the standard Oaxaca-Blinder decomposition technique which apportions to the gap into two components – explained and unexplained. The explained component reflects differences in the observed characteristics of males and females; while the unexplained reflects differences in returns. The latter is more problematic to interpret as these differences may be due to unobserved variables, discrimination, and/or different preferences for non-wage components by gender.

As the Oaxaca-Blinder approach may suffer from sample selection bias (wages can only be observed for the employed), we also apply the Heckman procedure to correct for this bias. This provides a predicted pay gap under the scenario that both males and females not in the labour force select into the labour market. We, then, switch to the semi-parametric approach of propensity score matching (PSM), which offers an alternative approach to test the reliability of our decomposition results. The PSM approach matches males and females based on their observed characteristics – this includes all personal, educational, household, region, occupation, industry and other job characteristics used in the earlier Oaxaca decompositions. The wages of the matched male observations provide the counterfactual wage for females, based on the returns to the characteristics that males are receiving.

We end our empirical endeavours with an assessment of how the gender pay gap changes across the wage distribution. In particular, we employ quantile regression to investigate how the gender pay gap (along with the proportion that is explained/unexplained) varies across various wage quantiles.

The format for the rest of the paper is as follows: A background on the New Zealand literature in this space is provided in Section two; Data and variable selection is described in Section three; The decomposition analysis, along with results corrected for sample selection bias are shown in Sections four and five; The PSM findings (as a test of robustness of our findings) are illustrated in Section six; while Section seven covers the quantile regressions; and the final section concludes.

2. New Zealand literature

There has been limited empirical work on understanding the gender wage ratio in New Zealand. Kirkwood and Wigbout (1999) made use of the first wave (1997) of the Income Survey (IS) via ‘tree analysis’ and found that about half of the earnings gap between men and women in full time employment could be explained by observed characteristics (such as education, occupation, ethnicity, marital status, etc.). The IS was added in 1997 as an annual supplement to the June quarter of the Household Labour Force Survey (a quarterly survey of around 15,000 households that began in March 1986).

The most substantial contributor to the New Zealand literature has been Sylvia Dixon. Her earliest work (1996ab; 1998) used the Household Economic Survey (HES) to investigate the distribution of earnings in New Zealand. She estimated Ordinary Least Squares (OLS) regressions with log of real hourly earnings as the dependent variable and generally found a significant gender wage differential. For instance, Dixon (1996b) found that, in 1995, the predicted earnings of females was 9.6 per cent lower than males, after controlling for other factors – which included age, age squared, educational characteristics, ethnicity, and part-time status.

In Dixon (2000), more covariates were added to the gender pay gap analysis (such as occupation and industry), analysis was extended to include the 1997 IS (in addition to the 1997 and 1998 HES), and wage regressions were replaced with decomposition frameworks. Such techniques (first developed by Oaxaca (1973) and Blinder (1973)) are now a standard method employed in the pay inequality literature; and apportion the pay gap either into endowments, characteristics, and residuals (a three-fold decomposition) or explained and unexplained (a two-fold approach).

In Dixon (2000), the total log hourly earnings gap equated to 15.3 per cent when using the HES (1997-1998), and 17.1 per cent when using the IS (1997). Initially, it was found that between 30 and 60 per cent of the gender wage differential could be attributed (explained) to differences in education and experience; and after information on occupation and industry was added to the model – the explained component rose to between 40 to 80 per cent. Dixon (2000) expected the pay gap to narrow in future years due to improvements in relative educational attainment of females, as well as the long-run expectation that male and female paid work patterns would gradually become more alike. Indeed, in a follow up paper, Dixon (2003) did find that the total gender pay gap had narrowed to 12.8 per cent. She argued that the decline in the pay gap was primarily driven by increases in the human capital of females (relative to males) and changes in the employment distribution of the two groups.

A final study worth mentioning is that by Alexander, Genç and Jaforullah (2004). They also made use of IS data, from both its inaugural year in 1997 and 2003. They estimated wage regressions via OLS, Heckit and Maximum Likelihood Estimation (MLE) – the latter two taking into account sample selection bias. Dixon (2000) did trial correcting for sample selection bias in her study as well, but found that the selection effect estimates were often insignificant, and very sensitive to the exclusion or inclusion of alternative ‘identifying’ variables. Alexander et al., (2004) found that, regardless of the estimation technique (i.e. whether correcting for selection or not), similar gender wage differentials were found – with approximately a 13 per cent gap in 1997 and 12 per cent gap in 2003.

To sum up the literature on examining the gender pay gap in New Zealand, while there appears to have been a flurry of estimations of the gap when the IS was first introduced in 1997, there has been no substantive analysis of the gender wage differential using data post-2003. We, therefore, provide a much-needed update using the latest available data in 2015.

3. Data and descriptives

Data

This study continues the tradition of using IS data in New Zealand when analysing pay inequality. This data source provides earnings information for a representative sample of approximately 15,000 New Zealand households (corresponding to roughly 30,000 individuals). The survey asks for details on, among other things, the respondent's pay and working hours. Where necessary, responses are imputed from a random donor of similar characteristics, with an 11.1 per cent imputation rate in the 2015 iteration[1]. Information from the accompanying Household Labour Force Survey (HLFS) provides a detailed picture of the labour force in terms of geographic, demographic, occupational, industry, and other job characteristics.

We limit our 2015 sample to the working age population (i.e. those aged 16-64) and drop a small number of wage earners with very low or high values for earnings and/or hours[2] to minimise the potential for measurement error influencing our estimates. We also exclude the self-employed – leaving us with a final sample of 13,737 (6834 males and 6903 females[3]).

Descriptives

Figures 1 and 2 show kernel density curves for the distribution by gender of log hourly wages and weekly hours worked, respectively. In Figure 1 the curve for the female wage distribution is steeper and higher than that for males, meaning that it is more clustered around the point of maximum density than the male distribution, which is more uniform. This indicates that male wages are more distributed over a range of values, relative to their female counterparts in the workforce. The point of maximum density is also further to the left for the female distribution, relative to the male distribution. This finding is reinforced by the lower value for the median usual hourly wage for women compared to men ($21 versus $24.21).

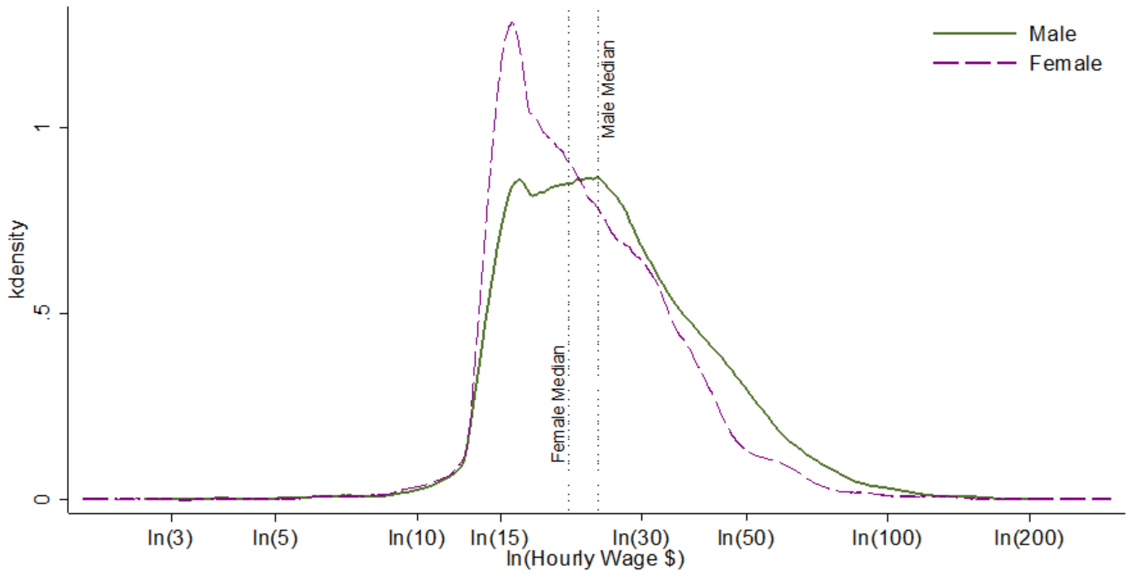
---

[1] See http://archive.stats.govt.nz/browse_for_stats/income-and-work/Income/NZIncomeSurvey_HOTPJun15qtr/%20Data%20Quality.aspx

[2] Specifically, we follow Dixon's (2003) thresholds of excluding hourly wage <$1 and >$500, and inflate these figures to 2015$. This mostly removes employed individuals who report zero wages, which may indicate a misclassification of their employment status. We also drop individuals reporting weekly hours in excess of 100.

[3] All sample sizes are random rounded to base three, due to Statistics New Zealand requirements regarding confidentiality assurance. Also, included in our sample are all imputed records.

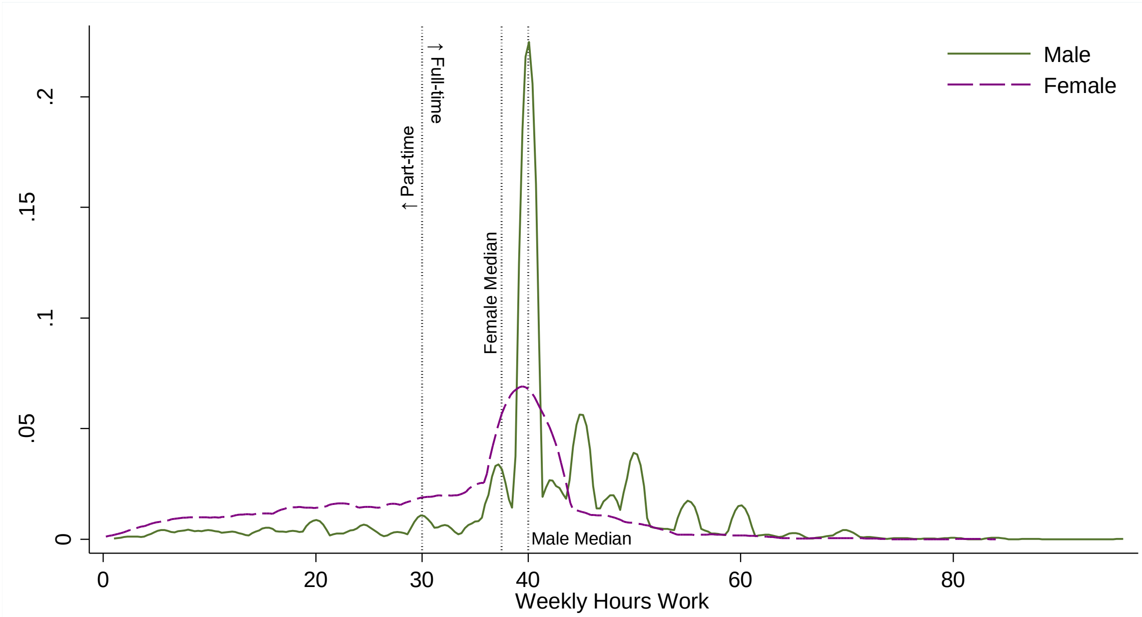
Figure 1: Usual hourly wage distribution, by gender (2015)



Source: 2015 IS. Author's compilation.

For usual weekly hours worked (Figure 2), the male curve is considerably steeper and higher than the female curve, suggesting more clustering in our male series. The density for females is flatter relative to that for males, indicating that weekly hours worked by females are more distributed in comparison to the clustering of male hours around the 40-hour mark. The female distribution also lies to the left of the distribution for males; median weekly hours worked are 37.5 for females and 40 for males.

Figure 2: Usual weekly hours distribution, by gender (2015)



Source: 2015 IS. Author's compilation.

A comprehensive descriptive portrait of the individuals in our sample is provided in Table 1. The 13,737 sample is fairly evenly split between the genders and the means and standard deviations for all variables used in the forthcoming analysis are provided for both the full sample, as well as separately for males and females. The final column in the table reflects whether the differences in characteristics between the gender sub-groups are statistically significant.

As Table 1 indicates, female employees, on average, receive a lower wage than their male counterparts – $25 per hour compared to $29 per hour. In terms of personal characteristics, females in our sample are marginally older, and there are minor differences in the ethnic makeup across the genders (with a little more Māori females than males; and a little less Asian females than males in the workforce). There are more marked differences across the genders with respect to household characteristics – with females close to three times as likely to be a sole parent and twice as likely to be widowed or separated/divorced, relative to males. It also appears that males in 2015 were more likely to be living in a household with children under the age of 6 (26.1 per cent of males compared to 18.1 per cent of females).

In previous studies, differences in educational attainment between males and females have often been found to be a contributing factor in explaining pay gaps – past evidence has usually found higher levels of educational attainment for males. However, based on Table 1, and as Dixon (2000) predicted, the educational divide in New Zealand has narrowed considerably – with females overtaking males in all qualification levels (barring post-school – which encompasses many vocational certificates and diplomas). Males are more likely to have no qualifications (16.3 per cent versus 14.2 per cent), and less likely to have a Bachelor's and postgraduate qualification. We can compare the figures in Table 1 to those reported by Dixon (2000) which used the 1997 wave of the IS. This comparison shows that in the 1997 sample, 14.3 per cent of males held a Bachelor's or postgraduate degree as their highest level of educational attainment (and 12.4 per cent of females); while in the 2015 sample, the comparable proportions were 22.5 per cent and 30.5 per cent, respectively.

In terms of occupational structure, males are more likely to be managers, trades workers, machinery operators or labourers; and women are more likely to be professionals, community and personal service workers, or in administration roles. There are also significant differences in gender distribution for the majority of the industry categories. Manufacturing and Construction, for example, appear to be male dominated; while Retail Trade and Education and Training appear to be female dominated sectors. Besides occupation and industry, one other job related characteristic provided in Table 1 is a dummy variable for working part-time. Females appear to be more than three times more likely to work part-time compared to males, 30.4 per cent versus 8.8 per cent.

Table 1: Variable definitions and descriptive statistics

Variable	Definition	Mean (Standard Deviation)			
		Full Sample	**Male**	**Female**	**Significant difference**
Hourly wage	Usual hourly total earnings ($)	27.0 (15.7)	29.0 (16.8)	25.0 (14.1)	***
Ln hourly wage	Natural logarithm of hourly usual total earnings	3.18 (0.45)	3.25 (0.47)	3.12 (0.42)	***
Weekly hours	Weekly usual total hours work	36.7 (12.5)	40.9 (11.1)	32.5 (12.5)	***
Personal characteristics					
Age	Age in years	41.2 (13.0)	40.7 (13.0)	41.7 (12.9)	***
Pakeha	Dummy variable: 1 = Pakeha; 0 otherwise	0.743 (0.437)	0.739 (0.439)	0.746 (0.436)	
Māori	Dummy variable: 1 = Māori; 0 otherwise	0.117 (0.321)	0.111 (0.314)	0.123 (0.328)	**
Pacific	Dummy variable: 1 = Pacific; 0 otherwise	0.067 (0.251)	0.069 (0.254)	0.065 (0.247)	
Asian	Dummy variable: 1 = Asian; 0 otherwise	0.115 (0.319)	0.120 (0.326)	0.109 (0.312)	**
MELAA	Dummy variable: 1 = MELAA; 0 otherwise	0.009 (0.092)	0.008 (0.092)	0.009 (0.093)	
Other ethnicity	Dummy variable: 1 = Other ethnicity; 0 otherwise	0.020 (0.141)	0.018 (0.135)	0.022 (0.147)	
Non-immigrant	Dummy variable: 1 = Born in NZ; 0 otherwise	0.726 (0.446)	0.719 (0.449)	0.732 (0.443)	*
Immigrant - Pasifika	Dummy variable: 1 = Born in Pacific Island countries; 0 otherwise	0.055 (0.227)	0.057 (0.231)	0.052 (0.223)	
Immigrant – Asia, Middle East, Africa	Dummy variable: 1 = Born in Asian, Middle East or African countries; 0 otherwise	0.098 (0.298)	0.104 (0.306)	0.092 (0.290)	**
Immigrant - Other	Dummy variable: 1 = Born in other countries (not listed above); 0 otherwise	0.121 (0.326)	0.120 (0.324)	0.123 (0.328)	
Household characteristics					
Joint parent	Dummy variable: 1 = Couple with one or more dependent children; 0 otherwise	0.345 (0.475)	0.382 (0.486)	0.307 (0.461)	***
Sole parent	Dummy variable: 1 = One parent with one or more dependent children; 0 otherwise	0.059 (0.236)	0.030 (0.172)	0.087 (0.283)	***
Children under 6	Number of children aged under 6 in the family	0.221 (0.549)	0.261 (0.600)	0.181 (0.491)	***
Children 6 - 14	Number of children aged 6-14 in the family	0.364 (0.744)	0.362 (0.749)	0.365 (0.738)	
Children 15 - 18	Number of children aged 15-18 (and not in full-time employment) in the family	0.102 (0.342)	0.092 (0.328)	0.111 (0.355)	***
Married/partnered	Dummy variable: 1 = Married/living as married; 0 otherwise	0.649 (0.477)	0.675 (0.468)	0.624 (0.484)	***
Widowed/separated/Divorced	Dummy variable: 1 = Widowed/separated/divorced; 0 otherwise	0.069 (0.254)	0.043 (0.203)	0.095 (0.294)	***
Never married	Dummy variable: 1 = Never married; 0 otherwise	0.281 (0.449)	0.282 (0.450)	0.280 (0.449)	
Educational attainment (highest qualification)					
No qualification	Dummy variable: 1 = No qualification; 0 otherwise	0.152 (0.359)	0.163 (0.369)	0.142 (0.349)	***
School	Dummy variable: 1 = Lower/upper secondary school qualification; 0 otherwise	0.243 (0.429)	0.230 (0.421)	0.255 (0.436)	***
Post school	Dummy variable: 1 = Post school qualification (level 1-7 certificate or diploma); 0 otherwise	0.339 (0.474)	0.381 (0.486)	0.298 (0.457)	***
Bachelor's	Dummy variable: 1 = Bachelor's degree (including Honours); 0 otherwise	0.180 (0.385)	0.153 (0.360)	0.207 (0.405)	***
Postgraduate	Dummy variable: 1 = Postgraduate qualification; 0 otherwise	0.085 (0.279)	0.072 (0.259)	0.098 (0.298)	***
Occupational Characteristics (ANZSCO Level 1)					
Dummy variables (8)	1 = Manager; 0 otherwise	0.130 (0.336)	0.170 (0.376)	0.090 (0.285)	***
	1 = Professional; 0 otherwise	0.238 (0.426)	0.195 (0.396)	0.280 (0.449)	***
	1 = Technician and Trades Worker; 0 otherwise	0.124 (0.329)	0.199 (0.400)	0.049 (0.215)	***
	1 = Community and Personal Service Worker; 0 otherwise	0.098 (0.297)	0.054 (0.225)	0.141 (0.348)	***

	1 = Clerical and Administrative Worker; 0 otherwise	0.132 (0.339)	0.063 (0.243)	0.201 (0.401)	***
	1 = Sales Worker; 0 otherwise	0.098 (0.298)	0.073 (0.261)	0.123 (0.328)	***
	1 = Machinery Operator or Driver; 0 otherwise	0.065 (0.247)	0.112 (0.315)	0.020 (0.138)	***
	1 = Labourer; 0 otherwise	0.115 (0.319)	0.134 (0.341)	0.097 (0.295)	***
Industry Classifications (ANZSIC Level 1)					
Dummy variables (19)	1 = Agriculture, Forestry and Fishing; 0 otherwise	0.045 (0.206)	0.060 (0.238)	0.029 (0.168)	***
	1 = Mining; 0 otherwise	0.004 (0.061)	0.006 (0.080)	0.001 (0.034)	***
	1 = Manufacturing; 0 otherwise	0.131 (0.337)	0.187 (0.390)	0.075 (0.264)	***
	1 = Electricity, Gas, Water and Waste Services; 0 otherwise	0.012 (0.108)	0.018 (0.131)	0.006 (0.079)	***
	1 = Construction; 0 otherwise	0.073 (0.260)	0.126 (0.332)	0.020 (0.141)	***
	1 = Wholesale Trade; 0 otherwise	0.041 (0.199)	0.056 (0.230)	0.027 (0.161)	***
	1 = Retail Trade; 0 otherwise	0.103 (0.304)	0.088 (0.283)	0.118 (0.322)	***
	1= Accommodation and Food Services; 0 otherwise	0.053 (0.224)	0.040 (0.196)	0.065 (0.247)	***
	1= Transport, Postal and Warehousing; 0 otherwise	0.045 (0.207)	0.062 (0.241)	0.028 (0.166)	***
	1 = Information Media and Telecommunications; 0 otherwise	0.019 (0.137)	0.020 (0.140)	0.018 (0.133)	
	1 = Financial and Insurance Services; 0 otherwise	0.022 (0.147)	0.018 (0.132)	0.026 (0.160)	***
	1 = Rental, Hiring and Real Estate Services; 0 otherwise	0.013 (0.115)	0.013 (0.112)	0.014 (0.117)	
	1 = Professional, Scientific and Technical Services; 0 otherwise	0.056 (0.230)	0.050 (0.218)	0.062 (0.241)	***
	1 = Administrative and Support Services; 0 otherwise	0.038 (0.190)	0.042 (0.200)	0.034 (0.181)	**
	1 = Public Administration and Safety; 0 otherwise	0.054 (0.227)	0.047 (0.212)	0.062 (0.240)	***
	1 = Education and Training; 0 otherwise	0.109 (0.311)	0.069 (0.254)	0.148 (0.355)	***
	1 = Health Care and Social Assistance; 0 otherwise	0.109 (0.312)	0.037 (0.190)	0.180 (0.384)	***
	1 = Arts and Recreation Services; 0 otherwise	0.039 (0.193)	0.021 (0.143)	0.057 (0.231)	***
	1 = Other services; 0 otherwise	0.034 (0.182)	0.040 (0.196)	0.029 (0.167)	***
Region					
Dummy variables (12)	1 = Northland Regional Council; 0 otherwise	0.037 (0.189)	0.035 (0.183)	0.039 (0.195)	
	1 = Auckland Regional Council; 0 otherwise	0.278 (0.448)	0.282 (0.450)	0.275 (0.446)	
	1 = Waikato Regional Council; 0 otherwise	0.082 (0.274)	0.083 (0.277)	0.080 (0.272)	
	1 = Bay of Plenty Regional Council; 0 otherwise	0.059 (0.236)	0.058 (0.234)	0.061 (0.239)	
	1 = Gisborne/Hawke's Bay Regional Council; 0 otherwise	0.056 (0.230)	0.056 (0.229)	0.057 (0.231)	
	1 = Taranaki Regional Council; 0 otherwise	0.037 (0.189)	0.037 (0.188)	0.038 (0.190)	
	1 = Manawatu-Wanganui Regional Council; 0 otherwise	0.059 (0.235)	0.056 (0.231)	0.061 (0.240)	
	1 = Wellington Regional Council; 0 otherwise	0.103 (0.304)	0.102 (0.303)	0.105 (0.306)	
	1 = Nelson/Tasman/Marlborough/West Coast Regional Council; 0 otherwise	0.052 (0.223)	0.051 (0.220)	0.054 (0.226)	
	1 = Canterbury Regional Council; 0 otherwise	0.132 (0.339)	0.137 (0.344)	0.127 (0.333)	*
	1 = Otago Regional Council; 0 otherwise	0.065 (0.247)	0.065 (0.247)	0.065 (0.247)	
	1 = Southland Regional Council; 0 otherwise	0.038 (0.191)	0.038 (0.191)	0.038 (0.191)	
Other job-related characteristics					
Part-time	Dummy variable: 1= Part-time (working less than 30 hours a week); 0 otherwise	0.197 (0.398)	0.088 (0.284)	0.304 (0.460)	***
Sample size		**13,737**	**6,834**	**6,903**	

Notes: ***, ** and * reflect the significance of the differences between the male and female subgroups, at the one per cent, five per cent, and ten per cent level, respectively.

4. Decomposition

Given the observable characteristics detailed in the previous section, the next step of our empirical endeavour is to investigate which variables (or more accurately put – gender differences in variables) can explain the gender pay gap, and how much of the gap is left unexplained.

We use the common decomposition approach in the literature on gender pay disparities introduced by Oaxaca (1973) and Blinder (1973). The process is to, first, separately estimate the wage equations (using the natural logarithm of usual hourly wages) for males in (1) and females in (2) as[4]:

$$\ln(w_i^m) = \beta^m X_i^m + \varepsilon_i^m \qquad (1)$$

$$\ln(w_i^f) = \beta^f X_i^f + \varepsilon_i^f \qquad (2)$$

where *m* and *f* superscripts denote males and females, the *i* subscript denotes the i^{th} wage earner, and *w* stands for hourly wages. *X* represents vectors of explanatory variables, shown in Table 1 above, which includes information on personal, educational, regional and household characteristics, as well as occupation, industry, and other job-related characteristics.
The gender pay gap is calculated in (3) and decomposed in (4):

$$\overline{\ln(w^m)} - \overline{\ln(w^f)} = \widehat{\beta^m}\overline{X^m} - \widehat{\beta^f}\overline{X^f} \qquad (3)$$

$$\overline{\ln(w^m)} - \overline{\ln(w^f)} = \widehat{\beta^m}\left(\overline{X^m} - \overline{X^f}\right) + \left(\widehat{\beta^m} - \widehat{\beta^f}\right)\overline{X^f} \qquad (4)$$

where $\widehat{\beta}$ stands for the vector of coefficients estimated in the wage equations. The first term on the right hand side of (4) is the part of the gender pay gap that can be explained by male-female differences in average characteristics (based on the explanatory variables outlined in Table 1). This ‘explained’ component can also be further broken down to show the contribution of different groupings of characteristics to the overall gap (as shown in Table 2).

The second component on the right hand side of (4) is the part of the gender pay gap left unexplained. This reflects differences in the returns to characteristics in the labour market and is more problematic to interpret. The unexplained component may indicate there are unobservable differences in the quality of characteristics between males and females, or differences in preference for non-wage components of jobs across gender, or discrimination against females in the labour market. For instance, Weichselbaumer and Winter-Ebmer (2005) argue that lower investment in on the job training, or more flexible and lower occupation levels of women could be voluntary choices made by some women, and these will not be observed in the data at hand, but could be responsible for part of the unexplained component.

[4] There are two weighting schemes that are possible with such a decomposition. The first uses the male wage structure when valuing the characteristics of men and women, and the opposite is true for the second. We use the first, which is the more commonly reported one in the gender pay gap literature.

Table 2: Oaxaca decomposition (pooled), dependent variable = ln hourly wage

	Explained	Unexplained
Model (A): With only personal characteristics		
Overall pay penalty = 12.71% ***	-1.13%***	13.84%***
Model (B): Model (A) + educational attainment		
Overall pay penalty = 12.71% ***	-3.88%***	16.59%***
Model (C): Model (B) + occupation and industry sector controls + other job related characteristics		
Overall pay penalty = 12.71% ***	1.81%**	10.90%***
Model (D): Model (C) + regional characteristics		
Overall pay penalty = 12.71% ***	1.97%**	10.74%***
Model (E): Model (D) + household characteristics		
Overall pay penalty = 12.71% ***	2.15%**	10.56%***

Note: *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. N = 13,737

Five specifications are presented in Table 2 (labelled Models A through to E), each of which includes additional explanatory variables in an iterative fashion. Our decomposition analysis begins with Model A which includes only the personal characteristics listed in Table 1, i.e. age, ethnicity and migrant status (and we also include age squared to allow age to play a non-linear role in our specification). Model B then replicates this analysis with the addition of educational attainment variables. Model C includes the covariates from Model B and adds occupation and industry sector controls, as well as other job related information on part-time status of the individual. Model D then controls for regional characteristics in an additive manner. Finally, the decomposition analysis culminates with Model E which includes all the aforementioned variables and household characteristics regarding sole/joint parenthood; marital status; and information on the number and ages of dependent children in the household (again details of all these variables are listed in Table 1).

From Table 2, it can be seen that, regardless of the model used, the pay gap equates to a penalty of 12.71 per cent for females. With the exception of Model B, incorporating additional controls in to the specification results in a decline in the unexplained component of the gender pay gap from 13.84 per cent to 10.56 per cent, a drop of 3.28 percentage points or 23.7 per cent.

With respect to Model B, which contains controls for personal characteristics and educational attainment, female characteristics appear to be better than males, as shown by the negative explained component. However, this model yields the largest unexplained component of 16.59 per cent. The main driver of the unexplained component appears to stem from age, which is also a proxy for experience in this analysis. More specifically, the unexplained total of 16.59 per cent is made up the following elements: 30.53 per cent age, and age squared; -0.74 per cent other personal characteristics; -0.004 per cent education; and a constant of -13.20 per cent. Hence, it appears that males are receiving a much higher rate of return for age, relative to their female counterparts. It is also worth noting that adding job specific information to the mix (via inclusion of occupation, industry dummies, and an indicator for part time status) in Model C does little to quell the role of age in terms of the unexplained component.

Adding regional controls (shown by Model D) has a negligible effect on both the overall results for the explained and unexplained, as well as their sub-components. Finally, when household characteristics are added (i.e. moving from model D to E) the explained component rises marginally from 15.5 per cent to 16.9 per cent of the total gap (1.97 and 2.15 percentage points out of 12.71 per cent, respectively). To delve further into the drivers of the unexplained figure, we breakdown this component of the pay gap into the following sub-parts: 7.41 per cent age; -0.83 per cent other personal characteristics; 1.02 per cent education; -2.94 percent occupation, industry and part-time status; -0.38 per cent region; 5.62 per cent household characteristics; and a constant of 0.66 per cent. There is one result shown in these figures that is worth highlighting. The role of age appears to have diminished with the addition of household characteristics, i.e. controlling for differences in marital status and childcare responsibility has helped reduce the magnitude of the large unexplained positive returns for age/experience found for males (relative to females) in Models A through to D.

It is important to note that, even for the fullest specification (model E), characteristics and endowments still only account for 16.9 per cent of the gap, leaving just over 83 per cent unexplained.

How does this compare with the international literature? Such comparisons are fraught with difficulty (see for example, Blau and Kahn, 2001; 2016). However, Christofides, Polycarpou, & Vrachimis (2013) consider the pay gap across 26 European countries using the Oaxaca decomposition and data from the 2007 European Union Statistics on Income and Living Conditions. They find considerable heterogeneity in the size of the raw gender pay gap and also in the unexplained component of the gap. The percentage of the gap unexplained varies from values similar to what we find here for Denmark (74.2 per cent), Germany (75.8 per cent) and Norway (87.2 per cent) to those that are considerably higher (Poland – more than 100 per cent [5]) or lower (United Kingdom, 45.3 per cent).

Similarly, the OECD in its report “Closing the Gender Gap” (OECD, 2012) find considerable variation in the unexplained component of the gender pay gap, with the unexplained component varying from 15 per cent in Australia to 137 per cent in Slovenia.

As detailed earlier, the unexplained residual can encompass any unobserved differences in characteristics or preferences between males and females as well as discrimination against females in the labour market. Therefore, the “unexplained” cannot be unproblematically equated with the extent of labour market discrimination against females. Such unobservables include personality, attitudes, motivation, and ambition for example. While many of these will be difficult to quantify, one set of unobservables that could be included in future research is the subject studied by those that undertook bachelor’s qualifications or higher. For instance, recent research by Frölich (2007) finds that the subject of degree was an important variable in explaining gender wage differences in the United Kingdom. Future research could use the Income Survey linked with Ministry of Education data in the Integrated Data Infrastructure provided by Statistics New Zealand to include this explanatory variable.

[5] In this case, the explained portion is negative, indicating that female characteristics are better than male characteristics.

5. Correcting for selection bias

The Oaxaca-Blinder approach may suffer from sample selection bias (Heckman, 1979) as wages can only be observed for individuals that are employed. The decision to enter the labour force may be systematically correlated with potential wages, meaning that limiting our analysis to only the employed may result in biased estimates. For example, one of the traditional explanations for a gender pay gap is different levels of experience between the genders; therefore, to understand the potential drivers of this difference, we have to better understand the factors associated with the decision to participate in the labour market and accumulate experience. Additionally, females are more likely to change their participation decision during child bearing and rearing years, and it is important to take this into account. Another aspect of the participation decision that is also relevant here are education levels. Given that there are rising levels of education for both males and females in New Zealand (and at a faster rate for the latter), and knowing that education and wages are positively correlated, changing levels of educational attainment will likely affect the participation decision, and thus influence the pay gap.

To correct our estimates for sample selection bias, we apply the Heckman procedure and do this for both females and males[6]. The procedure requires one additional step before (1)-(4) above. This is to separately estimate probit models for males in (5) and females in (6) as:

$$LFP^m = \vartheta^m Z^m \tag{5}$$
$$LFP^f = \vartheta^f Z^f \tag{6}$$

where *m* and *f* superscripts denote males and females, respectively and the full HLFS sample is utilised, i.e. not restricting analysis to the waged employees, as was done in Section 4. In equations (5) and (6) *LFP* stands for labour force participation (with =1 for wage earners, unpaid workers or volunteer job takers, self-employed and unemployed, and =0 for those not in the labour force). *Z* represents vectors of explanatory variables shown in Table 1, except for occupation, industry and other job related characteristics. Then for each male in (7) and female in (8), the probability of participating in the labour force is predicted as:

$$\widehat{LFP_j^m} = \widehat{\vartheta_1^m} Z_{1j}^m + \widehat{\vartheta_2^m} Z_{2j}^m + \cdots + \widehat{\vartheta_k^m} Z_{kj}^m \tag{7}$$
$$\widehat{LFP_j^f} = \widehat{\vartheta_1^f} Z_{1j}^f + \widehat{\vartheta_2^f} Z_{2j}^f + \cdots + \widehat{\vartheta_k^f} Z_{kj}^f \tag{8}$$

where k and j subscripts denote the k^{th} explanatory variable and the j^{th} male or female in the sample, respectively.

A selection-correction parameter for each male in (9) and female in (10) is generated as:

$$mills_j^m = \frac{(normalden(-\widehat{LFP^m}))_j}{1-(normal(-\widehat{LFP^m}))_j} \tag{9}$$
$$mills_j^f = \frac{(normalden(-\widehat{LFP^f}))_j}{1-(normal(-\widehat{LFP^f}))_j} \tag{10}$$

[6] Most prior literature has only corrected for selection bias for females. There are a few exceptions of recent studies that control for sample selection for both genders. See Perugini and Selezneva (2015), and Christofides, Li, Liu and Min (2003).

where *normalden and normal* denote the standard normal density function and the cumulative normal distribution function, respectively. The selection-correction indices (inverse mills ratios) $mills^m$ for males and $mills^f$ for females are added as additional variables into the decomposition process shown in (1)-(4), giving the decomposition results corrected for selection bias. These results are provided in Table 3, and included in this table are the uncorrected estimates for the full specification from Table 2, for comparison purposes.

Table 3: Oaxaca decomposition of full specification – with and without adjustment for sample selection bias

	Not corrected	Correction for females	Correction for males	Correction for females and males
Explained	2.15%**	2.15%**	2.46%**	2.46%*
Unexplained	10.56%***	18.00%***	2.10%*	9.54%***
Total gap	12.71%***	20.14%***	4.56%**	12.00%***
Inverse mills ratio (std error)	N/A	0.205** (0.106)	0.373*** (0.085)	Females: 0.205** (0.106) Males: 0.373*** (0.085)

Note: Full specification = Model E. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. N = 13,737.

Table 3 presents three new specification results, the decomposition corrected for sample selection bias for females only, males only and, in the last column, for both genders. A useful way to conceptualise these results is as follows. The working age population (All) for each gender is made up of those in the labour force (Employed (E) and Unemployed (U)) and those not in the labour force (NILF). The first column which provides the uncorrected results of the decomposition compares the pay gap between E^f and E^m, where *f* and *m* superscripts denote females and males, respectively[7]. The second column compares the same pay gap under the scenario where the females that are NILF now join the labour force; the third column repeats this exercise but changes the scenario to that of NILF males joining the labour force; and then the final column provides the predicted pay gap if both genders that are NILF join the labour force.

For both males and females in Table 3 we find that the inverse mills ratio is positive and significant. This means that there is positive selection into the labour market - those participating in the labour force have favourable unobservable characteristics (relative to those not in labour force) that positively affect their wages[8].

When correcting for selection bias for only females (column 2), the predicted pay gap rises to 20.14 per cent This is the scenario where all NILF females join the labour force. This result is expected as including females not in the labour force (who have less favourable unobservable characteristics) in the comparison reduces the average predicted wage of this group, relative to employed males. In a similar fashion in column 3, correcting for selection bias for only males reduces the predicted pay gap to 4.56 per cent. Again, this means that males not in the labour force have less favourable unobservable characteristics compared to males in the labour force, as the unobservables that would increase the likelihood of

[7] Noting that we exclude employed individuals with zero wages.

[8] Note that while we cannot show that the NILF group have poorer unobservables relative to those in the labour force, we can check the observable characteristics as a potential proxy. For instance, we find that 29.4 per cent of the NILF group have no school qualifications, compared to the 15.2 per cent of the group in the labour force. Additionally, 11.73 per cent of the NILF group have a bachelor's qualification or higher, compared to 23.29 per cent of the group in the labour force.

participating in the labour force simultaneously increase the likelihood of higher predicted wages. Including these males in the comparison substantially reduces the predicted gender pay gap.

Finally, as the last column in Table 3 shows, correcting for sample selection bias for both females and males reduces the gap marginally from 12.71 per cent (when uncorrected for both genders) to 12.00 per cent This decrease indicates that the males out of the labour force have slightly less favourable unobservables compared to the females out of the labour force. Once the selection adjustment has been taken into account, the explained proportion of the pay gap rises a little to 20.50 per cent (which corresponds to 2.46 percentage points out of 12.00 per cent).

6. Matching

Another way of assessing the gender pay gap in New Zealand is to apply the semi-parametric technique of propensity score matching (PSM). We follow Frölich (2007), who argues that the functional form assumptions inherent in the parametric Oaxaca decomposition may potentially give misleading results (see Barsky, Bound, Charles, & Lupton (2002); Mora (2008), and Ñopo (2008) for further discussion of this). In contrast, PSM does not specify linear regression functions and only simulates the adjusted mean wages for the common support subpopulation (Frölich, 2007). This distinguishes PSM from Oaxaca and allows PSM to serve as an alternative approach to test the reliability of our initial results from the Oaxaca decomposition. The process is to first estimate a probit model for males and females together:

$$f_i = \vartheta X_i \qquad (11)$$

where f_i is the gender dummy equal to 1 for female observations and 0 for males, and ϑ is a vector of coefficients. X_i is a vector of control variables that are the same ones as those in Oaxaca equations (1) and (2). The probability of being female, namely the propensity score, for each observation of males and females is predicted as:

$$\widehat{p_i} = \hat{\vartheta} X_i \qquad (12)$$

where $\hat{\vartheta}$ is a vector of the estimated coefficients from equation (11). Male observations are then matched to the female observations who have exactly the same (or the closest) propensity scores. Then, the wages (or average of those wages) of those matched male observations are assigned to those female observations. This provides a counterfactual for females' observations of the potential wage they would receive if they experienced the same wage returns to their characteristics that males are receiving.

The pay gap can then be broken down into explained and unexplained components. The unexplained part encompasses the difference between the females' mean counterfactual wages and the females' mean actual wages. This corresponds to $\left(\widehat{\beta^m} - \widehat{\beta^f}\right)\overline{X^f}$ in the Oaxaca decomposition from equation (4). The explained component reflects the difference between the males' mean actual wages and the females' mean counterfactual wages and corresponds to $\widehat{\beta^m}\left(\overline{X^m} - \overline{X^f}\right)$ from the Oaxaca decomposition in equation (4).

It is useful to point out that by applying PSM to the decomposition of the gender pay gap, the females' counterfactual wage is not estimated (unlike Oaxaca), but assigned using the matched males' actual wages. The results are provided in Table 4.

Table 4: PSM decomposition of full specification – with and without adjustment for sample selection bias

	Not corrected	Correction for females	Correction for males	Correction for females and males
Explained	3.54%***	3.54%***	4.26%***	4.27%***
Unexplained	9.17%***	16.60%***	0.30%	7.73%***
Total gap	12.71%***	20.14%***	4.56%**	12.00%***

Note: Variables used in PSM stem from the full specification of Model E. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. N = 13,737

As shown in Table 4, the gender pay gap is unchanged from Table 3 – as it is based on mean wage outcomes for each gender. What has changed a little is the proportion of the gap accounted for by our observable characteristics for the individual, household, region, industry, occupation and part-time status.

Focussing on just the final column (which are the results after the adjustment for sample selection bias for both males and females), we can see that 4.27 percentage points out of 12 per cent is now explained by the explanatory variables, compared to Table 3 where the comparable proportion was 2.46 percentage points out of 12 per cent. This is a jump from 20.50 per cent to 35.58 per cent of the total gap. Nevertheless, the pay gap still remains dominated primarily by the unexplained component.

7. Quantile regression

The purpose of this section is to explore gender wage disparities at different points in the wage distribution. More specifically to investigate the existence of both "sticky floors" (where the disparity is greater at the lower end of the distribution) and "glass ceilings". The latter refers to "a greater earnings gap at the top end of the distribution" (Chi & Li, 2008, p.244).

In this final section, we undertake an unconditional quantile decomposition and follow the approach by Firpo, Fortin and Lemieux (2009), which is an improved version of the approach detailed in DiNardo, Fortin and Lemieux (1996)[9]. Other relevant literature utilising this approach includes Chi and Li (2008), Ahmed and Maitra (2015), and Barón and Cobb-Clark (2010).

The first step is to calculate the re-centred influence function (RIF) for each male observation and each female observation, for each quantile of their respective distributions:

$$RIF_i^{m,\tau} = w^{m,\tau} + \frac{(\tau - 1\{w_i^m \leq w^{m,\tau}\})}{DEN(w^{m,\tau})} \qquad (13)$$

$$RIF_i^{f,\tau} = w^{f,\tau} + \frac{(\tau - 1\{w_i^f \leq w^{f,\tau}\})}{DEN(w^{f,\tau})} \qquad (14)$$

where, i is the i^{th} observation, τ is the τ^{th} quantile of the log wage distribution (males' or females'), $w^{m,\tau}$ is the log wage at the τ^{th} quantile of the males' log wage distribution, $\mathbf{1}\{\ \}$ is an indicator function, which is equal to 1 if $w_i^m \leq w^{m,\tau}$ is true, otherwise 0. w_i^m is the i^{th} male's log wage, and $DEN(w^{m,\tau})$ is the density at the τ^{th} quantile of the males' log wage distribution.

[9] An alternative method is the conditional quantile decomposition (see Machado and Mata (2005); and Melly (2005)); however it is generally recognised as problematic to implement as it is computationally intensive (Chi & Li, 2008).

As shown in Firpo et al. (2009), we then follow the Oaxaca steps detailed in (1) to (4), and replace the dependent variable of log wage with the calculated RIF obtained from equations (13) and (14). Equations (15) and (16) are then run for each quantile:

$$RIF_i^{m,\tau} = \beta^m X_i^m + \varepsilon_i^m \quad (15)$$
$$RIF_i^{f,\tau} = \beta^f X_i^f + \varepsilon_i^f \quad (16)$$

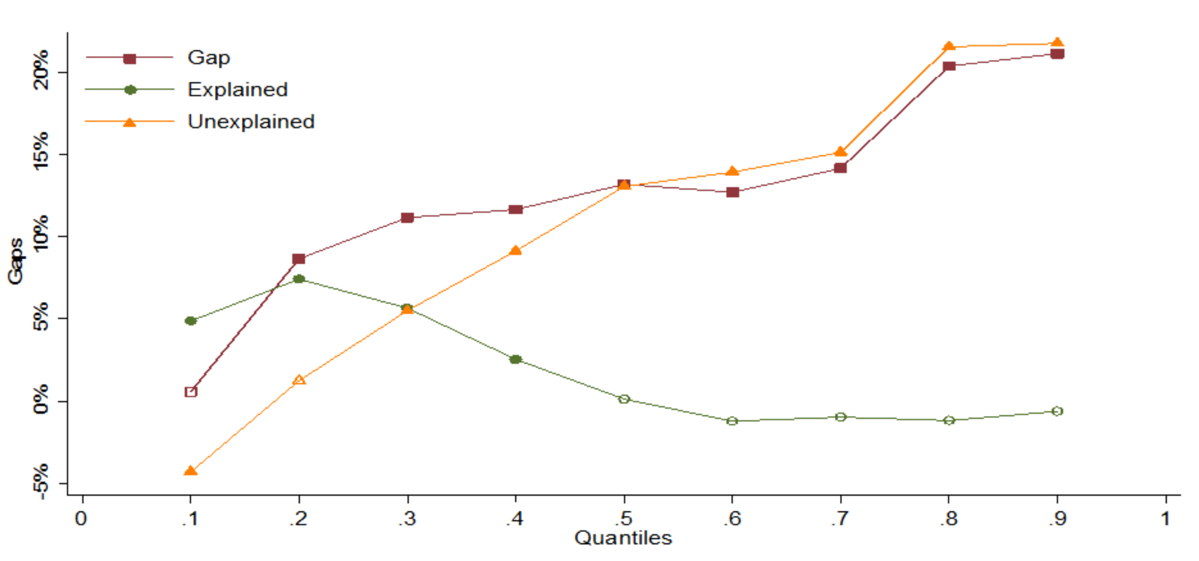
where *m* and *f* subscripts denote male and female wage earners, respectively, *w* stands for wages, and *X* represents vectors of explanatory variables shown in Table 1 above.

The gender pay gap at the τ^{th} quantile of the log wage distribution (males' or females') is calculated in (17) and decomposed in (18) as:

$$\overline{RIF^{m,\tau}} - \overline{RIF^{f,\tau}} = \widehat{\beta^m}\overline{X^m} - \widehat{\beta^f}\overline{X^f} \quad (17)$$
$$\overline{RIF^{m,\tau}} - \overline{RIF^{f,\tau}} = \widehat{\beta^m}\left(\overline{X^m} - \overline{X^f}\right) + \left(\widehat{\beta^m} - \widehat{\beta^f}\right)\overline{X^f} \quad (18)$$

where $\widehat{\beta}$ stands for the coefficients estimated in the RIF equations ((15)-(16)). The first term on the right hand side of (18) is the part of the calculated gender pay gap in (17) at the τ^{th} quantile of the log wage distribution (males' or females') that can be explained by male-female differences in means of those explanatory variables. The second term in (18) is the part of the gap left unexplained (reflecting differences in returns)[10].

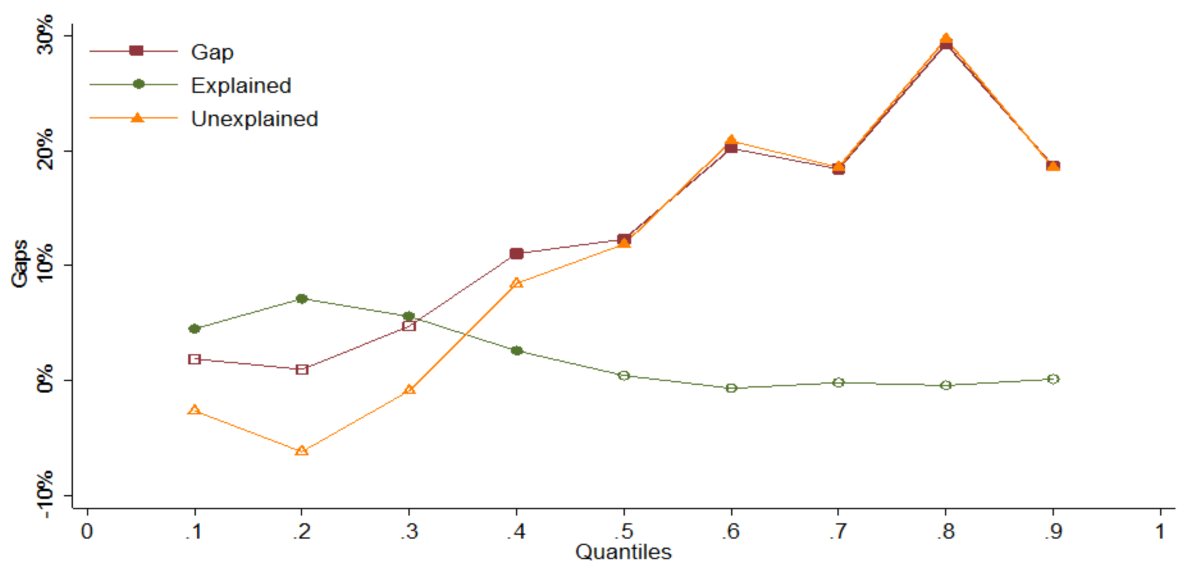
Figure 3: Gender pay gaps across wage distribution



Note: Hollow markers indicate insignificant gaps at 10 per cent significance level. Source: 2015 IS. Author's compilation.

[10] The specific stata command is rifreg which is available for download as an RIF-regression STATA ado file from Firpo, Fortin and Lemieux (2009): http://faculty.arts.ubc.ca/nfortin/datahead.html.

Figure 4: Gender pay gaps across wage distribution, with adjustment for sample selection bias



Note: Hollow markers indicate insignificant gaps at 10 per cent significance level. Source: 2015 IS. Author's compilation.

Figures 3 and 4 illustrate the total gender pay gap at the following wage cut-offs: 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th quantiles. These graphs also show the proportion of the gap that is explained and unexplained. Figure 3 provides the uncorrected results, while Figure 4 shows the results corrected for selection bias.

Regardless of which figure is viewed, there are a couple of noteworthy trends. First, the gender pay gap appears to increase as we move up the wage distribution – from zero per cent to 21.15 per cent in Figure 3 (when moving from the 10th to 90th quantile), and from zero per cent to 18.69 per cent in Figure 4. This is evidence in favour of the glass ceiling hypothesis. Other studies that also find strong evidence of only the glass ceiling and not the sticky floor include Kee (2006) for Australia; and Booth, Francesconi and Frank (2003) for the United Kingdom.

In a similar fashion to Kee (2006), we also find that the proportion of the gap that is unexplained rises as we move up the wage distribution. In particular, it is clear in both Figures 3 and 4 that the explained component tends to be statistically significant at the lower quantiles, and insignificant at the higher quantiles; while the reverse is true for the unexplained component. For instance, viewing the selection corrected results in Figure 4, we can see that the unexplained component is statistically insignificant for the first four quantiles; while the explained component then becomes statistically insignificant for the last five quantiles.

Why is there no evidence of a sticky floor? This is potentially due to the high relative minimum wage ratio in New Zealand. Most recent data from the OECD[11] (from 2013) shows that the minimum wage in New Zealand is 60 per cent of the median wage of full time employees. This high minimum wage ratio provides minimal room at the bottom of the wage distribution for sizable wage disparities by gender.

[11] See https://stats.oecd.org/Index.aspx?DataSetCode=RMW#

8. Conclusions

This research presents empirical evidence of the gender pay gap in New Zealand, based on Income Survey data from Statistics New Zealand. It is the first rigorous empirical look at the drivers of the pay gap since 2003. We find that the pay gap (based on 2015 data) is approximately 12 per cent and unchanged since this topic was last analysed in 2003. Additionally, regardless of the approach undertaken, the majority of the gap appears to be unexplained. This result persists even after correcting for selection bias.

While the unexplained portion dominates the pay gap, it is not a simple concept to tackle. The unexplained could relate to unobservables (such as subject of qualification), or differences in preferences for non-pecuniary aspects of the job by gender, or unconscious bias, or discrimination, or all of the above. Additionally, our methodology cannot deal with differences between men and women in terms of access to endowments, such as training and promotion opportunities, and this could mean that our decomposition approach is underestimating the size of the gender pay gap in general.

This research also illustrates that the size of the gender pay gap depends heavily on the location in the wage distribution. There is strong evidence pointing to a glass ceiling effect in New Zealand. Future work could delve further into the drivers of this outcome. For instance, it may be related to the differential impacts (by gender) of parenthood on labour market outcomes, and it would be useful to follow the evolution of life-course earnings by gender in New Zealand to assess the point at which the gender pay gap emerges.

References

Ahmed, S., & Maitra, P. (2015). A distributional analysis of the gender wage gap in Bangladesh. *The Journal of Development Studies*, *51*(11), 1444-1458.

Alexander, R., Genc, M., & Jaforullah, M. (2004). The Gender Wage Ratio in New Zealand. In B. Mahy, R. Plasman, & F. Rycx (Eds.), *Gender Pay Differentials Cross-National Evidence from Micro-Data* (pp.165-187). New York: Palgrave Macmillan.

Barón, J. D., & Cobb-Clark, D. A. (2010). Occupational segregation and the gender wage gap in private- and public-sector employment: a distributional analysis. *Economic Record*, *86*(273), 227-246.

Barsky, R., Bound, J., Charles, K. K., & Lupton, J. P. (2002). Accounting for the black–white wealth gap: a nonparametric approach. *Journal of the American Statistical Association*, *97*(459), 663-673.

Blau, F., & Kahn, L. (2001). Understanding international differences in the gender pay gap. *NBER Working Paper 8200.* Retrieved from https://www.nber.org/papers/w8200

Blau, F., & Kahn, L. (2016). The gender wage gap: Extent, trends and explanations. *NBER Working Paper 21913*. Retrieved from https://www.nber.org/papers/w21913

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, *8*, 436-455.

Booth, A. L., Francesconi, M., & Frank, J. (2003). A sticky floors model of promotion, pay, and gender. *European Economic Review, 47,* 295-322.

Chi, W., & Li, B. (2008). Glass ceiling or sticky floor? Examining the gender earnings differential across the earnings distribution in urban China, 1987–2004. *Journal of Comparative Economics*, *36*(2), 243-263.

Christofides, L. N., Li, Q., Liu, Z., & Min, I. (2003). Recent two-stage sample selection procedures with an application to the gender wage gap. *Journal of Business & Economic Statistics*, *21*(3,: 396-405.

Christofides, L. N., Polycarpou, A., & Vrachimis, K. (2013). Gender wage gaps, 'sticky floors' and 'glass ceilings' in Europe. *Labour Economics, 21,* 86-102.

DiNardo, J., Fortin, N. M., & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, *64*(5): 1001-1044.

Dixon, S. (1996a). The Distribution of Earnings in New Zealand: 1984-94. *Labour Market Bulletin, 1,* 45-100.

Dixon, S. (1996b). The Distribution of Earnings in New Zealand: 1984-95. In *Proceedings of the Seventh Labour, Employment and Work Conference*, November 1996, Wellington, Victoria University of Wellington.

Dixon, S. (1998). Growth in the Dispersion of Earnings: 1984-97. *Labour Market Bulletin*, *1&2*, 71-107.

Dixon, S. (2000). *Pay Inequality between Men and Women in New Zealand* (Occasional Paper 2000/1). Labour Market Policy Group, Department of Labour: Wellington.

Dixon, S. (2003). *Understanding reductions in the gender wage differential: 1997 – 2003*. Paper presented at the New Zealand Conference on Pay and Employment Equity for Women, Wellington, 28-29 June 2004.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953-973.

Frölich, M. (2007). Propensity score matching without conditional independence assumption—with an application to the gender wage gap in the United Kingdom. *The Econometrics Journal*, *10*(2), 359-

Heckman, J. (1979). Sample selection bias as specification error. *Econometrica, 47*, 153-162.

Kee, H. J. (2006). Glass Ceiling or Sticky Floor? Exploring the Australian Gender Pay Gap. *The Economic Record*, *82*(259), 408-427.

Kirkwood, H., & Wigbout, M. (1999). *An exploration of the difference in income received from wages and salaries by women and men in full-time employment*, Statistics New Zealand, Wellington.

Machado, J. A., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, *20*(4), 445-465.

Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour economics*, *12*(4), 577-590.

Mora, R. (2008). A nonparametric decomposition of the Mexican American average wage gap. *Journal of Applied Econometrics*, *23*(4), 463-485.

Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *The review of economics and statistics*, *90*(2), 290-299.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, *14*, 693-709.

OECD (2012). *Closing the Gender Gap: Act Now*. OECD Publishing. Retrieved from http://dx.doi.org/10.1787/9789264179370-en

Perugini, C., & E. Selezneva (2015). Labour Market Institutions, Crisis and Gender Earnings Gap in Eastern Europe. *Economics of Transition*, *23*(3), 517-564.

Weichselbaumer, D., & Winter-Ebmer, R. (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys*, *19*(3), 479-511.