



# Working Paper No. 15

The Gender Pay Gap in the New Zealand Public Service





# The Gender Pay Gap in the New Zealand Public Service

by Michelle A Gosse\*

## ABSTRACT

This paper examines the issue of gender pay gap in the New Zealand Public Service. Using data obtained from the Human Resources Capability Survey, the paper concludes that while on average, male employees earn more than female employees in the Public Service, since the 1970s the gender pay gap has been decreasing. The “Blinder-Oaxaca decomposition” method is used to argue that industry and occupation have consistently been found to have major effects on the gap. Both factors tend to exhibit gender-segregation, with male employees tending to work in higher paid industries and occupations, with occupation having the largest effect. In particular, the paper finds that the gender pay gap was 16.61% in 2000, falling to 15.51% in 2001 although when differences in human capital factors and employment characteristics were taken into account, the gap decreased to 4.7% for both years. Other findings include: (a) where male and female employees worked in the same occupations, male employees tended to earn more; (b) on average, male employees were older and had longer tenure than female employees, although this difference in human capital had mixed effects; (c) male employees had a different distribution across employers compared to female employees, again with mixed effects; and (d) there were similar proportions of male and female employees in ethnic minorities, with male minority employees appearing to average slightly higher earnings compared to female minority employees.

\* Michelle Gosse is Advisor in the Strategic Development Branch of the State Services Commission. This paper was submitted by the author as a Research Report in partial fulfilment of the Diploma in Applied Statistics at Massey University, 2001.

This, and other, Working Papers are available through the State Services Commission's website:

<http://www.ssc.govt.nz>

The State Services Commission welcomes any comments and suggestions you have about this paper. If you require further information about this paper please let us know. In both instances please contact:

Derek Gill, Branch Manager  
Strategic Development Branch  
State Services Commission  
Email: [derek.gill@ssc.govt.nz](mailto:derek.gill@ssc.govt.nz)  
Facsimile: +64 4 495 6699

**Note:** Information contained in this paper may be reproduced for personal or in-house use without formal permission or charge. However, it **must not** be used in a commercial context unless written permission has been given by the State Services Commission. Where any material from this paper is quoted in any context, it must be sourced – preferably to the Commission's website.

© *Crown copyright reserved 2000*

Publication of the Working Papers Series recognises the value of developmental work in generating policy options. The papers in this series are prepared for the purpose of informing policy development. The views expressed are those of the authors and should not be taken to be the views of the State Services Commission. The SSC view may differ in substance or extent from that contained in this paper.

# Table of Contents

<b>EXECUTIVE SUMMARY .....</b>	<b>1</b>
<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1 OVERVIEW OF THE REPORT.....	1
1.2 THE GENDER PAY GAP .....	2
<b>2 THE NEW ZEALAND HUMAN RIGHTS AND EEO CONTEXT.....</b>	<b>3</b>
2.1 THE LEGISLATIVE FRAMEWORK .....	3
2.1.1 <i>Legislation Covering Pay Discrimination</i> .....	3
2.1.2 <i>State Sector Act 1988</i> .....	4
2.1.3 <i>Human Rights Act 1993</i> .....	4
2.1.4 <i>Summary of the Current New Zealand Position</i> .....	6
<b>3. INTERNATIONAL FINDINGS ON GENDER DIFFERENCES IN PAY .....</b>	<b>6</b>
3.1 UNITED STATES .....	6
3.2 UNITED KINGDOM .....	6
3.3 AUSTRALIA .....	7
<b>4. NEW ZEALAND FINDINGS ON GENDER DIFFERENCES IN PAY.....</b>	<b>8</b>
4.1 EMPLOYMENT LEGISLATION CONTEXT .....	8
4.2 GENERAL LABOUR MARKET FINDINGS .....	9
4.3 INCOME INEQUALITY .....	10
4.4 FUTURE PREDICTIONS .....	10
4.5 THE GENDER PAY GAP IN THE NEW ZEALAND PUBLIC SERVICE .....	10
<b>5 FINDINGS FROM BLINDER-OAXACA DECOMPOSITIONS .....</b>	<b>11</b>
5.1 THE ORIGINAL DECOMPOSITIONS .....	11
5.2 EXTENSION OF THE ORIGINAL MODELS .....	12
5.2.1 <i>Models Incorporating Human Capital Variables</i> .....	12
5.2.2 <i>Models Incorporating Human Capital and Firm Variables</i> .....	12
5.2.3 <i>Models Incorporating Human Capital, Firm and Industry/Occupation Variables</i> .....	13
5.3 BLINDER-OAXACA RESEARCH WITH NEW ZEALAND DATA .....	15
<b>6. METHOD .....</b>	<b>16</b>
6.1 DATA SOURCE .....	16
6.2 DATA CLEANING: MISSING VALUES .....	16
6.3 VARIABLES USED IN THE MODEL.....	17
6.3.1 <i>The Dependent Variable</i> .....	17
6.3.2 <i>The Predictor Variables</i> .....	17
6.4 DECOMPOSITION METHODOLOGY .....	18
<b>7. RESULTS .....</b>	<b>18</b>
7.1 REGRESSION DIAGNOSTICS .....	18
7.2 MODEL SUMMARIES .....	19
7.3 CHOW TESTS.....	20
7.4 DECOMPOSITION ANALYSES .....	21
7.4.1 <i>The “Explained” Gender Pay Gap – Earnings Power Function</i> .....	21
7.4.2 <i>The “Unexplained” Gender Pay Gap – Discrimination Function</i> .....	23
7.5 SUMMARY OF FINDINGS.....	25
7.5.1 <i>Occupation</i> .....	25

7.5.2 <i>Human Capital</i> .....	25
7.5.3 <i>Employer</i> .....	26
7.5.4 <i>Region</i> .....	26
7.5.5 <i>Employment Characteristics</i> .....	26
<b>8. DISCUSSION AND COMMENTARY ON METHODOLOGY</b> .....	<b>26</b>
8.1 DISCUSSION OF RESULTS .....	26
8.2 CRITIQUE OF METHODOLOGY .....	27
<b>9. REFERENCES</b> .....	<b>29</b>
<b>APPENDIX A: THE BLINDER-OAXACA DECOMPOSITION METHOD</b> .....	<b>32</b>
A1 THE DECOMPOSITION MODEL .....	32
A2 THE QUANTIFICATION OF DISCRIMINATION .....	33
A3 HUMAN CAPITAL IN THE MODEL .....	34
A4 GENERAL LIMITATIONS OF DECOMPOSITION METHODS.....	35
A5 PROBLEMS WITH BLINDER-OAXACA DECOMPOSITION.....	35
A5.1 <i>The Index Number Problem</i> .....	35
A5.2 <i>The Use of Indicator Variables</i> .....	35
A5.3 <i>The Model Evaluates at the Average</i> .....	36
A5.4 <i>The Influence of Proxy Variables</i> .....	37
<b>APPENDIX B: REMOVING “CONSTANT” OCCUPATIONS</b> .....	<b>38</b>
2000 DATASET .....	38
2001 DATASET .....	39
<b>APPENDIX C: CREATION OF AGE VARIABLES FOR THE MODEL</b> .....	<b>40</b>
C1 LINEAR EFFECT OF AGE .....	40
C2 ADDITION OF A QUADRATIC AGE VARIABLE.....	41
C3 ADDITION OF A CUBIC AGE VARIABLE.....	42
C4 RECOMMENDATION .....	42
<b>APPENDIX D: REGRESSION DIAGNOSTICS OUTPUT</b> .....	<b>43</b>
2000 .....	43
<i>Male Normal Probability Plot</i> .....	43
<i>Male Residuals Plot</i> .....	44
<i>Female Normal Probability Plot</i> .....	44
<i>Female Residuals Plot</i> .....	45
2001 .....	45
<i>Male Normal Probability Plot</i> .....	45
<i>Male Residuals Plot</i> .....	46
<i>Female Normal Probability Plot</i> .....	46
<i>Female Residuals Plot</i> .....	47
<b>APPENDIX E: REGRESSION DIAGNOSTICS WITH HOURLY SALARY</b> .....	<b>48</b>
E1 FINDINGS .....	48
E2 RECOMMENDATIONS .....	49

## **Executive Summary**

On average, male employees earn more than female employees. This effect is termed the “gender pay gap”, conventionally expressed as the percentage difference in earnings between male and female employees. Since the 1970s the gender pay gap has been decreasing. Over the past 30 years, a significant amount of international work has attempted to explain the gender pay gap. Many studies have used the “Blinder-Oaxaca decomposition” method, the focus of this report. Industry and occupation have consistently been found to have major effects on the gender pay gap. Both factors tend to exhibit gender-segregation, with male employees tending to work in higher paid industries and occupations, with occupation having the largest effect.

The current study examines the New Zealand Public Service, coded within the industry division of Government Administration and Defence, analysing data for 2000 and 2001 separately. The gender pay gap was 16.61% in 2000, falling to 15.51% in 2001. When differences in human capital factors and employment characteristics were taken into account, the gap decreased to 4.7% for both years. That is, the adjusted earnings level for female employees was 95.3% of the average male earnings for both 2000 and 2001, once human capital and employment differences were accounted for.

The average male advantage in hourly earnings was \$1.18 in 2000 and \$1.17 in 2001. These figures equate to an average male earnings advantage of just over \$2465 in annual base salary in 2000, and just over \$2438 in 2001. The adjusted male advantages are \$0.34 in hourly earnings (\$703.11 gross annual salary) in 2000 and \$0.35 hourly earnings (\$732.89 gross annual salary) in 2001.

Consistent with previous research, occupation had the largest effect on the gender pay gap. Proportionately more male employees worked in higher paid occupations with proportionately more female employees in lower paid occupations. Where male and female employees worked in the same occupations, male employees tended to earn more. This may be due to male employees tending to hold more senior positions within occupations. On average, male employees were older and had longer tenure than female employees, although this difference in human capital had mixed effects. Male employees also had a different distribution across employers compared to female employees, again with mixed effects. There were similar proportions of male and female employees in ethnic minorities (non-New Zealand European ethnicities), with male minority employees appearing to average slightly higher earnings compared to female minority employees.

There were similar overall proportions of male and female employees employed in regions outside Wellington, and females employed in occupations outside Wellington averaged higher earnings than males employed in occupations outside Wellington. Fixed-term employment agreements provided slightly higher average earnings for male employees. There was a male earnings advantage of 10.4% in 2000 and 9.4% in 2001 for being on an individual agreement.

## **1. Introduction**

### **1.1 Overview of the Report**

This study investigates the gender pay gap in the New Zealand Public Service. It is separated into ten major parts, plus appendices. Section 1 introduces the gender pay gap terminology, which is central to the discussions and analysis shown in this report. Section 2 summarises the New Zealand anti-discrimination legislation, which provides the legislative context for

interpreting findings of gender pay gaps. Section 3 summarises the gender pay gap findings from the UK, United States, and Australia. Section 4 summarises the recent gender pay gap findings from New Zealand. Section 5 summarises previous gender pay gap work using the Blinder-Oaxaca decomposition. Section 6 provides detailed notes on the methodology used in the analysis. Section 7 presents the results of the decomposition of the gender pay gap in the New Zealand Public Service, with Section 8 summarising the results and the methodology used. Section 9 contains the references.

There are five appendices. Appendix A discusses the analytical methodology (Blinder-Oaxaca decomposition) that is followed in this report for investigating the gender pay gap in the New Zealand Public Service. Appendix B shows the occupations that were removed from each dataset so that the same regression model was used for each decomposition. Appendix C details the method used to determine the measures of age that should be included in the two decompositions. Appendix D contains the normal probability plots and studentised residuals plots for each regression. Appendix E examines the regression diagnostics for the alternative method of using the untransformed hourly salary.

The coefficients and means for each variable for the male and female models in each year, and example decomposition calculations, are available from the author upon request.

## **1.2 The Gender Pay Gap**

Historically men have earned more – on average – than women even when hours of employment are controlled. This difference between male and female earnings is termed the “gender pay gap”, and is normally represented as either the ratio of average female earnings to average male earnings ( $\frac{Avg_{FE}}{Avg_{ME}}$ ) or the percentage point gap in earnings ( $100 - (\frac{Avg_{FE}}{Avg_{ME}} * 100)$ ). In this report,

unless specified otherwise, the gender pay gap refers to the percentage point gap in male and female earnings.

From a simple employment equality perspective, there is improvement in the gender pay gap when the ratio nears unity (or the percentage point gap decreases). A difference in male and female earnings is undesirable from an economic perspective. It is accepted that labour market discrimination has effects on all employees. Black (1995) has developed an equilibrium job search model that suggests employer discrimination on pay or hiring has effects on minority and non-minority workers. According to this model, the presence of such “discriminating” employers reduces the earnings of minority workers compared to non-minority workers even when the minority workers are employed in “non-discriminating” firms. The mere presence of “discriminating” employers has the effect of lowering labour market returns for the minority workers. This, in turn, is influenced by the lower returns to job search for minority workers compared to non-minority workers, resulting in a lower worker/employer match compared to that for non-minority workers. The model also suggests that as the proportion of minority workers rises, the earnings of non-minority workers will decrease.

From the gender pay gap perspective, females are the minority worker group. An imbalance in male and female earnings, as demonstrated by a gender pay gap, implies inequality in the labour market. The identification of the sources of inequality is important, as the policy responses (if any) used to address the imbalance will alter depending on the nature and degree of influence of the myriad of factors identified as influencing the gender pay gap. For example, differences in male and female earnings within the same occupation may suggest the use of anti-discrimination

policies to reduce the imbalance. However, differences in the wages between similar occupations may suggest that policies around comparable worth would be more effective.

## **2 The New Zealand Human Rights and EEO Context**

Equal Employment Opportunities (EEO) and protections based on the concept of “human rights” are interrelated yet distinct. In the case of EEO, the onus is on the employer to take a proactive approach in identifying and eliminating any organisational practice(s) that could lead to inequality in employment between individuals or groups. Section 58 of the State Sector Act 1988 places this onus on the individual departments within the New Zealand Public Service (see Section 2.1.2 below for more detail). Enforcement of EEO tends to be by way of monitoring, review, and guidance by internal EEO committees and by external agencies, such as the EEO Trust and the State Services Commission (for Public Service departments only). Compliance with EEO procedures is demonstrated with reference to the employer’s overall workforce, such as the percentage of women or Maori employed.

On the other hand, the Human Rights Act 1993 promotes an individual rights-based approach to freedom from discrimination in various areas (including employment) on 13 specified grounds. In general, this Act promotes similar or same treatment for different groups and individuals. Section 73 of the Act does, however, allow “measures to ensure equality” for persons or groups. Enforcement of the Human Rights Act 1993 is largely by way of individual complaints and legal processes.

While women are a target group for EEO- and human rights-based policies, so typically are other groups defined on the basis of ethnicity (e.g., Maori in New Zealand, Blacks and Hispanics in the US), disability, age (e.g., older persons), and sexual orientation (gay/lesbian)<sup>1</sup>. Regardless of country, EEO policies tend to be initiated by governments (federal and/or state) with the aim of reducing the discrimination experienced by disadvantaged groups, for example, the reduction of the higher unemployment rates for these groups. Thus, the attitude of countries towards EEO is illustrated by the timing and impact of legislation designed to further these aims.

This study discusses protections from a gender perspective only.

### **2.1 The Legislative Framework**

The first legislation to introduce anti-discriminatory principles was the Race Relations Act 1971, followed by the Human Rights Commission Act 1977. These two pieces of legislation affected both private and public sector organisations in New Zealand.

#### ***2.1.1 Legislation Covering Pay Discrimination***

The concept of equal pay for equal work had been introduced in the public sector in the 1960s (Hyman, 1992). The enactment of the Equal Pay Act 1972 legislatively spread this principle to the private sector. While the Act provided for equal remuneration for females when performing work that was substantially similar to males (i.e., work of equal value), as well as equal remuneration for the same jobs, only this more narrow interpretation of the Act was actually enforced in practice (see, e.g., Review Committee, 1979, cited in Hyman, 1992).

Union pressure led to the then Labour Government introducing the Employment Equity Act in 1990, which was subsequently repealed by the National Government in the same year, after the 1990 general election. Section 28 of the Employment Equity Act 1990 required all public sector employers, and private sector employers with at least 50 workers, to develop and implement an

EEO programme. The repeal of this Act also removed the sole statutory EEO requirement of private sector employers.

### **2.1.2 State Sector Act 1988**

In 1984 central government (i.e., Public Service) organisations pledged to lead EEO development through the Statement by Government Employing Authorities on Equal Employment Opportunities. This statement said:

*Indirect discrimination occurs when the outcome of rules, practices and decisions which treat people equally in fact reduce significantly the chances of a particular group of people from obtaining a benefit or an opportunity. This happens because people are not identical. Employing authorities in the government sector have a responsibility to ensure that groups such as women, ethnic minorities and disabled persons can as far as possible achieve equality with other members of the community (SSC, 1995, p. 48).*

Little actual EEO progress was made by government organisations, however, until the State Sector Act 1988, which enshrined the principles of EEO into legislation governing the Public Service. As well as the promotion of EEO, the Act also emphasised the Public Service as a good employer and the concept of merit – the employment of the person best suited to the job. Section 56(2) defines a good employer as “*an employer who operates a personnel policy containing the provisions generally accepted as necessary for the fair and proper treatment of employees in all aspects of their employment*”. The Act states a minimum of eight factors that should be included in these provisions:

- good and safe working conditions;
- an EEO programme;
- the impartial selection of suitably qualified people for appointment;
- recognition of the aims, aspirations and employment requirements of the Maori people, and the need for greater involvement of the Maori people in the Public Service;
- opportunities for the enhancement of abilities of individual employees;
- recognition of the aims and aspirations, and the cultural differences, of ethnic or minority groups;
- recognition of the employment requirements of women; and
- recognition of the employment requirements of people with disabilities.

Section 58(3) of the State Sector Act 1988 defined an EEO programme as one “*aimed at the identification and elimination of all aspects of policies, procedures, and other institutional barriers that cause or perpetuate, or tend to cause or perpetuate, inequality in respect to the employment of any persons or group of persons*”. The Act also requires Chief Executives of the Public Service departments to, on an annual basis: develop and publish an EEO programme for their department; ensure compliance throughout their department with the programme; and report on the extent to which their department was able to meet their programme. Thus, each department has the responsibility to ensure its compliance with the State Sector Act 1988. The State Services Commission, through the Act, has an oversight role in monitoring the compliance of departments with the Act, including its own compliance.

### **2.1.3 Human Rights Act 1993**

This Act came into force on 1 February 1994, replacing the Race Relations Act 1971 and the Human Rights Commission Act 1977. This Act contains 13 grounds for unlawful employment

discrimination<sup>2</sup>, and binds all employers, including Public Service departments. The grounds are:

- gender (including pregnancy and childbirth);
- marital status (including de facto or common law marriage);
- religious belief;
- ethical belief;
- colour;
- race;
- ethnic or national origin;
- *age* (from 16 years);
- *disability*;
- *political opinion*;
- *employment status*;
- *family status* (including presence or absence of children, one's relatives); and
- *sexual orientation*.

The factors in italics were introduced in 1993, so discrimination on any of these six grounds was legal prior to the introduction of the Act.

The Human Rights Act 1993 prohibits both direct and indirect discrimination. Direct employment discrimination occurs when a person is discriminated against on one of the 13 prohibited grounds. An example of direct discrimination is when a person is denied employment or a promotion because of his/her responsibilities for dependents. Indirect discrimination occurs when apparently non-discriminatory practices have the effect of different treatment of a person or group of people. An example of indirect employment discrimination is when an employer's building is not accessible for a person with disabilities.

There are three general grounds on which discrimination on prohibited grounds can be legal. These are:

- **Genuine occupational qualification:** The employee is required to have a particular characteristic related to job performance that is a prohibited ground. For example, a sexual health doctor may be appointed or refused employment on the basis of gender.
- **Reasonable accommodation:** The employer must make reasonable efforts to accommodate a person who could be prevented from performing a job because of one of the prohibited grounds. The test of reasonableness depends on the circumstances. For example, the employer may have to ensure that the place of employment is accessible for people in wheelchairs.
- **Unreasonable disruption:** The employer is not expected to accommodate a person's needs where this would cause unreasonable disruption to the employer's activities. For example, an employer selling plants would not be required to terminate pollen-producing stock to accommodate an employee or prospective employee with allergies.

These exemptions do not apply where the job applicant or employee is the best person for the job (merit principle) and cannot carry out some job duties on the basis of one of the prohibited grounds, but also where those duties could be carried out by another employee without causing unreasonable disruption to the employer.

### **2.1.4 Summary of the Current New Zealand Position**

Both public and private sector employers cannot discriminate on any of the 13 prohibited grounds in the Human Rights Act 1993. Discrimination in employment can occur in hiring, promotion or pay increase processes. As one of the prohibited grounds is gender, the legislative environment in New Zealand does not generally support a gender pay gap. Public Service departments are also subject to the State Sector Act 1988, which has a higher anti-discrimination expectation due to the requirement of these Chief Executives to follow EEO principles. These factors suggest that New Zealand should experience a low effect of discrimination on any gender pay gap, and that the gender pay gap in the Public Service should be lower than the equivalent gap in the private sector.

## **3. International Findings on Gender Differences in Pay**

The purpose of this section is to show the international context in which gender pay gap findings for New Zealand should be placed. The three countries that are discussed below are the United States, the United Kingdom, and Australia. These countries have been chosen as most of the gender pay gap statistical modelling, discussed later in this report, has been performed on data from these countries.

### **3.1 United States**

Blau and Kahn (2000) provide a good summary of the gender pay findings for the US from the mid-1970s. The gender pay gap for full-time employees decreased between the late 1970s and the early 1990s, with a plateau from the mid-1990s. An analysis of 10-year age cohorts found that the gender pay gap tended to decrease for most cohorts over time, indicating that each new cohort of women entering the workforce fares better economically than the previous cohorts. There are also indications that the gender pay gap increases when the cohort reaches the 35-44 years band, rising again in the next band. This suggests that the wages of women in full-time employment are penalised during the child bearing and rearing years. The segregation of women across occupations has also decreased, by similar levels between 1970 and 1980, and between 1980 and 1990. Most of the decrease was due to the movement of women into male-dominated occupations, especially into the male service and white-collar occupations rather than blue-collar jobs. This reduction of occupational segregation continued in the 1990s, but at a slower pace.

Other factors that helped reduce the US gender pay gap included women's increased human capital (education and work experience) and wage convergence, with women receiving a higher increase in indexed wages compared to men between 1978 and 1998. For both genders, wage inequality (as measured by the standard deviation of the natural log of wages) also increased between 1978 and 1998, at approximately the same rate.

Fortin and Lemieux (1998) examined the gender pay gap for the period 1979 to 1991. They found a convergence in the skew of the male and female wage distributions, suggesting that the gender pay gap did not decrease at a uniform rate across the wage distribution. An analysis of the gender pay gap by percentiles, for 1979 and 1991, showed that the gender pay gap decreased more for the 40<sup>th</sup> to 75<sup>th</sup> percentiles. The gender pay gap was smallest, for both periods, below the 25<sup>th</sup> percentile. The relative position of women improved due to increased female labour market experience and improvements in log wage position relative to men.

### **3.2 United Kingdom**

The trends were similar to the US. In the UK, the gender pay gap for women in full-time employment has shrunk from 35% of median hourly pay in 1970 to 20% in 1994 (Harkness, 1996). During this time period, the average age of women in full-time employment has increased

so that their age profile resembles that for working men. This change has been mainly due to the increased labour market participation of women of child-bearing age. Women also increased their educational attainment to such a level that there was no qualification gap between men and women, aged less than 35 years, in the 1990s. However, older women continue to make the largest contribution to part-time employment and, perhaps because of this, women in part-time employment now tend to be less qualified than both men and women in full-time employment. Between 1973 and 1993 real average hourly earnings increased for men and women, with both part-time and full-time employed women experiencing a higher increase than men. Wage inequality also increased over this period, although not for women employed part-time, and this effect was larger for men than for women.

A comparison of female earnings to male earnings deciles in 1973, 1983 and 1993 shows that females employed full-time improved their decile standings. In 1973, 88% earned less than the male median earnings and over 45% had earnings placing them in the bottom male decile. In 1993, these figures were 67% and 17% respectively. The percentage of women employees in the top male decile almost doubled from 1.3% in 1973 to 2.4% in 1993.

One major difference between the UK and countries like the US and New Zealand is the only recent (1999) introduction of an across-the-board minimum wage. From the 1950s to their official disestablishment in 1993, Wages Councils set the minimum pay rates, but only in low-paying industries (Dex, et al, 2000). These industries contained a high proportion of women and a low proportion of trade unionism. There is some evidence that the removal of the Wages Councils caused reductions in pay and conditions, especially for the very low paid (Craig, Rubery, Tarling and Wilkinson, 1982, cited in Dex, et al, 2000). Given these findings, and that women disproportionately receive lower wages than men (also due to their higher part-time work status, as well as to their occupational and industrial distributions), Dex et al (2000) predict that the minimum wage will cause the gender pay gap to decrease. The largest gain is expected for female part-time workers in manual occupations, with an expected decrease in the gender pay gap of 2 percentage points. There is little change for other major occupational groups because the minimum wage affects all low paid workers and the minimum rate of £3.60 is very low.

### **3.3 Australia**

In the ten years to June 2000, the Australian male unemployment rate was static at 6.4% and the female unemployment rate decreased from 6.9% to 6.2% (Preston, 2000a). Both the male and female labour market participation rates also changed; the female rate increased from 52% to 55% and the male rate decreased from 75.7% to 72.6%. In 2000, while females represented 44% of the Australian labour force, 57% of females were in full-time employment<sup>3</sup> compared to 88% of males and 73% of part-time jobs held by women. Over this decade, females were employed in 60% of new jobs, representing 61% of new part-time jobs and 58% of new full-time jobs. Ninety-percent of male employment growth was in casual jobs, with 54% of this growth in the part-time category. While the corresponding figure for women was much lower at 37%, 81% was in the part-time category. There is also evidence of increasing wage inequality from 1991 to 1998, which was more marked for men. Preston attributes this to a decrease in (male) trade union membership.

There was no change in the gender pay gap at the national level using average weekly ordinary time earnings (based on a four-quarter moving average), which remained essentially static at 15.6% between 1991 and 1999 (Preston, 2000a). Western Australia was consistently associated with the largest gender pay gap during this time, with a gap of 21.4% in 1999. While South Australia had the smallest gender pay gap for the first four years of the period, it decreased to plateau at the national average from 1996. These findings should be interpreted in light of the

labour market deregulation that occurred in Australia, in 1992 with amendments to the Industrial Relations Act 1988 and in 1994 when the Industrial Relations Reform Act 1993 took effect.

Preston (2000b) paid a closer look at the gender pay gap in Western Australia (WA). In May 2000, the gap in the full-time labour market was 16% nationally and the WA gap was about 22%. While the WA gender pay gap has been larger than the national gap over the period of analysis (February 1990 to February 2000), there was a significant increase in the WA gap in 1993 and 1994. As noted above, the national gender pay gap remained static. In 1996, after the WA gender pay gap had experienced its most dramatic increase, men in WA earned 3.8% more than their male counterparts in other states and females earned 6.8% less compared to their female counterparts, when human capital factors such as marital status, age of children, and highest qualification were taken into account. While WA females earned 13.3% less than WA males in 1990, the 1996 gap was 18.5%. This occurred in spite of the higher improvement of human capital for females compared to males over the same time period. These results are interesting in the context that WA had the most vigorous promotion of individualistic industrial relations bargaining, and most of the state legislative reforms were enacted in 1993 (Industrial Relations Amendment Act 1993, Minimum Conditions of Employment Act 1993, and Workplace Agreements Act 1993, all cited in Preston, 2000c).

## **4. New Zealand Findings on Gender Differences in Pay**

### **4.1 Employment Legislation Context**

New Zealand's employment history could also be considered a history of unionism. The formation of craft-based unions predated the 1878 Trade Union Act<sup>4</sup> by over a decade, and unions for some semi-skilled and unskilled workers were started in the 1870s (Deeks, Parker and Ryan, 1994). The working class ensured that the Liberal Party was elected to power in 1890, leading to the Industrial Conciliation and Arbitration Act 1894. Both employers and unions, comprising a minimum of seven workers in an industry, could register under this Act, with the registered union becoming the legal representative of the workers in that industry. The Act introduced compulsory arbitration of industrial disputes, and became the main process for wage fixing.

Because of the Great Depression of the late 1920s and early 1930s, the compulsory arbitration provisions of the Industrial Conciliation and Arbitration Act 1894 were repealed, resulting in large decreases in wage rates (Deeks, Parker and Ryan, 1994). With the election of the first Labour Government in 1935, this Act was amended so that union membership effectively became compulsory in 1936. Minimum wages were introduced with effect from 1 April 1946. Initially covering male and female workers aged 21 years and over, excluding some general classes such as apprentices, this age was reduced to 20 years in 1970. The Industrial Conciliation and Arbitration Amendment Act 1936 also introduced the 40-hour working week (Szakats, 1988).

The next biggest change to employment law came when the fourth Labour Government introduced the Labour Relations Act in 1987 (Deeks, Parker and Ryan, 1994). This Act removed secondary bargaining<sup>5</sup>, encouraged collective agreements, and required registered unions to have a minimum of 1000 members. The Act, however, retained the 40-hour maximum working week for awards. Four years later, the National Government replaced the Labour Relations Act with the Employment Contracts Act. The Employment Contracts Act 1991 introduced three major changes: (a) effectively, the introduction of voluntary unionism; (b) the introduction of individual employment contracts; and (c) the extension of personal grievance procedures to non-union members.

These legislative changes have had an impact on the number of unions and union density (the percentage of wage and salary earners who are union members). Between December 1985 and September 1989, which saw the introduction of the Labour Relations Act 1987, while the number of unions fell from 259 to 112 (57% decrease) the union density actually rose from 66% to a peak of 73% – an increase of almost 11% (Harbridge, 1993, cited in Deeks, Parker and Ryan, 1994). The introduction of the Employment Contracts Act 1991 was associated with a significant decline in both union numbers and union density, so that there were only 58 unions and union density was only 46% in December 1992. In 2000, the Employment Contracts Act was replaced by the Employment Relations Act.

The other recent major legislative impacts on wages were the changes to personal income tax rates in the 1980s. There were decreases to the top tax rates in 1986 and 1988, and the introduction of a goods and services tax of 10% in 1986, raised to 12.5% in 1989 (Dixon, 1996a).

## **4.2 General Labour Market Findings**

During the 1973 to 1977 implementation of the Pay Equity Act 1973, the gender pay gap decreased from 27.9% to 21.5% (Hyman, 1992). This reduction trend slowed after 1977, leading eventually to the introduction of the Employment Equity Act 1990, subsequently repealed.

From 1984 to 1994, the female participation rate in employment increased from 43% to 49%, the percentage of earners with no formal qualifications decreased from 40% to 23%, and the percentage with post-graduate qualifications increased from 35% to 43% (Dixon, 1996a). These findings suggest that earnings should increase over the period due to the increase in human capital by wage and salary earners. Between 1993 and 1996 the seasonally adjusted male participation rate increased from around 73% to just below 75%, compared to a 1986 participation rate of about 79% (Dixon, 1996b). The seasonally adjusted female participation rate increased from around 53% in 1993 to about 57% in 1996. Between 1986 and 1996, there was a decrease in the proportion of prime-aged (i.e., 25-54 years) males, in the proportion of Maori and Pacific peoples males, and in the proportion of adults without formal qualifications in employment. There was an increase in the proportion of prime-aged females, and earners aged 60 to 64 years. Most of the growth in female employment, however, was in part-time employment, and in 1999 over 13% of part-time male and female workers wanted full-time employment (Morrison, 2001).

Real wage and salary earnings actually decreased between 1984 and 1994, with a 2.7% (4.5%) decrease in mean (median) weekly earnings and a 6.5% (8.0%) drop in mean (median) hourly earnings (Dixon, 1996a). The weekly decrease was smaller because weekly hours increased over the period. Paid overtime hours also declined, more substantially for men, although this was not a noticeable cause of the decline in real earnings. There was increased weekly earnings dispersion for males over the period, although this did not translate to an increase in hourly earnings dispersion as there was an increased proportion of males in part-time employment and an increased proportion of males and females reporting at least 45 hours of work per week.

The gender wage gap decreased between 1984 and 1994 because the average real hourly earnings for males decreased more than the average real hourly earnings for females (Dixon, 1996a). Wage and salary earners aged 15 to 24 years experienced the largest decrease in average real hourly earnings, with earners aged 55 years and over experiencing the smallest decrease. Overall, the decrease affected part-time and full-time workers equivalently, although the median hourly earnings of full-time female earners increased by 1.0% compared to a decrease of 6.5% for part-time females. In 1984, the gender pay gap for median earnings was 21%, reducing to

11% in 1992, but increasing to 14% in 1994. Inequality in earnings also rose between 1995 and 1997 (Dixon, 1998)

### **4.3 Income Inequality**

The work by Dixon (1996a, 1996b), summarised in the previous section, indicates that income inequality has increased in New Zealand from 1984 (the earliest year considered in Dixon's work). A number of working papers commissioned by the New Zealand Treasury examine this income inequality.

The late 1980s saw the most substantial growth in income inequality, although it increased at a slower rate in the 1990s (O'Dea, 2000). This rate of growth in inequality was not experienced by other OECD countries. As wages and salaries represent around 80% of the sources of personal income, these two sources are the predominant causes of income inequality. At the household level, between 1983/86 and 1995/98 there was a large decrease in the proportion of middle-income (\$30,000-\$100,000) households, with a corresponding increase in the proportions of both low- and top-income households. Most of the increase occurred at the low-income end of the distribution. Because households are comprised of differing numbers of individuals, however, changes in household memberships can have a large influence on income at the aggregated household level of measurement. Changes in household proportions explained 17% of the increased inequality of household incomes between 1985/86 and 1995/96.

Acemoglu (2001) suggests some key factors that have caused the increase in wage inequality. He believes the major determinant has been technical change that is skill-biased, i.e., change that provides more benefits to skilled or more highly educated workers and less benefits to unskilled or less educated workers. As this type of technological change will continue, wage inequality produced from this source will continue to increase. Three additional determinants of wages are discussed: increased education of workers; changes in patterns of rent sharing (e.g., level of unionisation, industry effects on average wages); and the quality distribution of jobs ("good jobs" versus "bad jobs"). US evidence suggests that these three factors are secondary to technological change in explaining changes to wage inequality. The trends in these factors suggest that wage inequality in New Zealand will continue to increase.

### **4.4 Future Predictions**

In 1997, the Ministry of Women's Affairs, in conjunction with the New Zealand Institute of Economic Research (NZIER), published two related forecast reports on the gender pay gap. The analyses were performed on the basis of industry rather than occupation, and produced forecasts on the gender pay gap in the New Zealand labour market to 2001. The first report suggested that the gender pay gap would slightly increase at the aggregated industry level (Cook and Briggs, 1997). The second report suggested a slightly decreased gender pay gap in 2001, with a narrowing gap in the manufacturing and the transport/communication industries, and a widening gap in the business/financial services and other community social services industries (Barnett, 1997). The forecasts for the Public Administration/Defence industry predicted that the gender pay gap would actually increase in the time period to 2001. The suggested primary cause of this trend is a higher forecast ratio of female-to-male hours worked compared to other industries, coupled with a low forecast ratio of female-to-male wages (predicted to be the second lowest out of nine industry groups).

### **4.5 The Gender Pay Gap in the New Zealand Public Service**

The limitation of these studies is that they have tended to assess the New Zealand labour market as a whole. The New Zealand Public Service has specific EEO requirements that are not

reflected in legislation governing the private sector since the repeal of the Employment Equity Act in 1990. Therefore, an examination of the gender pay gap in the New Zealand Public Service is of particular interest.

The State Services Commission has examined trends in the employment participation of women in the New Zealand Public Service. The proportion of women in the Public Service has been gradually increasing over the 1990s. In June 2001, women represented 56.5% of the New Zealand Public Service, compared to 45.7% of the June 2001 employed labour force (State Services Commission, 2001). By comparison, in 1996 54.7% of the Public Service and 44.9% of the employed labour force were women (State Services Commission, 2000).

The SSC's Human Resource Capability surveys have also examined the gender pay gap in the New Zealand Public Service, using gross annual base salary. The June 2001 pay gap was 17% for the Public Service, compared to 16% for the employed labour force (State Services Commission, 2001). The June 2000 pay gaps were larger at 19% and 17%, for the Public Service and the employed labour force respectively (State Services Commission, 2000). However when the pay gaps within aggregated occupation categories<sup>6</sup> were examined, the largest gender pay gap was 17.2% for "Managers" in 2000 and 16% for "Managers" in 2001. The smallest gender pay gaps were for 4.5% for "Personal and Protective Services Workers" in 2000 and only 1% for "Trades and Production Workers" in 2001.

This project uses the same datasets as those for the 2000 and 2001 Human Resource Capability surveys to perform Blinder-Oaxaca decompositions for the gender pay gap in the New Zealand Public Service (refer to Appendix A for detail on the Blinder-Oaxaca decomposition method, the assumptions of the method, and some problems with this methodology).

## **5 Findings From Blinder-Oaxaca Decompositions**

This section discusses the development of the variable specifications of the Blinder-Oaxaca model, starting with the original models.

### **5.1 The Original Decompositions**

Oaxaca (1973) used a 12-variable model to investigate the gender wage gap in the natural-log of hourly wages ( $\ln-W$ ). The model for females only contained the number of children, and both male and female models contained another eleven sets of variables, of which nine sets were comprised of indicator variables. These variable categories were: experience<sup>7</sup> (linear and quadratic terms); education (linear and quadratic); worker class; industry (2-digit); occupation (2-digit); presence of health conditions (indicator variable); part-time status (indicator variable); migration; marital status; urban area type; and region. The wage differentials were calculated separately for whites and blacks, with a log differential of 43% for whites (non-log differential of 54%) and a log differential of 40% (49%) for blacks. Industry, in particular, and also occupation and class of worker (union, government, or self-employed), had the largest effects on the gender pay gap.

Blinder (1973) reported the results from two decompositions: a white/black wage differential and a male/female wage differential for whites. The second decomposition will be examined here. The structural wage regression used 12 variables, in which only the two age variables were continuous: age (linear and quadratic); region; local labour market; migration; health conditions; education (indicator variables); occupation; union membership (indicator variable); veteran status (indicator variable); seasonal employment (indicator variable); vocational training (indicator variable); and tenure (indicator variables). The reduced form regression used ten variables and,

again, only the age variables were continuous: age (linear and quadratic); region; local labour market; migration; health conditions; seasonal employment (indicator variable); siblings (two indicator variables); father's education; parent's economic status; and childhood residence.

The structural differential found large influences of age, education, and local labour market conditions (Blinder, 1973). Age explained most of the difference in the gender pay gap, as women's wages did not tend to rise over the life span whereas the wages for men did tend to rise. The other two largest contributors were education and local labour market conditions. While men and women had the same average endowments of these factors, men received greater returns for education and were less affected by the local labour market conditions. In the reduced form differential, age accounts for the whole gender pay gap. Again, men were less affected by the local labour market conditions compared to women.

## **5.2 Extension of the Original Models**

Over the past decades, the Blinder-Oaxaca method has been used as the basis for many comparisons of male and female wages. The major types of decomposition used over the past decade have incorporated human capital variables, human capital plus firm (employer) variables, and human capital plus industry/occupation variables. The remainder of this section summarises the findings from these three types of study.

### ***5.2.1 Models Incorporating Human Capital Variables***

While the original Blinder and Oaxaca models included firm-based and industry/occupation variables, one study used only human capital variables. Waldfogel (1998) found that even when the gender pay gap had been closing, the pay gap between women with children and women without children has been widening in the US. This phenomenon is called the "family gap". When only women with children are considered, sole parent mothers fared worst, including women who had been previously married. The two decompositions performed by Waldfogel (on 1980 and 1991 data) indicate that the presence of children has such a negative effect on women's earnings that if there was no difference between women with children and women without children, especially for human capital, then the family gap in pay would still exist. Waldfogel suggests that this result is mediated by the provision of maternal leave, as countries with better maternal leave policies (e.g., longer leave provisions, paid leave provisions) have smaller gender pay gaps, possibly due to the positive effect of maternal leave provisions on tenure and work experience.

### ***5.2.2 Models Incorporating Human Capital and Firm Variables***

The addition of firm-based variables in the model, such as number of employees, provides more information on the source of differences in male and female wages.

Chauvin and Ash (1994) used a sample of American business school graduates to decompose the gender pay gap in base pay, contingent pay<sup>8</sup>, and total pay. (Three indicator variables relating to occupation in the professional, technical, and sales categories were included in the model.) The authors found a 9% unexplained pay gap for total pay, and no unexplained pay gap for base pay, adjusted for differences in means. Further decomposition work showed that the unexplained pay gap in total pay was due to gender differences in contingent pay; this gap disappeared when contingent pay was added to the model. This suggests that the size of the gender pay gap is strongly dependent on the type of pay data used as the dependent variable.

Swaffield (2000) found that full-time education of up to five years duration had a positive effect on women's hourly wage, but no effect on that for men. Unemployment had a negative impact

on the hourly wage of both males and females, but duration of unemployment of over one year had no additional effect. While employment in a male-dominated occupation<sup>9</sup> increased the female hourly wage, the women employed in such occupations were penalised more heavily for exiting the paid labour force. The interaction between being employed in a male-dominated occupation and time spent outside the labour force had a significantly negative effect on the female hourly wage.

One primary advantage of using employer-employee linked data is the ability to include a measure of the proportion of women in the employer's workforce within the model. The consistent finding from these studies is that a predominantly female workforce – at the employer level – significantly decreases both male and female wages. The following two studies used linked employer-employee data.

Reilly and Wirjano (1998) used a Mincerian analysis of the 1979 Canadian data from the General Segmentation Study. Individuals were included if they worked at least 30 hours per week. The authors found that the largest single (negative) effect on wages was the proportion of women in the employer's workforce. Education and experience both increased wages, as did tenure to a lesser degree. In a later Blinder-Oaxaca decomposition, the proportion of females in the employer workforce was found to account for 26% of the mean gap in log wages. The practical significance of these results is increased by the fact that males tended to be employed in male-dominated workplaces and females in female-dominated workplaces. The implication from the model is that increased workplace segregation will increase the gender pay gap.

Reiman (2001) used data from the 1995 Australian Workplace Industrial Relations Survey (AWIRS95) initially to construct a regression of log wages against a set of 36 human capital and employer-based variables (incorporating 28 indicator variables). He found that males earned 7.56% higher than females, indicating that a gender pay gap existed in Australia in 1995. Working full-time ( $\geq 35$  hours per week), and working for an employer with a predominantly female workforce, were associated with lower wages. Higher years of schooling and having English as the primary language were associated with higher wages. With further analysis, Reiman found six variables that produced statistically significant differences between males and females. Compared to males, females were paid more in full-time work, when they had children aged less than five years, and in metropolitan areas. Females were paid less in South Australia and when the employer was at least partially foreign-owned. Reiman's model reduced the gender wage gap from 13.4% to 8.1%, which is a reduction of 39%. In dollar terms, the adjusted gender pay gap equated to \$1.23 Australian dollars per hour, before tax.

### ***5.2.3 Models Incorporating Human Capital, Firm and Industry/Occupation Variables***

Groshen (1991) used the data from five American Industry Occupational Wage Surveys in a Blinder-Oaxaca decomposition. Because different industry-based surveys were used, the data periods range from 1974 to 1983. While the industries were decomposed separately, the five findings actually relate to three different years. Occupation was found to be highly segregated, and wages were found to be strongly related to the proportion of females in the occupation. Occupation was found to account for over half the observed gender wage gap. While males and females who worked in the same occupation for the same employer (termed a "job cell") earned roughly the same amount, most occupations were segregated and within employers most occupations were totally segregated.

Bayard, Hellerstein, Neumark, and Troske (1999) matched employee records from the 1990 Worker-Establishment Characteristics Database constructed from the American Decennial

Census to employers listed in the US Census Bureau's American Standard Statistical Establishment List. The authors could not replicate Groshen's finding of sex segregation causing the gender pay gap. In this later study, while females were found to be segregated into lower paying industries and occupations, the largest influences on the gender pay gap arose from the lower wages of females compared to males for the same job cell and in job cell segregation. The level of occupation disaggregation in the model was important; as occupation was more highly specified (from a 1-digit level, representing 13 occupations through to a mix of 3- and 4-digit occupations, representing 491 occupations) the size of the gender coefficient decreased.

Macpherson and Hirsch (1995) examined the influence of feminised occupations on the gender pay gap, primarily using 1983 to 1993 data from the American Current Population Survey Outgoing Rotation Group. Occupation was analysed at the 3-digit level, and the proportion of female workers in each occupation was calculated. Macpherson and Hirsch found significantly lower wage rates for all workers in feminised occupations (containing at least 75% women) and also in masculinised occupations (containing at least 75% men). The lowest average wage is associated with the group of feminised occupations. For all occupation groups (0-25% women, 25-50% women, 50-75% women, 75-100% women) the average female wage was lower than the average male wage. When extra variables associated with job characteristics were introduced into the model, the influence of occupation feminisation on wages was reduced. The proportion of females in a job effectively acts as a proxy variable for differences in job characteristics (e.g., physical requirements), worker-based productivity differences, and preferences for job characteristics. For example, feminised jobs typically have a lower requirement for training (and, therefore, lower levels of human capital).

Finally, the following three studies also incorporate public sector variables into their models. Naur and Smith (1996) used three 10-year cohorts in their decomposition of Danish employees. They found that the youngest cohort (aged 20 to 29 years) had the smallest gender pay gap in 1980, but this gap widened in the ten years to 1990. Only the oldest cohort (aged 40 to 49 years) experienced a decreased gender pay gap. The middle cohort consistently had the largest gender pay gap in the 1980s due to the low wages paid to women, especially in the Danish public sector that was the primary employer of the middle and youngest cohort women. In comparison, the lower pay for women in the oldest cohort was mainly due to a lack of human capital. These findings suggest that occupational and industrial factors can override increased human capital (e.g., higher educational attainment) investment by women.

Further work by Gupta, Oaxaca and Smith (1998), using data from the Danish Longitudinal Sample, found little change in the gender pay gap in either the private or public sector between 1983 and 1994, with a decrease in public sector wages compared to the private sector over the time period. There was also an increase in wage dispersion, particularly in the public sector. Given that Danish women were concentrated on lower paying jobs, this increase in wage dispersion should have increased the gender pay gap. In both the private and the public sectors there was an increased return on qualifications, higher for females compared to males, so this human capital factor reduced the gender pay gap for both sectors. Qualification was decomposed into education, experience, and occupational position factors. The biggest contributor to the qualification effect was the increased labour experience of women in the public sector, and the increased educational attainment of women also decreased the gender pay gap in this sector. Changes in occupational position should have slightly decreased the gender pay gap, although this effect was countered by relative pay changes within the occupational groups.

The finding that increased educational attainment of women decreases the gender pay gap is robust across different cultures. For example, Sung, Zhang and Chan (2000) decomposed Hong

Kong census data from 1981, 1986, 1991 and 1996, incorporating seven occupational groups at the 1-digit level. The authors also used the Brown, Moon and Zoloth extension to the Blinder-Oaxaca method to separate intra- and inter-occupational wage differences. From 1981 to 1996 there was a decrease in the gender pay gap from 29% to 16.1%, partly due to increased female educational attainment and also due to the Hong Kong economy shifting from manufacturing to services-based, resulting in females shifting to the more highly paid services occupations. Regarding occupation, the gender pay gap is mainly intra-occupational rather than inter-occupational, with inter-occupational effects actually reducing the gender pay gap.

### **5.3 Blinder-Oaxaca Research With New Zealand Data**

Little published research has been performed in New Zealand using the Blinder-Oaxaca decomposition to investigate the gender pay gap. Primary New Zealand research using this method is outlined below in chronological order.

Dixon (1996a) performed a number of decompositions of salaried and waged employees using the Household Economic Survey (HES). The model included human capital, industry, and occupational variables. Initial decomposition models containing an indicator variable for gender was entered into the model, although latter analyses decomposed males and females separately. Qualifications and gender had the largest effect on earnings, with university qualifications and being males producing higher wages. Part-time work status was associated with lower hourly income.

Dixon (1998) examined the changes in income inequality between 1984 and 1997 using the Household Economic Survey (HES) conducted by Statistics New Zealand. She found that the gender pay gap in average hourly earnings decreased, as did the gender pay gaps in full-time weekly earnings and median hourly earnings. The reason for this reduction was that female earnings increased more than male earnings from 1984 to 1995, with no reduction evident in 1996 or 1997.

Dixon (2000) used the Blinder-Oaxaca decomposition method to examine the gender pay gap of New Zealand salaried and waged employees aged 20 to 59 years. Due to the low hourly pay gap between part-time and full-time female earners, part-time employees were also included in the analysis. Qualification (4 indicator variables), ethnicity (3 indicator variables), country of birth (2 indicator variables), part-time status, region (based on Regional Council), industry (2-digit or 3-digit level, included as indicator variables), and occupation (2-digit or 3-digit level, included as indicator variables) were included in the models. Two models decomposed log wages using 2-digit industry and occupation classifications. The HES data showed a log wage gap of 0.136, with between 14% and 30% of the gap attributable to industry and 4% and 10% attributable to occupation<sup>10</sup>. The Household Labour Force Survey Income Supplement (IS) data gave a log wage gap of 0.171, with between 20% and 24% of the gap attributable to industry and -9% and 5% attributable to occupation. When the decomposition of the IS data was based on 3-digit industry and occupation classifications, the log wage gap remained at 0.171, with a lower effect of industry (between 12% and 18%) and a larger effect of occupation (between 6% and 23%).

In other modelling work on the gender pay gap, Kirkwood (1998) used a tree analysis to examine 1997 earnings data on males and females in full-time employment (defined as at least 30 hours per week, not including the self-employed). A pruned tree with 12 terminal nodes explained 29% of average earnings. The most important variable was occupation (based on the one-digit New Zealand Standard Classification of Occupation – NZSCO – group) followed by hours worked. Age, highest education qualification, and sex were also found to be important in the model, although industry was not. In a subsequent standardisation procedure using these four

variables plus ethnicity, hours worked was found to have the greatest influence on average weekly earnings. The standardisation model also decreased the gender pay gap from 21% to 14% due to a better model specification resulting from the inclusion of additional variables. These findings suggest that gender is a secondary explanatory factor for earnings.

## 6. Method

### 6.1 Data Source

The data used in this study come from the 30 June 2000 and 30 June 2001 Human Resource Capability (HRC) data collections for the State Services Commission. The data consist of unit record observations for each person employed during the previous year in the New Zealand Public Service<sup>11</sup>, including entrants and exits. While part-time employees are included in the dataset (and can be identified from an hours variable), casual employees<sup>12</sup>, contractors<sup>13</sup>, and Chief Executives are excluded. The variables included in the analysis are those variables contained within the HRC datasets.

Each year has been regressed separately. The 2001 year has been analysed as a check on the relative reliability of the 2000 coefficients. Each observation is a unique Public Service employee.

### 6.2 Data Cleaning: Missing Values

There were 22,822 observations in the 2000 dataset and 24,060 observations in the 2001 dataset. For both datasets the same cleaning techniques were applied. As the decomposition methodology is multivariate, an observation requires information on all variables in the model in order to be included in the analysis. Instead of relying on the statistical process to remove the observations with missing values, or imputing the missing values, observations with missing values were removed prior to the analyses. Table 1 below shows the effect of the data cleaning process on observation numbers.

**Table 1 Data Cleaning Steps and Numbers of Observations Affected**

<b>Cleaning Step</b>	<b>2000 Dataset</b>	<b>2001 Dataset</b>
Initial number of observations	37,386	38,930
- Non-current Employees <sup>14</sup>	-7,462	-7,549
- Missing date of birth	- 362	- 453
- Missing occupation	- 230	- 163
- Missing ethnicity	-4,037	-3,996
- Part-time work status <sup>15</sup>	-2,181	-2,359
- Age < 16 years or > 64 years	- 59	- 103
- Single gender occupations <sup>16</sup>	- 233	- 247
Observations included in analysis	<b>22,822</b>	<b>24,060</b>

## 6.3 Variables Used in the Model

### 6.3.1 The Dependent Variable

The natural log of hourly wages ( $\ln\_wages$ ) was the dependent variable in the decompositions. The HRC dataset contains annualised salary. The conversion on salary was performed as:

$$\ln\_wages = \ln\left(\frac{AnnualSalary}{2088}\right)$$

The salary information comes from the employing department, and is the employee's annual base salary on the payroll system. This means that "above base" earnings that can be expected to change from one pay period to another, such as overtime and "at risk" pay, are omitted from the annual salary figures for all employees. The denominator (2088) is based on a 40-hour working week multiplied by 52.2 weeks (the average number of weeks in a year). The same denominator has been applied to all observations because the decompositions examine full-time employees only.

### 6.3.2 The Predictor Variables

The dataset for the 2000 and 2001 collections contained 191 and 194 predictor variables respectively. The difference in predictor variables was due to a slightly different specification of 131 occupational indicator variables and the addition of the Serious Fraud Office as an indicator variable in the employer category. The same reference categories were used for the 2001 decomposition.

Eight categories of predictor variable were included in the decompositions:

1. **Age** was included as two continuous variables, encompassing a linear form and a quadratic ( $age^2$ ) form. Refer to Appendix C for a detailed analysis of the effect of age on  $\ln\_wages$ , using the 2000 dataset.
2. **Ethnicity** was included as five indicator variables, where ethnicity was defined using a priority classification system<sup>17</sup>. The indicator variables were: NZ Maori, Pacific peoples, Asian, Non-New Zealand European, and Other. The reference category is *New Zealand European/Pakeha*.
3. **Occupation** was included as 129 (2000 decomposition) or 131 indicator variables (2001 decomposition)<sup>18</sup>. While there were a higher number of occupations in the analysis dataset at the 5-digit NZSCO level, the decompositions omit the indicator variables containing only males or only females (refer to Appendix B for more information). The reference category is "*Careers, Transition and Employment Adviser*" (NZSCO code = 33511), an occupation not limited to the New Zealand Public Service, and one which represents 1.5% of the dataset and has a female proportion of 52.2% (2000 dataset analysis).
4. **Tenure** was based on tenure with current employer, and was included as a continuous variable. Tenure was moderately correlated with age (0.45) and  $age^2$  (0.44),  $p < .001$  for both correlations (2000 dataset analysis).
5. **Region** was included as 16 indicator variables: Northland, Auckland, Waikato, Bay of Plenty, Gisborne, Hawkes Bay, Taranaki, Manawatu-Wanganui, West Coast, Canterbury, Otago, Southland, Tasman, Nelson, Marlborough, and Overseas. The reference category is *Wellington*.

6. **Employer** was included as 36 indicator variables. The reference category is the *Ministry of Pacific Island Affairs*, which has a female proportion of 48%, has an average salary close to the Public Service mean, and has a limited number of different occupation groups (2000 dataset analysis). To protect the identity of departments, the department names have been anonymised into the three categories (Small, Medium and Large based on employee numbers). For example, the variable “Small 1” indicates the first small department in the model.
7. **Term of employment** was included as one indicator variable for fixed-term employment<sup>19</sup>. The reference category is *open-term employment*.
8. Type of **employment agreement** was included as one indicator variable for individual agreement.<sup>20</sup> The reference category is *collective agreement*, and includes employees on expired collective agreements.

#### 6.4 Decomposition Methodology

This research uses the standard Blinder-Oaxaca decomposition shown in equation 5 in Appendix A. An alternative pooled method, suggested by Oaxaca and Ransom (1994), uses the cross-product matrices as the weighting matrix in the regression. This weighting method is assumed to create the wage structure that would have occurred if wage discrimination had not existed. As the current research is concerned with the gender pay gap given historical wage discrimination, the standard decomposition weighting method is used.

While both datasets come from census information on Public Service employees, the data cleaning process resulted in the removal of observations with missing values. Given this, and the fact that the data arise from an observational study rather than an experiment (i.e., values of the predictor variables are not fixed experimentally), the decompositions are based on random effects models rather than fixed effects models. This distinction simply means that the decomposition findings are estimates of the relationship between the predictor variables and  $\ln\_wages$ .

SPSS version 10 was used for the data analysis, and the full model was specified in each regression.

## 7. Results

### 7.1 Regression Diagnostics

The 2000 and 2001 decompositions were checked for violation of the five major regression assumptions: (i) absence of influential cases; (ii) linearity of  $\ln\_wages$  and the predictor variables; (iii) normal distribution of  $\ln\_wages$ ; (iv) constant variance of  $\ln\_wages$ ; and (v) independence of observations. A normal probability plot of the residuals was constructed for each male and female regression model (refer Appendix D). Visual inspection of these plots suggests that  $\ln\_wages$  meets the assumption of coming from a normal distribution. A scatter plot of studentised residuals versus standardised predicted values was constructed for each male and female regression model (refer Appendix D). Visual inspection of these plots suggests that the assumption of a linear relationship between  $\ln\_wages$  and the matrix of predictor variables, and the assumption that  $\ln\_wages$  has a constant variance, were met.

One-sample tests of fit can be used to determine whether a particular sample of observations fits a specified distribution. The one-sample Kolmogorov-Smirnov test for goodness-of-fit was used to determine whether the distribution of the standardised residuals from each model fit the normal distribution<sup>21</sup>. The alternative hypothesis ( $H_A$ ) for this test was that the cumulative density function of the standardised residuals from each model did not equal the normal

distribution for at least one observation. In this instance, the statistic was calculated as the largest difference in absolute value between the observed and normal distribution functions. The analyses showed that none of the models produced a normal distribution of standardised residuals. For 2000, the test statistic was 5.577 ( $p < .001$ ) for the male model and 6.962 ( $p < .001$ ) for the female model. For 2001, the test statistic was 5.804 ( $p < .001$ ) for the male model and 7.305 ( $p < .001$ ) for the female model.

Cook's D was used to check the assumption of an absence of influential cases for each male and female regression model. The largest Cook's D statistic in each model was .024 for the 2000/male model, .060 for the 2000/female model, .019 for the 2001/male model, and .035 for the 2001/female model. These are extremely small Cook's D values and suggest the absence of individual influential points on the analyses. In fact, with such large numbers of observations it would be unusual to find individual observations of particularly high influence.

Finally the Durbin-Watson statistic, shown in Table 2, was used to check the assumption of independence of observations for each male and female regression model. The formula of the Durbin-Watson test statistic  $d$  is given by (Ott, 1993):

$$d = \frac{\sum_{t=1}^{n-1} (\hat{\epsilon}_{t+1} - \hat{\epsilon}_t)^2}{\sum_t \hat{\epsilon}_t^2}$$

where  $n$  is the total number of time points,  $\epsilon_t$  is the residual for observation  $t$ , and  $\epsilon_{t+1}$  is the successive residual (observation  $t+1$ ). In the absence of serial correlation, the expected value of  $d$  is 2.0, with values of  $< 1.5$  suggesting positive serial correlation and values  $> 2.5$  suggesting negative serial correlation. The test statistic was 2.0 for the two 2000 models and 1.8 for the two 2001 models. Therefore, the 2000 and 2001 models show no serial correlation.

Appendix E uses hourly salary as the dependent variable in the 2000/male model, showing that the regression diagnostics are worse when the untransformed salary variable is used. Accordingly, the transformed (natural log) of hourly salary has been used in the analyses.

## 7.2 Model Summaries

The summaries of the models in the 2000 and 2001 decompositions are shown in Table 2 below. As shown by the adjusted  $R^2$  values, the models have good explanatory power, explaining 71.1% of the variance in  $\ln\_wages$  for male employees in 2000 and 72.9% in 2001, and 65.7% of the variance in  $\ln\_wages$  for female employees in 2000 and 66.2% in 2001. The F-test in each of the four ANOVA tests (not shown) was statistically significant at  $p < .001$ , indicating that a relationship existed between  $\ln\_wages$  and all the predictor variables in each model.

**Table 2. Model summaries**

		2000/male	2000/female	2001/male	2001/female
	R	0.846	0.814	0.856	0.817
	R <sup>2</sup>	0.717	0.662	0.733	0.667
	Adjusted R <sup>2</sup>	0.711	0.657	0.729	0.662
	Std. Error of the Estimate	0.1845	0.1674	0.1855	0.1751
Change Statistics	R <sup>2</sup> Change	0.717	0.662	0.733	0.667
	F Change	136.901	124.193	153.900	132.540
	df1	191	191	194	194
	df2	10346	12092	10856	12814
	Sig. F Change	0.000	0.000	0.000	0.000
	Durbin-Watson	1.987	1.986	1.829	1.823

### 7.3 Chow Tests

As the male and female models were regressed separately in order to decompose the effects between the groups, a test for the difference in these effects is required. The Chow test examines the equality of parameters between two subgroups (Hardy, 1993). The null hypothesis is that the parameters are equal, meaning that all the independent variables have uniform effects for both subgroups. The formula of the Chow test is:

$$F = \frac{(RSS_{pooled} - \sum RSS_j) / k + 1}{\sum RSS_j / n_1 + n_2 - 2k - 2}$$

where  $RSS_{pooled}$  is the residual sum of squares (RSS) in the pooled regression (ignoring gender differences),  $\sum RSS_j$  is the sum of the RSS from the two subgroup (gender-based) regressions,  $k$  is the number of predictor variables in the model and  $n_1$  and  $n_2$  are the number of observations in the subgroups. The Chow test statistic follows an F-distribution with  $k+1$  and  $n_1+n_2-2k-2$  degrees of freedom.

Accordingly, a Chow test was performed on data for each year separately, and values for the parameters are shown in Table 3 below. For the 2000 data,  $F(192, 22438) = 7.589$ , and for the 2001 data,  $F(195, 23670) = 6.815$ , the Chow tests were significant at  $p < .0001$ . These significant F-tests indicate that not all the independent variables have uniform effects for males and females, but do not indicate where these differences lie. Given the effect of large group sizes, compared to the number of predictor variables used in this calculation, it is not surprising that the F-tests lead to the rejection of the null hypothesis of uniform effects of parameters.

**Table 3 Chow tests comparing the two subgroups, 2000 and 2001**

	2000	2001
RSS <sub>pooled</sub>	735.695	809.240
RSS <sub>j</sub>	690.834 (Σ[352.134 + 338.700])	766.219 (Σ[373.556 + 392.663])
k	191	194
n <sub>1</sub> +n <sub>2</sub>	22822 (Σ[10538 + 12284])	24060 (Σ[0011051 + 13009])
F	7.589	6.815
df (k+1 , Σ[n <sub>1</sub> +n <sub>2</sub> -2k-2])	df = (192, 22438)	df = (195, 23670)

#### 7.4 Decomposition Analyses

Table 4 below shows the means and standard deviations for the non-indicator variables in the models for the 2000 and 2001 decompositions. For the 2000 decomposition, the gender pay gap was 16.61% (i.e., 3.0505-2.8844), equating to an average earnings advantage for male employees of \$1.18 per hour or just over \$2465 in gross annual salary. Male employees tended to be older than females and have longer tenure. For the 2001 decomposition, the gender pay gap was lower at 15.51%, equating to an average earnings advantage for male employees of around \$1.17 per hour or just over \$2438 in gross annual salary. On average, male employees still tended to be older than female employees and have longer tenure.

**Table 4. Mean values of non-indicator variables**

	2000		2001	
	Males	Females	Males	Females
ln_wages	3.0505 (0.3433)	2.8844 (0.2858)	3.0735 (0.3561)	2.9184 (0.3013)
age	42.55 (9.96)	39.61 (10.65)	42.17 (10.10)	39.37 (10.76)
age <sup>2</sup>	1909.66 (851.37)	1682.58 (873.44)	1880.40 (851.55)	1665.84 (874.43)
tenure	10.37 (9.55)	6.66 (6.35)	10.05 (9.62)	6.55 (6.52)

<sup>a</sup> Standard deviations are shown in brackets.

##### 7.4.1 The “Explained” Gender Pay Gap – Earnings Power Function

71.4% of the 2000 gender pay gap was due to differences in “earnings power” between male and female employees, and this had decreased to 70.0% of the gap in 2001. As shown in Appendix A, the earnings power function of the Blinder-Oaxaca decomposition is given by  $(\bar{X}_{ij}^M - \bar{X}_{ij}^F)\beta_{ij}^M$ . As males are the comparator group in these decompositions, any functions with positive numbers are associated with a higher earnings power for male employees (therefore, increasing the gender pay gap) and functions with negative numbers indicate a higher earnings power for female employees (therefore, decreasing the gap). Earnings power functions close to zero indicate areas containing similar proportions of males and females and/or areas that have little effect on male earnings due to a small associated coefficient.

Table 5 below partitions the overall earnings power function into the different contributor categories of predictor variable. Clearly, occupation had the largest influence, explaining 34.9% and 31.4% of the pay gap for 2000 and 2001 respectively. The next largest source of influence was human capital, proxied by two age variables and one tenure variable, at 28.8% and 29.1% respectively. Employer also played a large role, explaining 7.8% of the pay gap in 2000 and 10.6% in 2001. Notably, ethnicity, region, and the two employment factors (fixed-term versus open-term agreement, collective versus individual agreement) had only minor effects.

I now discuss the contribution of individual indicator variables to the gender pay gap.

In 2000, using the highest classification of occupations, the only occupation category that acted to decrease the gender pay gap was “Trades and Production Workers” (contributed -0.16% of the gender pay gap). The “Managers” category was associated with the largest occupational increase (15.49%), followed by “Associate Professionals” (7.87%), “Science/ Technical” (3.99%), “Professionals” (3.41%), “Personal/Protective Services Workers (2.00%), “Customer Services Clerks” (1.49%), and “Office Clerks” (0.80%). There were 90 occupations that contributed to widening the gender pay gap in the earnings power function. Unsurprisingly then, two of the largest contributions from individual occupations were the “Managers” categories of Administration Manager (NZSCO code = 12222) and General Manager (12111). Administration Manager was the fifth most common job for men in the Public Service and General Manager was the 17<sup>th</sup>, compared to seventh and 30<sup>th</sup> respectively for women. Conversely, 39 occupations effectively acted to decrease the pay gap, as they employed proportionately more women, with the “Office Clerks” occupation of Secretary (41141) making the largest contribution. This occupation was the sixth most common job for women in the Public Service, but only the 58<sup>th</sup> most common for men.

**Table 5. The earnings power function of each contribution source**

<b>Contribution source</b>	<b>2000</b>	<b>2001</b>
Human capital (age, tenure)	0.0478 (28.8%)	0.0451 (29.1%)
Ethnicity <sup>22</sup>	0.0021 (1.2%)	0.0020 (1.3%)
Occupation	0.0579 (34.9%)	0.0488 (31.4%)
Region	-0.0007 (-0.4%)	-0.0010 (-0.6%)
Employer	0.0130 (7.8%)	0.0165 (10.6%)
Employment term	0.0004 (0.2%)	-0.0004 (-0.2%)
Employment agreement	-0.0019 (-1.2%)	-0.0025 (-1.6%)
<b>Total earnings power</b>	<b>0.1187</b> <b>(71.4%)</b>	<b>0.1085</b> <b>(70.0%)</b>
<b>Total gender pay gap:</b>	<b>0.1661</b> <b>(100%)</b>	<b>0.1551</b> <b>(100%)</b>

Notes: <sup>(a)</sup> The individual means and coefficients are available from the author upon request.

<sup>(b)</sup> Percentage contributions to the gender pay gap are shown in brackets.

In 2001, all major occupational groups acted to increase the gender pay gap. The “Managers” category was again associated with the largest occupational increase (contributed 16.77% of the gender pay gap), followed by “Science/Technical” (7.16%), “Professionals” (3.34%), “Customer Services Clerks” (1.97%), “Associate Professionals” (1.11%), “Office Clerks” (0.64%), “Personal/Protective Services Workers” (0.44%), and “Trades and Production Workers” (0.03%). There were 86 occupations that contributed to widening the gender pay gap in the earnings power function. Again, the “Manager” occupations of Administration Manager (12222) and General Manager (12111) provided the largest occupation-based increases. Conversely, 45 occupations effectively acted to decrease the pay gap, with the “Office Clerk” occupation of Secretary (41141) again making the largest contribution to decreasing the gap.

Employers were split into the three groups “small”, “medium”, and “large”, based upon relative employee numbers. In 2000, the small department category effectively decreased the gender pay gap (contributed -1.06% of the gap), with medium and large departments increasing it (3.43% and 5.47% respectively). 21 employers contributed to widening the gap, most noticeably the employer “Large 2”. The largest decrease in the gap was for the employer “Large 6”.

In 2001, the small department category again contributed to decreasing the gender pay gap (contributed -1.43% of the gender pay gap), with medium and large departments increasing it (5.24% and 6.83% respectively). Only 17 of the 37 employers contributed to widening the gender pay gap, with the largest individual contributions originating from two large service delivery departments, “Large 1” and “Large 2”. The largest decrease in the gap was again associated with the employer “Large 6”.

#### **7.4.2 The “Unexplained” Gender Pay Gap – Discrimination Function**

28.6% of the 2000 gender pay gap was due to the unexplained residual (“discrimination”), and this had increased to 30.0% of the gap in 2001. As seen in Appendix A, the “unexplained residual” portion of the Blinder-Oaxaca decomposition is given by  $[(\beta_{ij}^M - \beta_{ij}^F)x_{ij}^F + (\beta_0^M - \beta_0^F)]$ . Any differences in male and female earnings from this function, therefore, arise from the difference in male and female coefficients. In other words, the function is comparing the “value” of being male compared to the “value” of being female. As males are the comparator group in these decompositions, any functions with positive numbers are associated with a higher earnings power for male employees (therefore, increasing the gender pay gap) and functions with negative numbers indicate a higher earnings power for female employees (therefore, decreasing the gap). “Discrimination” functions close to zero indicate areas associated with similar earning weights for male and female employees and/or areas that employ only small proportions of men.

Table 6 below partitions the overall unexplained residual into the different contributor categories of predictor variable. Again, occupation was the largest contributor category in both 2000 and 2001, at 29.4% and 28.0% respectively. In other words, around 30% of the difference in male and female earnings is due to an occupation-based salary advantage for male employees.

**Table 6. The unexplained residual associated with each contribution source**

<b>Contribution source</b>	<b>2000</b>	<b>2001</b>
Human capital (age, tenure)*	0.0112 (6.7%)	-0.0187 (-12.1%)
Ethnicity	0.0001 (0.1%)	0.0013 (0.8%)
Occupation	0.0488 (29.4%)	0.0434 (28.0%)
Region	-0.0131 (-7.9%)	-0.0120 (-7.8%)
Employer	-0.0198 (-11.9%)	0.0151 (9.7%)
Employment term	0.0029 (1.7%)	0.0029 (1.9%)
Employment agreement	0.0173 (10.4%)	0.0146 (9.4%)
<b>Total unexplained residual</b>	<b>0.0474</b> <b>(28.6%)</b>	<b>0.0466</b> <b>(30.0%)</b>
<b>Total gender pay gap:</b>	<b>0.1661</b> <b>(100%)</b>	<b>0.1551</b> <b>(100%)</b>

Notes: (a) The individual means and coefficients are available from the author upon request.

(b) Percentage contributions to the gender pay gap are shown in brackets.

(\*) These values include the difference in male and female intercepts.

In 2000, all major occupation categories acted to increase the gender pay gap. The largest occupational increase was from “Office Clerks” (contributed 13.99% of the gender pay gap), followed by “Associate Professionals” (7.94%), “Professionals” (3.02%), “Customer Services Clerks” (2.02%), “Managers” (1.76%), “Science/Technical” (0.44%), “Personal/Protective Services Workers” (0.17%), and “Trades and Production Workers” (0.02%). There were 93 occupations that contributed to widening the gender pay gap in the unexplained residual. The two occupations providing the largest earnings bias favouring male employees were the “Office Clerks” occupations of Secretary (NZSCO code = 41141) and General Clerk (41443). These two occupations had comparatively larger male coefficients. Conversely, 36 occupations effectively acted to decrease the pay gap, as they were associated with an earnings bias favouring female employees, with the “Personal/Protective Services Workers” occupation of Caregiver (51316) making the largest individual contribution (due to a comparatively larger female coefficient).

In 2001, “Trades and Production Workers” had a nil effect on the gender pay gap (contributed 0.00% of the gender pay gap). The largest occupational increase was again from “Office Clerks” (13.00%), followed by “Associated Professionals” (6.04%), “Professionals” (4.01%), “Managers” (2.28%), “Customer Services Clerks” (1.70%), “Personal/Protective Services Workers” (0.56%), and “Science/Technical” (0.42%). 85 occupations contributed to widening the gender pay gap in the unexplained residual. The two occupations providing the largest earnings bias favouring male employees were, again, the two “Office Clerks” occupations of Secretary (41141) and General Clerk (41443). Conversely, 46 occupations effectively acted to decrease the pay gap, as they were associated with an earnings bias favouring female employees, with the “Associate Professionals” category of Quarantine and Agriculture Ports Officer (33312) making the largest occupational contribution here.

Ethnicity and employment term continued to have only minor influence on the gender pay gap. However, employment agreement accounted for around one-third of the unexplained residual due to the larger male coefficients.<sup>23</sup> Therefore, there is an earnings premium for being on an individual agreement, especially for male employees, even when occupation has been taken into account.

## **7.5 Summary of Findings**

The gender pay gap in the New Zealand Public Service, as measured by the difference in geometric means between the log wages of male and female employees, decreased from 16.6% in 2000 to 15.5% in 2001. In other words, female employees earned 83.4% of male earnings in 2000 and 84.5% in 2001. In 2000, this average male earnings advantage was \$1.18 per hour or just over \$2465 in gross annual salary. In 2001, this advantage had decreased slightly to around \$1.17 per hour or just over \$2438 in gross annual salary.

Blinder-Oaxaca decompositions for gender were performed for 2000 and 2001 data separately. These decompositions showed that 71.4% of the 2000 gender pay gap was due to differences in “earnings power” between male and female employees, and this decreased to 70.0% of the gap in 2001. These findings suggest that the adjusted gender pay gap, as measured by the unexplained residual, is only 4.7% for both 2000 and 2001 when gender differences in human capital (age, tenure, ethnicity) and employment characteristics (occupation, region, employer, employment term and employment agreement) are taken into account. This residual pay gap equates to an average male earnings advantage of \$0.34 per hour or just over \$703.11 in gross annual salary for 2000, and \$0.35 per hour or \$732.89 in gross annual salary for 2001.

### **7.5.1 Occupation**

Occupation, which was measured at the 5-digit NZSCO level, was the largest contributor to the gender pay gap. This finding is consistent with previous studies that have incorporated measures of occupation. Male and female employees tended to work in different occupations, with men tending to work in higher paid occupations such as the Manager categories of 12222 (Administration Manager) and 12111 (General Manager). There was also an overall male premium attached to occupations, so that even when men and women worked in the same occupation the male employees tended to be higher paid. As human capital variables and employment characteristic variables were included in the model, the cause of this result is somewhat unclear. However seniority in the job,<sup>24</sup> qualifications, and direct measures of experience were not included in the decompositions due to a lack of data. Given that employees in the New Zealand Public Service tend to be older than those in the general labour market, that women have only relatively recently entered tertiary education in higher numbers, and that women tend to take more time out of employment for child bearing and rearing, the age and tenure measures used in the decompositions may not have adequately specified the human capital factors relevant to explaining gender differences in wages for the same occupation.

### **7.5.2 Human Capital**

While male employees averaged higher levels of human capital as measured by age, age<sup>2</sup> and tenure, the wage-based return on these three variables is difficult to interpret. There was a higher return to male employees in 2000 and a higher return to female employees in 2001 (6.7% compared to -2.1%). This alteration, in effect, is likely to be mainly due to two factors:

- The annual Public Service turnover rate was 22% for 2001. This is the actual turnover between the 2000 and the 2001 data. It is likely that not all exiting employees were replaced, and that replacements had different human capital to the original employees.

This is especially true for male employees who are generally older and tend to have longer tenure. This effect is supported by the differences in mean age and tenure between the 2000 and 2001 datasets as shown in Table 4.

- The misspecification of human capital in the decompositions. As noted earlier, seniority, qualifications, and experience were not included in the decompositions due to a lack of data. Thus, the decompositions may not have adequately specified the human capital factors relevant to explaining wages.

There were similar proportions of male and female employees in ethnic minorities. The returns for being in an ethnic minority were small but positive for both the earnings power function and the unexplained residual, suggesting that the gender pay gap also occurs for ethnic minority employees.

### **7.5.3 Employer**

Even with occupation incorporated into the decompositions, employers also played a role in explaining the gender pay gap. Just like occupation, there was some gender segregation between employers, although the returns from this segregation tended to benefit male employees in 2000 with the earnings advantage shifting to female employees in 2001. The reason for this is unknown, but may be related to the misspecification problems outlined in the human capital summary above. It is unlikely to result from either the addition of the Serious Fraud Office in the 2001 dataset due to the very small number of employees (< 40) in this department, or from the use of the Ministry of Pacific Island Affairs as the reference category. The Ministry of Pacific Island Affairs contains a very small number of staff (< 40), the proportions of male and female employees were similar in the datasets, and the average salary in the organisation was similar to the overall Public Service average.

### **7.5.4 Region**

While the overall proportions of male and female employees in different regions were similar, there was a female earnings advantage for working outside Wellington (the region reference category). This finding suggests that females employed in occupations outside Wellington may average higher earnings than male employed in occupations outside Wellington. Many of the highest-paid occupations in the New Zealand Public Service are in Wellington, particularly those “Manager” occupations containing disproportionate numbers of males. It appears that occupations based in Wellington could be the primary determinants of the gender pay gap.

### **7.5.5 Employment Characteristics**

Similar proportions of male and female employees were on fixed-term agreements. Being on a fixed-term agreement was associated with slightly higher earnings for male employees. While similar proportions of male and female employees were on individual agreements, there was a male earnings advantage of 10.4% in 2000 and 9.4% in 2001 for being on an individual agreement. This finding is significant, given that occupation has already been taken into account.

## **8. Discussion and Commentary on Methodology**

### **8.1 Discussion of Results**

The decompositions performed in this study produced a New Zealand Public Service gender pay gap of 16.61% for June 2000 and 15.5% for June 2001, based on the geometric means of the natural log of salary. 71% of the 2000 pay gap was due to male employees having more earnings power that dropped to 70% of the gap in 2001. These results need to be compared to the Human

Resource Capability survey results to determine the extent of any bias introduced by the data cleaning used in this project. Using the arithmetic means of hourly salary, the pay gap was 16.95% for 2000 and 16.04% for 2001. These results also compare reasonably favourably to the gender pay gap reported in the two State Services Commission Human Resource Capability surveys of 19% for 2000 and 17% for 2001. To reiterate, this project used the Human Resource Capability survey datasets, although employees with missing information were removed from the decompositions. A large difference in the arithmetic means would suggest that the data cleaning process was biased.

Occupation had the largest effect on the gender pay gap, with more than 30% of the gap for both 2000 and 2001 arising from differences in the male and female participation rates in the different Public Service occupations, with males tending to work in higher paid occupations, and a further 28-29% attributable to a male earnings advantage within occupations. While this earnings advantage occurs once human capital factors such as age and tenure are taken into account, variables such as seniority within occupation were not analysed due to a lack of data. Given that male employees tended to have higher levels of human capital such as tenure, the male earnings advantage is suggestive of a tendency for male employees to be more senior than female employees within occupations, rather than male employees earning more at the same level of seniority in the same occupation.

While many people think of the New Zealand Public Service as a purely Wellington industry, one department is entirely domiciled in Auckland (Serious Fraud Office) and only five have all their staff based in Wellington (Crown Law Office, Ministry of Defence, the Treasury, Ministry of Women's Affairs, Ministry of Youth Affairs). There was little difference in the proportions of male and female employees working outside Wellington; however, there was an earnings advantage for women in these regions of almost 8% for each year. Given that the head offices of each department are in Wellington (apart from the Serious Fraud Office), that higher paid occupations tend to be in head offices rather than regional offices or sites, and that males have a higher share of the higher paid occupations, this regional finding suggests that much of the gender pay gap is produced by male and female employment differences in Wellington. In other words, occupations that contain proportionately more employees outside of Wellington – and have larger numbers of employees (such as customer services officers) – are more likely to pay women a relatively higher salary. These may also be occupations where human capital factors – such as age – have relatively lower impact on salary.

Finally, an individual employment agreement was associated with an earnings advantage for men (10.4% in 2000, 9.4% in 2001) even though similar proportions of males and females were on individual employment agreements. This effect is independent of occupation, employer, region, tenure, and age. This finding suggests that individual employment agreements provide better working conditions for male employees.

## **8.2 Critique of Methodology**

The study findings are based on the econometric Blinder-Oaxaca decomposition method. Section 5.4 of this report discusses the problems associated with decompositions methods and a summary of the relevant commentary is reproduced here. First, it has been argued that decomposition methods can only examine post-hiring wage discrimination. Even when occupation is incorporated as a set of indicator variables, occupation is assigned after hiring the employee. Therefore, the decomposition cannot measure discrimination in the hiring decision and, as such, possibly underestimates the gender pay gap (assuming that discrimination in hiring operates in favour of men at the expense of women).

Second, the method is affected by the index number problem, as the choice of reference group (male employees or female employees) affects the results. In practice, this problem is most frequently negated by all studies reporting results based on the male wage structure, and this was the approach adopted for this study. Thus the findings from this study are directly comparable to those from other studies using the Blinder-Oaxaca decomposition.

Third, as the decomposition is effectively a comparison of two identically specified regression models that typically incorporate categorical variables (such as ethnicity, occupation), an omitted reference category is required for each set of categorical variables. In some cases, the choice of reference category is obvious (e.g., NZ European for ethnicity), and in other cases the decision is much more arbitrary (e.g., the choice for employer, occupation). Where the decision is less clear-cut it could be preferable to include all categories. This option is not possible using a regression methodology as the model would be over-specified<sup>25</sup>. A related problem was that occupation indicator categories that contained only male or only female employees had to be removed from the analysis (this had no effect on the decomposition results as the mean and associated coefficient for the other gender's regression model are both zero, causing that occupation to provide no contribution to either the earnings function or the unexplained residual). In both these cases information contained in the dataset has been omitted from the analysis. Given these two problems associated with categorical variables and the large, artefactual effects of proxy variables on regression models, an examination of the usefulness of regression analyses on the gender pay gap would be timely.

Finally, as the regression models in the decompositions are based on the ordinary least squares method, the decomposition results are a statement about the effects on the earnings of the "average" male employee compared to the "average" female employee. These "average" results may not be reflective of the results for employees at other earnings percentiles (e.g., 10<sup>th</sup> percentile, 75<sup>th</sup> percentile) and, therefore, may be of less utility for understanding the earnings dynamics of employees earning away from the mean.

The decomposition does provide useful and interesting information. Rather than replacing the method with an alternative analysis technique, a more useful approach would appear to be supplementing the analysis with a method more attuned to incorporating categorical variables such as the Classification and Regression Trees (CART) method. A decision tree method seems to be ideally suited to such a research question as the group of interest (male versus female employees) is already known in the dataset and a set of binary decision rules (yes/no) are suitable for analysis that combine continuous and categorical data. A second major advantage of such a supplementary methodology is that it does not require the data to meet all the rigorous assumptions associated with regression such as the dependent variable having a normal distribution and constant variance. Alternative data mining techniques could also prove valuable.

## 9. References

- Acemoglu, D. (2001). *Human capital policies and the distribution of income: A framework for analysis and literature review*. Working Paper 01/03, The Treasury: Wellington.
- Barnett, J. (1997). *Gender wage gap: Part Two. An assessment of the relative impact of each industry*. NZIER for the Ministry of Women's Affairs: Wellington.
- Bayard, K., Hellerstein, J., Neumark, D., and Troske, K. (1999). *New evidence on sex segregation and sex differences in wages from matched employee-employer data*. NBER Working Paper W7003, National Bureau of Economic Research: Massachusetts.
- Black, D.A. (1995). Discrimination in an equilibrium search model. *Journal of Labor Economics*, 13(2), 309-334.
- Blau, F.D. and Kahn, L.M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, 14(4), 75-99.
- Blinder, A.S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4), 436-455.
- Chauvin, K.W. and Ash, R.A. (1994). Gender earnings differentials in total pay, base pay, and contingent pay. *Industrial and Labor Relations Review*, 47(4), 634-648.
- Cook, D. and Briggs, P. (1997). *Gender wage gap: Scenarios of the gender wage gap*. NZIER for the Ministry of Women's Affairs: Wellington.
- Deeks, J., Parker, J. and Ryan, R. (1994). *Labour and employment relations in New Zealand*. (2<sup>nd</sup> ed.). Longman Paul: Auckland.
- Dex, S., Sutherland, H. and Joshi, H. (2000). Effects of minimum wages on the gender pay gap. *National Institute Economic Review*, 173, 80-88.
- Dixon, S. (1996a). The distribution of earnings in New Zealand 1984-94. *Labour Market Bulletin 1996:1*, 45-100.
- Dixon, S. (1996b). Labour force participation over the last ten years. *Labour Market Bulletin 1996:2*, 71-88.
- Dixon, S. (1998). Growth in the dispersion of earnings: 1984-97. *Labour Market Bulletin, 1998:1and2*, 71-107.
- Dixon, S. (2000). *Pay inequality between men and women in New Zealand*. Occasional Paper 2000/1, Labour Market Policy Group: Wellington.
- Fortin, N.M. and Lemieux, T. (1998). Rank regressions, wage distributions, and the gender gap. *Journal of Human Resources*, 33(3), 610-643.
- Groshen, E.L. (1991). The structure of the female/male wage differential: Is it who you are, what you do, or where you work? *Journal of Human Resources*, 26(3), 457-472.
- Gupta, N.D., Oaxaca, R.L., and Smith, N. (1998). *Wage dispersion, public sector wages and the stagnating Danish gender wage gap*. Working Paper 98-18, Centre for Labour Market and Social Research: Denmark.
- Hardy, M.A. (1993). *Regression with dummy variables*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-093. Newbury Park, CA: Sage.
- Harkness, S. (1996). The gender earnings gap: Evidence from the UK. *Fiscal Studies*, 17(2), 1-36.

- Hyman, P. (1992). *The earnings gap and pay equity: Developments in five countries*. Proceedings of the 5<sup>th</sup> Labour, Employment and Work in New Zealand Conference, Victoria University: Wellington.
- Jones, F.L. (1983). On decomposing the wage gap: A critical comment on Blinder's method. *Journal of Human Resources*, 18(1), 126-130.
- Juhn, C., Murphy, K.M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3), 410-442.
- Kirkwood, H. (1998). *Exploring the gap: An exploration of the difference in income received from wages and salaries by women and men in full-time employment*. Statistics New Zealand: Wellington.
- Lambert, S.F. (1993). Labour market experience in female wage equations: does the experience measure matter? *Applied Economics*, 25, 1439-1449.
- Laroche, M. and Mérette, M. (2000). *Measuring human capital in Canada*. Working Paper 2000-05, Canada Finance: Ontario.
- Levi, M.D. (1973). Errors in the variables bias in the presence of correctly measured variables. *Econometrica*, 41(5), 985-986.
- Macpherson, D.A. and Hirsch, B.T. (1995). Wages and gender composition: Why do women's jobs pay less? *Journal of Labor Economics*, 13(3), 426-471.
- Madden, D. (1999). *Towards a broader explanation of male-female wage differences*. Working Paper 99-11, Department of Political Economy: College Dublin.
- Miller, P.W. (1994). Effects on earnings of the removal of direct discrimination in minimum wage rates: A validation of the Blinder decomposition. *Labour Economics*, 1, 347-363.
- Morrison, P.S. (2001). Employment. *Asia Pacific Viewpoint*, 42(1), 1-24.
- Naur, M. and Smith, N. (1996). *Cohort effects on the gender wage gap in Denmark*. Working Paper 96-05, Centre for Labour Market and Social Research: Denmark.
- Nesterova, D.V. and Sabirianova, K.Z. (1999). *Investment in human capital under economic transformation in Russia*. Working Paper 99/04, Economic Education and Research Consortium: Russia.
- Nielsen, H.S. (1998). *Two notes on discrimination and decomposition*. Working Paper 98-01, Centre for Labour Market and Social Research: Denmark.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693-709.
- Oaxaca, R.L. and Ransom, M.R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, 61, 5-21.
- Oaxaca, R.L. and Ransom, M.R. (1997). *Identification in detailed wage decomposition*. Working Paper 97-12, Centre for Labour Market and Social Research: Denmark.
- O'Dea, D. (2000). *The changes in New Zealand's income distribution*. Working Paper 00/13, The Treasury: Wellington.
- Ott, R.L. (1993). *An introduction to statistical methods and data analysis* (4<sup>th</sup> edition). Duxbury Press: California.

- Preston, A.C. (2000a). *Deregulation and relative wages: Stability and change in Australia*. Discussion Paper 00/4, Women's Economic Policy Analysis Unit, Curtin University of Technology: Perth.
- Preston, A.C. (2000b). *Equal pay in W.A.* Discussion Paper 00/6, Women's Economic Policy Analysis Unit, Curtin University of Technology: Perth.
- Preston, A.C. (2000c). *The Hayekian paradigm: Some distributive consequences*. Proceedings of the Newcastle 2000 conference, Association of Industrial Relations Academics of Australia and New Zealand.
- Reilly, K.T. and Wirjanto, T.S. (1998). *Does more mean less? The male/female wage gap and the proportion of females at the establishment level*. Working Paper 98-04, Centre for Labour Market and Social Research: Denmark.
- Reiman, C. (2001). *The gender wage gap in Australia: Accounting for linked employer-employee data from the 1995 Australian workplace industrial relations survey*. Discussion Paper 54, National Centre for Social and Economic Modelling: University of Canberra.
- State Services Commission. (2000). *State Services Commission human resource capability survey of public service organisations and selected state sector organisations, as at 30 June 2000*. State Services Commission: Wellington. (<http://www.ssc.govt.nz/siteset.htm>)
- State Services Commission. (2001). *Human resource capability: survey of public service departments as at 30 June 2001*. State Services Commission: Wellington. (<http://www.ssc.govt.nz/siteset.htm>)
- Suen, W. (1997). Decomposing wage residuals: Unmeasured skill or statistical artifact? *Journal of Labor Economics*, 15(3), 555-566.
- Sung, Y, Zhang, J. and Chan, C. (2000). *Gender wage differentials and occupational segregation in Hong Kong, 1981-1996*. HIEBS Working Papers: Hong Kong Institute of Economics and Business Strategy.
- Swaffield, J. (2000). *Gender, motivation, experience and wages*. Discussion Paper 0457, Centre for Economic Performance: London School of Economics and Political Science.
- Szakats, A. (1988). *Law of employment*. (3<sup>rd</sup> edition). Butterworths: Wellington.
- Waldfoegel, J. (1998). *Understanding the "family gap" in pay for women with children*. *Journal of Economic Perspectives*, 12(1), 137-156.

## Appendix A: The Blinder-Oaxaca Decomposition Method

This section will introduce the decomposition model, the model assumptions, and will discuss some problems with the general model.

### A1 The Decomposition Model

From the 1970s, repeated attempts have been made to statistically identify the causes of the gender pay gap. Most gender pay gap modelling has used multiple regression for the methodology, with some measure of natural log (log) earnings as the dependent variable. In particular, since the separate publication in 1973 of the two landmark papers by Oaxaca and Blinder, the type of comparative regression modelling used has been a decomposition method. In this subcategory of “logged linear” modelling, gender is not included as an indicator variable<sup>26</sup> in the regression. Instead, the model estimates the separate coefficients for males and females on each independent variable. For each gender, the equation is the standard multiple regression model:

$$(1) y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j + \varepsilon$$

Using the standard multiple regression assumption of  $E(\varepsilon) = 0$ , the expected value of  $y$  is given by:

$$(2) E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j$$

In the Blinder-Oaxaca decomposition, the male and female equation 2 is defined as:

$$(3) \ln(\hat{W}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots + \beta_j x_j$$

where  $\ln(\hat{W})$  is log wages (frequently log hourly wages),  $\beta_0$  is the intercept,  $\beta_j$  is the coefficient of the  $j$ th variable, and  $x_j$  is the mean of the  $j$ th variable. The means are calculated from the dataset under analysis in the normal way, so that a mean is also produced for any indicator variable.<sup>27</sup>

Thus the difference in the gender wage gap (i.e., the Blinder-Oaxaca decomposition) arises from:

$$(4) \ln(\hat{W}^M) - \ln(\hat{W}^F) = (\beta_0^M - \beta_0^F) + (\beta_1^M x_1^M - \beta_1^F x_1^F) + (\beta_2^M x_2^M - \beta_2^F x_2^F) \dots + (\beta_j^M x_j^M - \beta_j^F x_j^F)$$

This model can be summarised as:

$$(5) \ln(\hat{W}^M) - \ln(\hat{W}^F) = [(x_{ij}^M - x_{ij}^F) \beta_{ij}^M] + [(\beta_{ij}^M - \beta_{ij}^F) x_{ij}^F + (\beta_0^M - \beta_0^F)]$$

where  $x_{ij}^M$  is the vector of means from the male equation,  $x_{ij}^F$  is the vector of means for the female equation,  $\beta_{ij}^M$  is the vector of coefficients from the male equation and  $\beta_{ij}^F$  is the vector of coefficients from the female equation.

Equation 5 is equivalent to Oaxaca (1973) equation 13. Thus, the gender pay gap in this logged linear method is due to two functions:

- the actual difference in the variables (e.g., occupation, age), as shown by the difference in mean values between males and females (the “earnings power” function, which shows the level of the variable); and
- the difference in the effects of these variables, as shown by the differences in the coefficients for males and females (the “discrimination” function or unexplained residual). This function contains the difference in the intercept terms from both the male and female equations.

The main advantage of this decomposition method, over traditional linear regression methods incorporating an indicator variable for gender, is the inclusion of this second function. In the traditional methodology, the coefficients of the non-gender predictor variables are the same for males and females, by definition, with the difference between the genders captured in the one-gender indicator variable. That is, the traditional regression assumes that the non-gender terms are equivalent for males and females whereas the Blinder-Oaxaca decomposition allows these terms to vary.

The validity of the Blinder-Oaxaca decomposition was tested by Miller (1994) who compared the earnings of Australian workers between 1973 and 1989.<sup>28</sup> The cross-sectional methodology meant that a matched employee-employee dataset was not used. The decomposition method detected improvement in log female wages compared to log male wages in the “discrimination” function of the equation, as predicted by the changes in Australian labour law.

The gender wage gap shown by this method is derived from the geometric mean rather than the arithmetic mean for male and female wages. The arithmetic mean is given by the equation  $\frac{x_1 + x_2 \dots + x_n}{n}$ . The geometric mean is given by the equation  $\sqrt[n]{x_1 * x_2 \dots * x_n}$ . The geometric mean is used instead of the arithmetic mean because the predictor variable is in log form and logarithms alter algebraic operations. Therefore, care is needed when comparing the gender pay gap results between a logarithm method (e.g., the Blinder-Oaxaca decomposition) and a non-logarithm method.

## A2 The Quantification of Discrimination

Statistical modelling has attempted to determine the underlying drivers of the gender pay gap. As the pay gap is based on the differences in average earnings between males and females, investigations have centred on attempting to model the reasons for this result using earnings as the dependent variable. There are two types of parameters typically included in these models:

- Those that have prima facie validity in explaining earnings, but may indicate an indirect discriminatory effect on earnings (e.g., years in workforce, educational attainment, industry, occupation); and
- Those that are person-specific and suggest a direct discriminatory effect on earnings (e.g., ethnicity).

Gender discrimination has been deemed to exist where the effect of a variable differs depending on whether one is male or female.<sup>29</sup> In statistical modelling, the effect is shown by the coefficient assigned to the variable by the analysis. In some of the literature, a difference in effects between the male and female models has been assumed to mean that one group receives

greater returns from the labour market (in this case, higher wages) compared to the other group even when the value of the variable under consideration (e.g., years in workforce) is the same. This quantification of discrimination is simplistic for three main reasons:

- It assumes the variables that explain the difference between male and female wages have been measured exhaustively. Models of the gender pay gap have changed across time and national borders, with variables added and dropped. Even with apparently detailed models, research still tends to show an unexplained gap. Clearly, we have an imperfect collection of variables that are used to model the gender pay gap.
- It assumes the variables have been measured without error. Variables in models of the gender pay gap have been respecified over time.
- It assumes that any difference in effects (coefficients) must be purely due to discrimination. This ignores a number of alternative explanations. First, the direction of the effect may actually operate in reverse (e.g., a women in a low paid job decides to have children because her exit from the labour market has a smaller effect on her total income than if she was in a higher paid job). Second, misspecification of even one variable in the model will alter the effect of all variables. This alteration occurs regardless of whether the misspecification is due to an included variable (e.g., inadequate proxy variable) or an omitted variable. Third, an unmeasured variable may influence both the explanatory variables and hourly wages.

Even extremely detailed models of the gender pay gap have tended to reduce, rather than remove, the unexplained gap (“unexplained residual”) between male and female hourly wages. This residual is simply defined as the portion of the gender pay gap that is represented by the second term in the decomposition model. The existence of an unexplained residual is undesirable for two main reasons, both of which are concerned with model misspecification:

- The residual typically has a larger effect on the gender pay gap than many of the individual variables in the model. In many models it accounts for more than 10% of the gap between male and female earnings. A large residual could mean that important factors have been omitted from the model under consideration.
- One or more variables in the model may not be a valid proxy. While the data sources that have been used to model the gender pay gap contain reliable data, and the variables have been included because of their face validity, there may be problems with the construct validity. Construct validity is the extent to which the variable represents the “thing” one is trying to measure, for example, the use of tenure to proxy – or represent – work experience. Obviously, a direct measure of work experience (e.g., years in occupation) would have more construct validity than tenure because the measure would contain no error, but no data is available for that ideal measure so the proxy is used instead. The greater the number of proxy variables used, and the more error associated with each proxy, the larger the unexplained residual.

### **A3 Human Capital in The Model**

From an economic perspective, human capital is essentially the interaction between education and experience at the microeconomic (individual) level, which can be aggregated up to indicate human capital at the macroeconomic (e.g., country) level. Here, education and experience are used as proxies for the human capital factors of skills and abilities (and productivity). This type of analysis traditionally uses a Mincer equation (named after Mincer’s 1974 work, see e.g., Fortin and Lemieux, 1998). The basic Mincer equation is:

$$(6) \ln(\hat{W}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2$$

where log wages is predicted by schooling ( $X_1$ ), and experience<sup>30</sup> is fitted using a linear ( $X_2$ ) and a quadratic ( $X_2^2$ ) term (see e.g., Nesterova and Sabirianova, 1999). This type of model is an income-based method, as the relationship is between human capital and log wages (Laroche and Mérette, 2000). The other two methods of measuring human capital are cost-based, where the depreciated value of inputs (investments such as education and health) are used, and output-based, where school enrolment rates, average years of schooling, or adult literacy rates are used. The Blinder-Oaxaca decomposition also uses an income-based model of human capital, as the predictor variables include human capital factors.

#### **A4 General Limitations of Decomposition Methods**

Decomposition methods are performed on datasets of employees that capture human capital, personal, and employer characteristics. Even when a panel study has been used to examine wage gap differences between time-1 and time-2, the analysis is performed on existing employees. This means that gender pay decompositions suffer from the weakness that they only measure post-hiring wage discrimination. Given that the hiring process is based on human capital considerations and is also prone to the influence of discrimination, and that minorities tend to have lower human capital levels anyway, the decompositions probably underestimate the impact of discrimination on wages.

Madden (1999) compared three wage decompositions of the British 1995 Family Resources Study: the standard Blinder-Oaxaca method, Blinder-Oaxaca method incorporating estimates of employment probabilities, and a Blinder-Oaxaca method that includes a measure of selectivity bias (the inverse Mills ratio). In contrast to the findings of Neumann and Oaxaca, Madden found that discrimination at labour market entry was of considerable importance, although selectivity bias was not demonstrated.

#### **A5 Problems With Blinder-Oaxaca Decomposition**

##### **A5.1 The Index Number Problem**

The most cited problem (e.g., Oaxaca, 1973) is termed the “index number problem”. In other words, the choice of reference group in the model affects the results produced by the decomposition. To illustrate this point, the most common method to define the gender pay gap is shown in equation 5, reproduced here:

$$\ln(\hat{W}^M) - \ln(\hat{W}^F) = [(x_{ij}^M - x_{ij}^F)\beta_{ij}^M] + [(\beta_{ij}^M - \beta_{ij}^F)x_{ij}^F + (\beta_0^M - \beta_0^F)]$$

In this method, the male wage structure is used to assess the gender pay gap. The alternative is to use the female wage structure, altering equation 5 to:

$$(7) \ln(\hat{W}^M) - \ln(\hat{W}^F) = [(x_{ij}^M - x_{ij}^F)\beta_{ij}^F] + [(\beta_{ij}^M - \beta_{ij}^F)x_{ij}^F + (\beta_0^M - \beta_0^F)]$$

The most common way to remove this problem has been to report the results based on the male wage structure, thereby standardising the literature. This tradition is repeated in the current study.

##### **A5.2 The Use of Indicator Variables**

Jones (1983) has suggested that the Blinder-Oaxaca decomposition has two main faults: the intercept and indicator variable coefficients are influenced by the reference group(s) used for the indicator variable(s) in the model; and the intercept is influenced by the choice of scale for continuous variables in the model. Jones suggests that the problem is so critical that interpretation of the intercept is meaningless.

There is no estimate invariance for the coefficients and intercept when indicator variables are included in the model (Oaxaca and Ransom, 1997). Interpretable results occur when only one category of indicator variable (e.g., ethnicity) is included in the regression, as the intercept result is simply added to the coefficient to produce the required estimate. When more than one category of indicator variable is included in the model, the individual effects of the indicator variables cannot be determined as it is unclear how the intercept term should be applied. In both cases, however, the overall decomposition result and the coefficients of the non-indicator variables are invariant. Finally, Oaxaca and Ransom suggest that the continuous variable problem does not occur in practice as the scale of these variables is not as arbitrarily specified. For example, experience is always measured in years.

Nielsen (1998) proposes that the 1 and 0 values of indicator variables should be replaced with the proportion of observations in each group. This substitution ensures that the decomposition results are invariant to the choice of reference group. Regardless of the number of groups, the overall decomposition result and the coefficients and intercept are invariant to reference group. For example, with two indicator variables the discrimination portion of the model becomes:

$$\alpha_1[(\hat{\beta}_0^M - \hat{\beta}_0^F) + (\hat{\beta}_{1j}^M - \hat{\beta}_{1j}^F) + (\hat{\beta}_{2j}^M - \hat{\beta}_{2j}^F)] + \alpha_2[(\hat{\beta}_0^M - \hat{\beta}_0^F) + (\hat{\beta}_{1j}^M - \hat{\beta}_{1j}^F)] + \alpha_3[(\hat{\beta}_0^M - \hat{\beta}_0^F) + (\hat{\beta}_{2j}^M - \hat{\beta}_{2j}^F)] + \alpha_4[\hat{\beta}_0^M - \hat{\beta}_0^F]$$

where  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ , and  $\alpha_4$  are the proportions of females in each of the four possible groups,<sup>31</sup>  $\beta_{1j}$  is the coefficient of the first indicator variable and  $\beta_{2j}$  is the coefficient of the second indicator variable. The advantage of this method of specification is that the contribution of the individual levels of the indicator variables can be estimated. This is the method of choice where the decomposition model contains more than one indicator variable, and where the effect of the indicator variables is desired.

### ***A5.3 The Model Evaluates at the Average***

Because a regression method is used, the equations produced in the model evaluate the gender pay gap at the mean male and mean female wages. Juhn, Murphy, and Pierce (1993) analysed the wage information of full-time employed male workers from the US Current Population Survey (1964 to 1990 years) and the 1960 decennial census. Wage information was deflated to 1963/1964 dollars. Decompositions incorporating education and experience were run on the 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles, although most comparisons were made between the 10<sup>th</sup> and 90<sup>th</sup> percentiles, the 10<sup>th</sup> and 50<sup>th</sup>, and the 50<sup>th</sup> and 90<sup>th</sup> percentiles. They found that increased inequality in wages was due to increasing skill prices (education, experience, occupation), apparently due to increased demand for skilled and not unskilled employees.

Suen (1997) argues that the Juhn, Murphy, and Pierce (1993) amendment to the Blinder-Oaxaca decomposition – the decomposition of the residual into changes in the dispersion of the residual wage distribution (standard deviation) and percentile ranks – is incorrect. This means that the associated interpretation of the dispersion of the residual wage distribution as equivalent to returns to skill (i.e., price effects) and the interpretation of the percentile ranks as equivalent to levels of unmeasured skill (i.e., quantity effects) are also incorrect. Suen points out that

dispersed distributions have thicker tails. That is, the percentile rank for female wages will tend to increase as wage inequality increases, even if price and quantity effects remain constant. Suen suggests that an increase in the price of skill would be shown by larger wage gains for individuals with large wage residuals.

#### ***A5.4 The Influence of Proxy Variables***

As stated in Section A5.2, a proxy variable is one used as a “stand-in” or approximation for a variable that is harder to measure or collect. For example, age is often used to proxy work experience. Swaffield (2000) found that the use of potential labour market experience, rather than actual experience, increased the unexplained residual. This finding suggests that the use of any proxy measures may artificially inflate the unexplained residual in decomposition models<sup>32</sup>.

Lambert (1993) examined the use of proxy measures of experience. All four proxies of experience used (age, Mincer estimation of experience, Mincer estimation allowing for current dependent children, and Mincer estimation allowing for both current and past dependent children) produced significant biases in the coefficients, not only biasing the experience variable and the constant but also, as an artefact, decreasing the effect of education in the same model. The first two specifications of experience (age, Mincer estimation without inclusion of children) produced the largest biases. The implication of these findings is that the use of even one misspecified variable could have significant effects throughout the model. The result is that the ordinary least squares estimate of the misspecified variable will be biased towards zero, even if all other variables are measured without error (Levi, 1973).

## **Appendix B: Removing “Constant” Occupations**

### **2000 Dataset**

Initially the 2000 dataset included 198 occupations. This was reduced to 129 after occupations containing solely men or solely women were removed from the models. The list of “sole gender” occupations below does not mean that these occupations are segregated by gender in the Public Service, as only a subset of employees – albeit a large subset – were incorporated into the first model.

There were 16 occupations that contained no men, which related to 47 women in the dataset. These observations were removed from the analysis dataset, and the affected occupations were: 12213 – Production Manager – Manufacturing, 12221 – Health Services Manager, 21412 – Resource Management Planner, 22251 – Dietitian/Public Health Nutritionist, 33121 – Insurance Representative, 33614 – Sub-editor, 41131 – Reserved, 41223 – Survey Interviewer, 41312 – Dispatch and Receiving Clerk, 41421 – Mail Sorting Clerk, 41441 – Reserved, 51234 – Catering Counterhand, 51312 – Health Assistant, 74323 – Canvas Worker, 82631 – Sewing Machinist, 91111 – Cleaner. Thirty-eight percent of the occupations contained only one employee, although two occupations contained eight employees each.

There were 53 occupations that contained no women, which related to 186 men in the dataset. These observations were removed from the analysis dataset, and the affected occupations were: 12214 – Transport Manager, 12215 – Forest Manager, 12267 – Other Catering Services Manager, 21141 – Geologist, 21455 – Other Mechanical Engineer, 21461 – Chemical Engineer, 21483 – Cartographer and Photogrammetrist, 22116 – Forestry Scientist, 22132 – Agricultural Consultant, 24511 – Minister of Religion, 31141 – Telecommunications Technician, 31422 – Launch Master, 31423 – Other Ships’ Deck Officer/Pilot, 32261 – Other Health Associate Professional, 33131 – Real Estate Agent/Property Consultant, 33171 – Valuer, 33642 – Instrumentalist, 33692 – Sports Coach/Trainer, 41212 – Audit Clerk, 51221 – Chef, 51235 – Kitchenhand, 52111 – Sales Assistant, 61133 – Grounds/Green Keeper, 61134 – Gardener, 61211 – Dairy Farmer/Farm Worker, 61214 – Pig Farmer/Farm Worker, 61221 – Mixed Livestock Farmer/Farm Worker, 61241 – Apiarist/Apiary Worker, 61251 – Crop and Livestock Farmer/Farm Worker, 61312 – Forest Hand, 71121 – Carpenter/Joiner, 71231 – Plumber, 71241 – Painter/Decorator/Paperhanger, 71311 – Electrician, 72122 – Sheet-Metal Worker, 72231 – Fitter and Turner, 72311 – Machinery Mechanic, 74321 – Furniture Upholsterer, 81231 – Welder/Flame-Cutter, 81411 – Timber Processing Machine Operator, 81612 – Boiler Attendant, 82641 – Launderer, 82643 – Drycleaner, 82741 – Baked Goods/Cereals Producing Machine Operator, 82751 – Fruit/Vegetable/Nut Processing Machine Operator, 82773 – Other Food Products Processing Machine Operator, 82931 – Metal Goods Assembler, 82932 – Plastic/Rubber Goods Assembler, 82941 – Wood and Related Materials Products Assembler, 83311 – Farm Machinery Operator, 91211 – Courier/Deliverer, 91512 – Builder’s Labourer, 91514 – General Labourer. While one occupation contained 20 employees, the majority of occupations contained less than four.

## 2001 Dataset

There were 14 occupations that contained no men, which related to 23 women in the dataset. These observations were removed from the analysis dataset, and the affected occupations were: 12211 – Senior Education Manager, 21142 – Geophysicist, 21412 – Resource Management Planner, 22121 – Biochemist, 22212 – Resident Medical Officer, 22251 – Dietitian/Public Health Nutritionist, 22315 – Public Health and District Nurse, 33614 – Sub-Editor, 41131 – Reserved, 41212 – Audit Clerk, 51234 – Catering Counter Assistant, 74323 – Canvas Worker, 82631 – Sewing Machinist, and 91111 – Cleaner. Sixty-four percent of the occupations contained only one employee, although one occupation contained five employees.

There were 58 occupations that contained no women, which related to 224 men in the dataset. These observations were removed from the analysis dataset, and the affected occupations were: 11411 – Special-Interest Organisation Administrator, 12214 – Transport Manager, 12215 – Forest Manager, 12261 – Supply and Distribution Manager, 12267 – Other Catering Services Manager, 21111 – Physicist, 21141 – Geologist, 21455 – Other Mechanical Engineer, 21461 – Chemical Engineer, 21481 – Surveyor, 21483 – Cartographer and Photogrammetrist, 22116 – Forestry Scientist, 22132 – Agricultural Consultant, 22216 – Radiologist/Radiation Oncologist, 22316 – Occupational Health Nurse, 24511 – Minister of Religion, 31141 – Telecommunications Technician, 31422 – Launch Master, 31423 – Other Ships’ Deck Officer/Pilot, 33171 – Valuer, 33242 – Building Control/Consents Officer, 33421 – Employment Programme Teaching Associate Professional, 33692 – Sports Coach/Trainer, 41331 – Transport Clerk, 41423 – Postal Deliverer, 51221 – Chef, 61133 – Grounds/Green Keeper, 61134 – Gardener, 61211 – Dairy Farmer/Farm Worker, 61214 – Pig Farmer/Farm Worker, 61221 – Mixed Livestock Farmer/Farm Worker, 61241 – Apiarist/Apiary Worker, 61251 – Crop and Livestock Farmer/Farm Worker, 61312 – Forest Hand, 71121 – Carpenter/Joiner, 71231 – Plumber, 71241 – Painter/Decorator/Paperhanger, 71311 – Electrician, 72122 – Sheet-Metal Worker, 72231 – Fitter and Turner, 72311 – Machinery Mechanic, 74321 – Furniture Upholsterer, 81231 – Welder/Flame-Cutter, 81411 – Timber Processing Machine Operator, 81612 – Boiler Attendant, 82121 – Concrete Worker, 82413 – Joiner’s Benchhand, 82641 – Launderer, 82643 – Drycleaner, 82741 – Baked Goods/Cereals Producing Machine Operator, 82751 – Fruit/Vegetable/Nut Processing Machine Operator, 82773 – Other Food Products Processing Machine Operator, 82931 – Metal Goods Assembler, 82932 – Plastic/Rubber Goods Assembler, 82941 – Wood and Related Materials Products Assembler, 91211 – Courier/Deliverer, 91512 – Builder’s Labourer, 91514 – General Labourer. While one occupation contained 22 employees, the majority of occupations contained less than four.

## Appendix C: Creation of Age Variables for the Model

These analyses were performed using the 2000 dataset, and were assumed to hold for the 2001 dataset.

### C1 Linear Effect of Age

Initially a simple linear regression of age against  $\ln\_wages$  was fitted separately for males (M) and females (F). The model summaries are shown in Tables CM1 and CF1 below. As the tables show, a model incorporating only a linear form of age accounts for 7.4% and 5% of the variance in  $\ln\_wages$ , for males and females respectively.

**Table CM1. Model summary of simple linear regression of age against  $\ln\_wages$ , males**

Model Summary <sup>b</sup>				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.273 <sup>a</sup>	.075	.074	.3287

a. Predictors: (Constant), Age as at 30 June 2000  
b. Dependent Variable: LN\_WAGE

**Table CF1. Model summary of simple linear regression of age against  $\ln\_wages$ , females**

Model Summary <sup>b</sup>				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.224 <sup>a</sup>	.050	.050	.2787

a. Predictors: (Constant), Age as at 30 June 2000  
b. Dependent Variable: LN\_WAGE

Tables CM2/CF2 and CM3/CF3 below show the ANOVA result and the coefficients of the simple linear regression, respectively for males and females. While the linear form of age explains only a small proportion of the variance in  $\ln\_wages$ , it is a significant effect of age.

**Table CM2. ANOVA of simple linear regression of age against  $\ln\_wages$ , males**

ANOVA <sup>b</sup>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	93.364	1	93.364	864.104	.000 <sup>a</sup>
	Residual	1158.482	10722	.108		
	Total	1251.846	10723			

a. Predictors: (Constant), Age as at 30 June 2000  
b. Dependent Variable: LN\_WAGE

**Table CF2. ANOVA of simple linear regression of age against ln\_wages, females**

ANOVA <sup>b</sup>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	50.827	1	50.827	654.218	.000 <sup>a</sup>
	Residual	957.850	12329	7.769E-02		
	Total	1008.677	12330			

a. Predictors: (Constant), Age as at 30 June 2000  
 b. Dependent Variable: LN\_WAGE

**Table CM3. Coefficients of simple linear regression of age against ln\_wages, males**

Coefficients <sup>a</sup>						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.650	.014		190.022	.000
	Age as at 30 June 2000	9.363E-03	.000	.273	29.396	.000

a. Dependent Variable: LN\_WAGE

**Table CF3. Coefficients of simple linear regression of age against ln\_wages, females**

Coefficients <sup>a</sup>						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.645	.010		273.747	.000
	Age as at 30 June 2000	6.024E-03	.000	.224	25.578	.000

a. Dependent Variable: LN\_WAGE

Normal probability plots and plots of studentised residuals against standardised predicted values (figures not shown) clearly indicated that a higher order form of age was required in both the male and female regression models.

## C2 Addition of a Quadratic Age Variable

A multiple linear regression was performed for each gender, with both a linear and a quadratic form of age as the predictor variables. Tables CM4 and CF4 below show the model summary for each multiple regression. The adjusted R<sup>2</sup> has been improved from 7.4% to 11.1% for males and from 5% to 9.9% for females. The standard error of the estimate has been reduced from 0.3287 to 0.3221 for males and from 0.2787 to 0.2715 for females.

**Table CM4. Model summary of multiple regression of age against ln\_wages, males**

Model Summary <sup>b</sup>				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.334 <sup>a</sup>	.112	.111	.3221

a. Predictors: (Constant), AGE2, Age as at 30 June 2000  
 b. Dependent Variable: LN\_WAGE

**Table CF4. Model summary of multiple regression of age against ln\_wages, females**

Model Summary <sup>b</sup>				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.315 <sup>a</sup>	.099	.099	.2715

a. Predictors: (Constant), AGE2, Age as at 30 June 2000  
 b. Dependent Variable: LN\_WAGE

For both the male and female multiple regressions, the regression model and the coefficients were significant ( $p < .001$  for all).

Normal probability plots and plots of studentised residuals against standardised predicted values (figures not shown) clearly indicated an improved model fit over the simple linear regression. However, heteroscedasticity was now evident in both the male and female models.

### C3 Addition of a Cubic Age Variable

Finally, a multiple linear regression was performed for each gender, with a linear, a quadratic, and a cubic form of age as the predictor variables. The addition of a higher order polynomial did not improve the model. Little improvement was noted for the model, with no improvement in the adjusted  $R^2$  for males (11.1%) and only a slight improvement for females (from 9.9% to 10.4%). The standard error of the estimate was not reduced for males (remained at 0.3221), and only slightly improved for females (from 0.2715 to 0.2707). There was no improvement noted for the residual diagnostics. Both the normal probability plots and the plots of studentised residual against standardised predicted values (not shown) were similar to those produced for the model that incorporated linear and quadratic forms of age, including the heteroscedasticity problem.

### C4 Recommendation

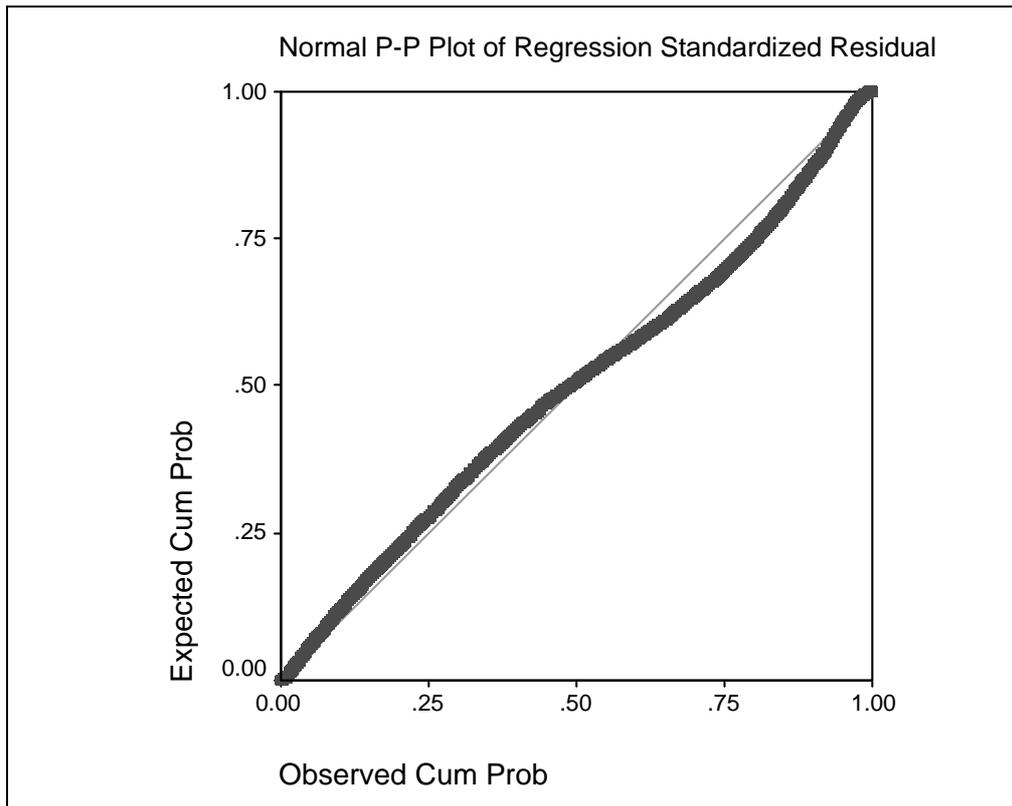
As the increased complexity of the effect of age on ln\_wages through adding a cubic variable is not countered by improved residuals, the model includes only the linear and quadratic forms of age.

## Appendix D: Regression Diagnostics Output

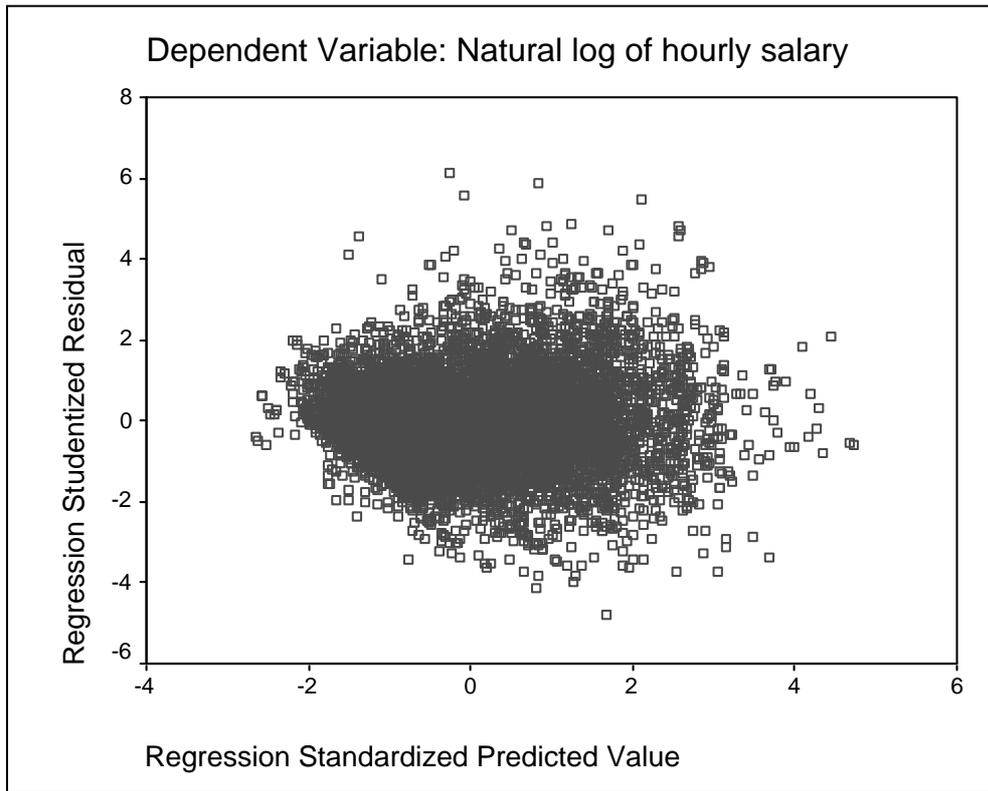
Normal probability plots and residual plots were requested for all four regression models examined in this study. These plots were used to check for any violation of the regression assumptions of normal distribution of regression residuals, linear relationship between  $\ln\_wages$  and the predictor variables, and constant variance of  $\ln\_wages$ .

**2000**

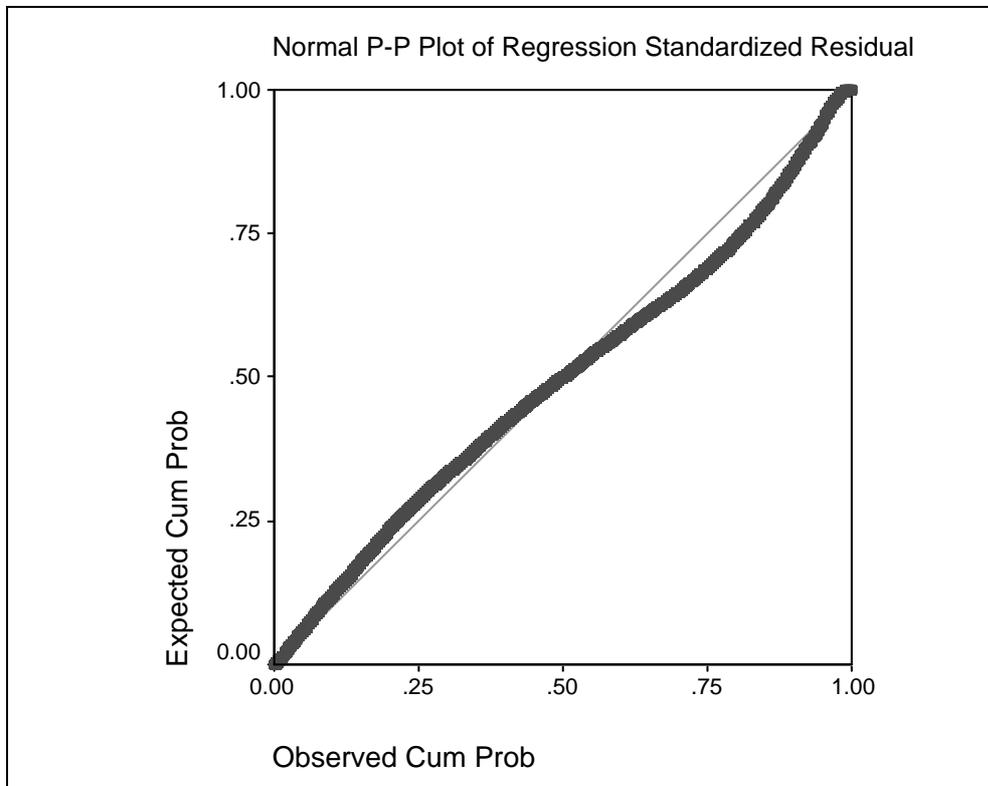
### *Male Normal Probability Plot*



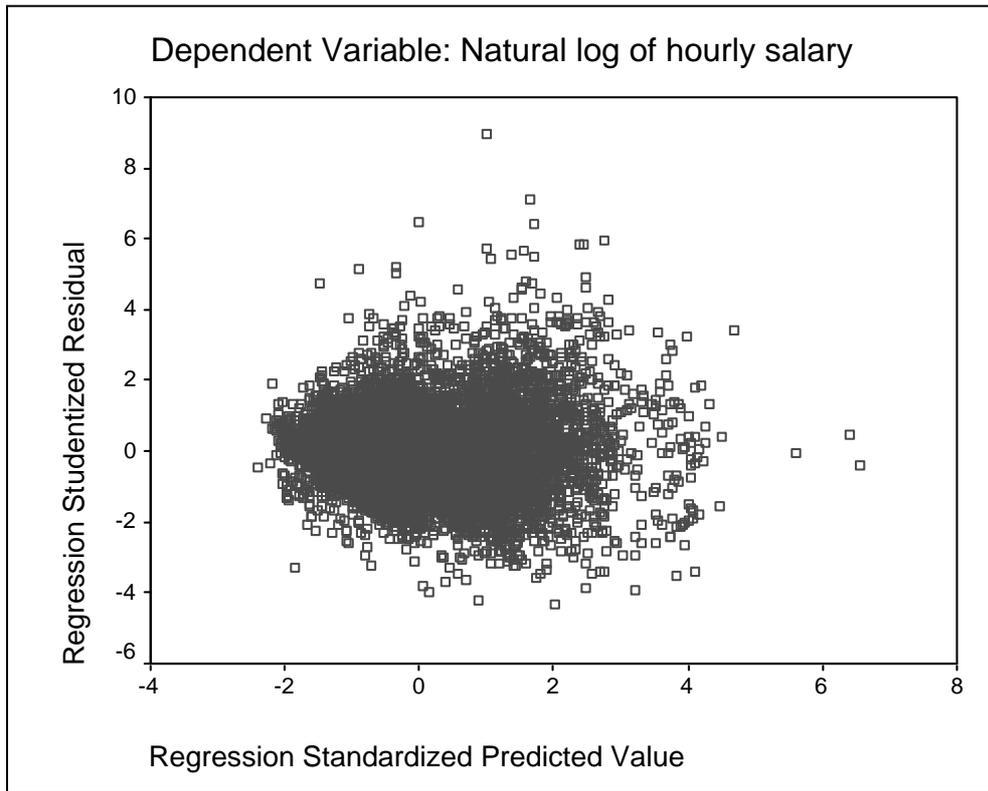
**Male Residuals Plot**



**Female Normal Probability Plot**

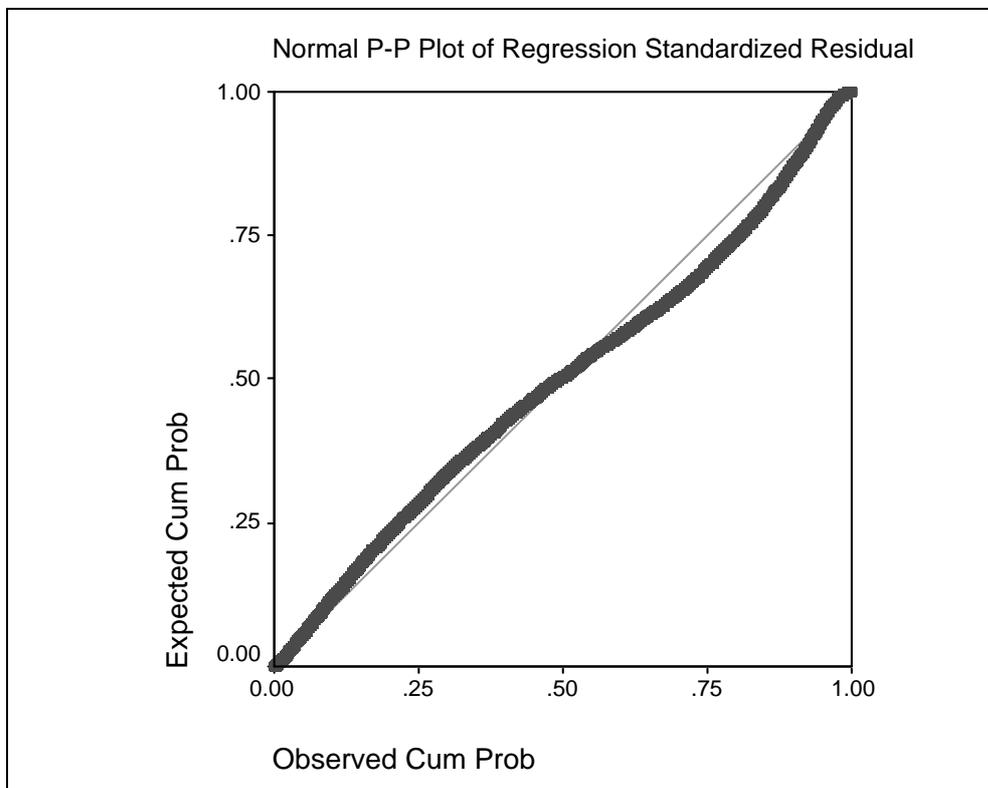


**Female Residuals Plot**

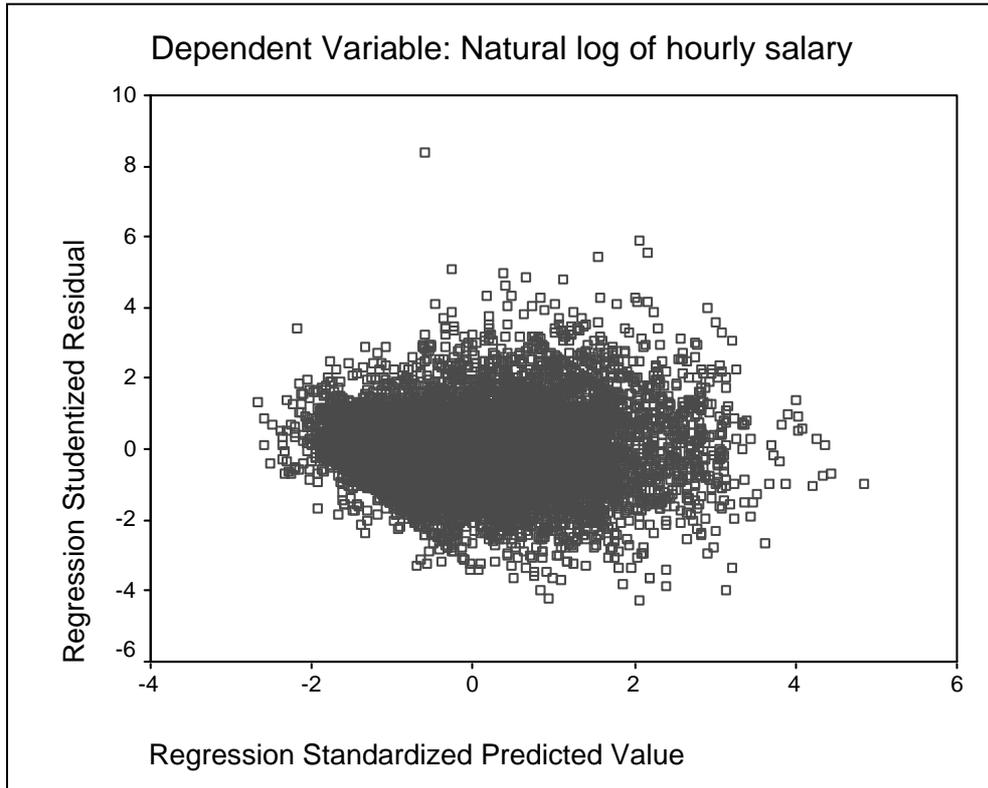


**2001**

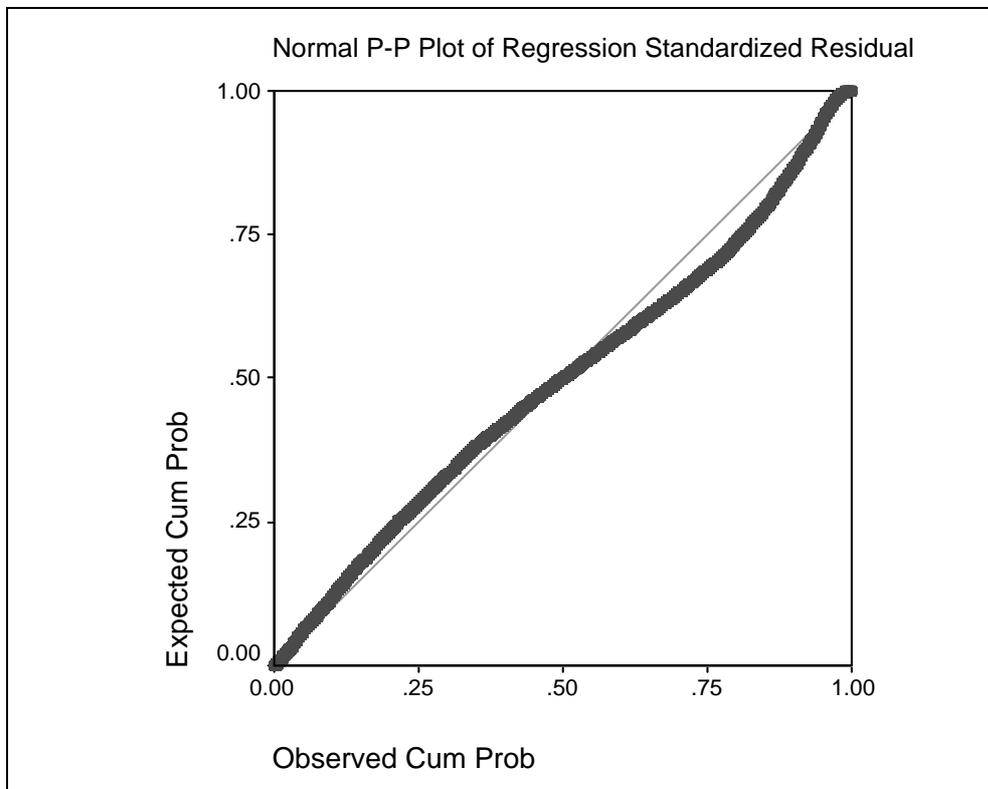
**Male Normal Probability Plot**



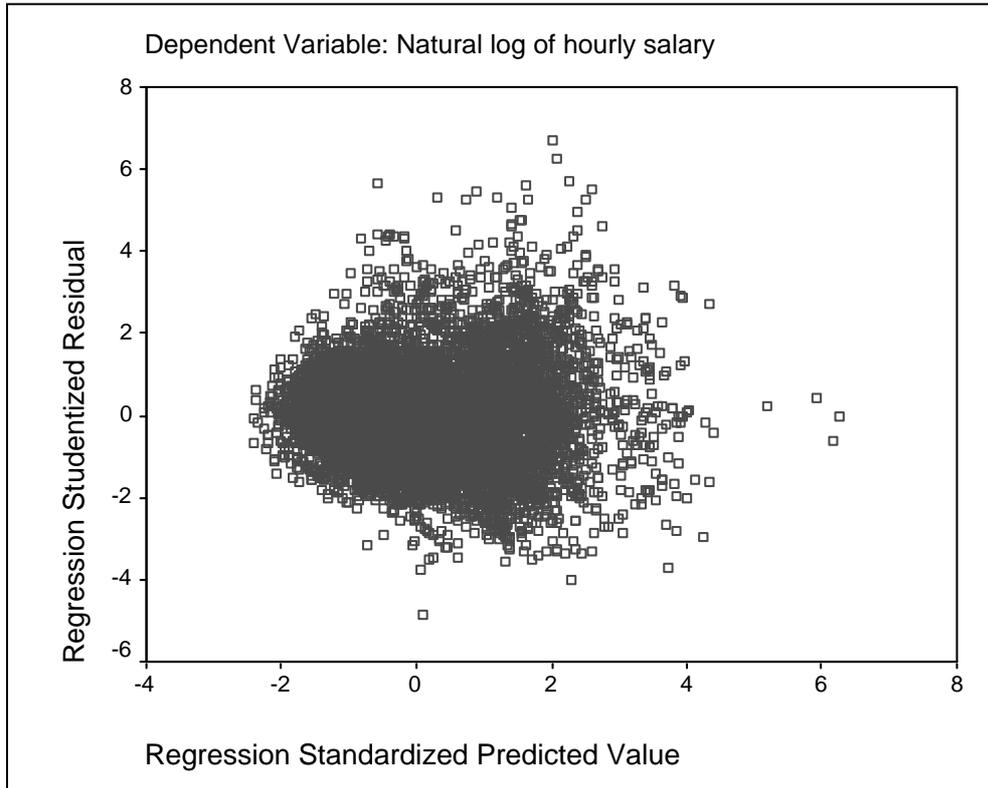
**Male Residuals Plot**



**Female Normal Probability Plot**



*Female Residuals Plot*



## Appendix E: Regression Diagnostics with Hourly Salary

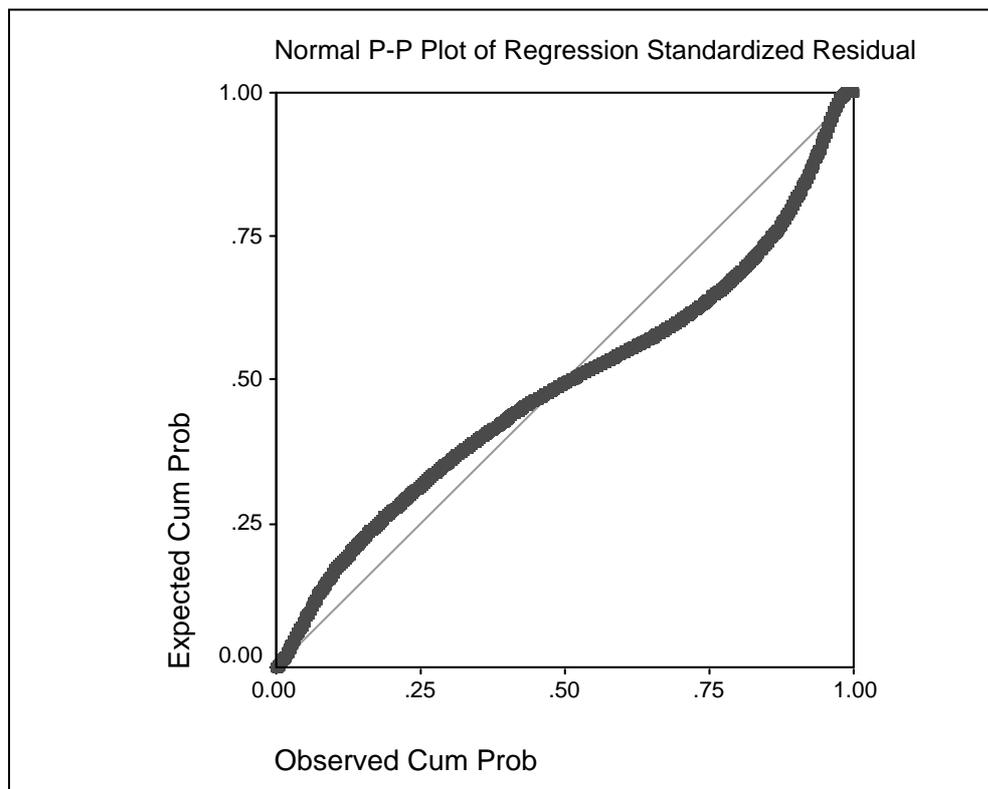
The 2000/male model was run using hourly salary as the dependent variable (“untransformed model”). The objective of this analysis was to establish the appropriateness of using the natural log of hourly salary, rather than simply hourly salary, as the dependent variable.

### E1 Findings

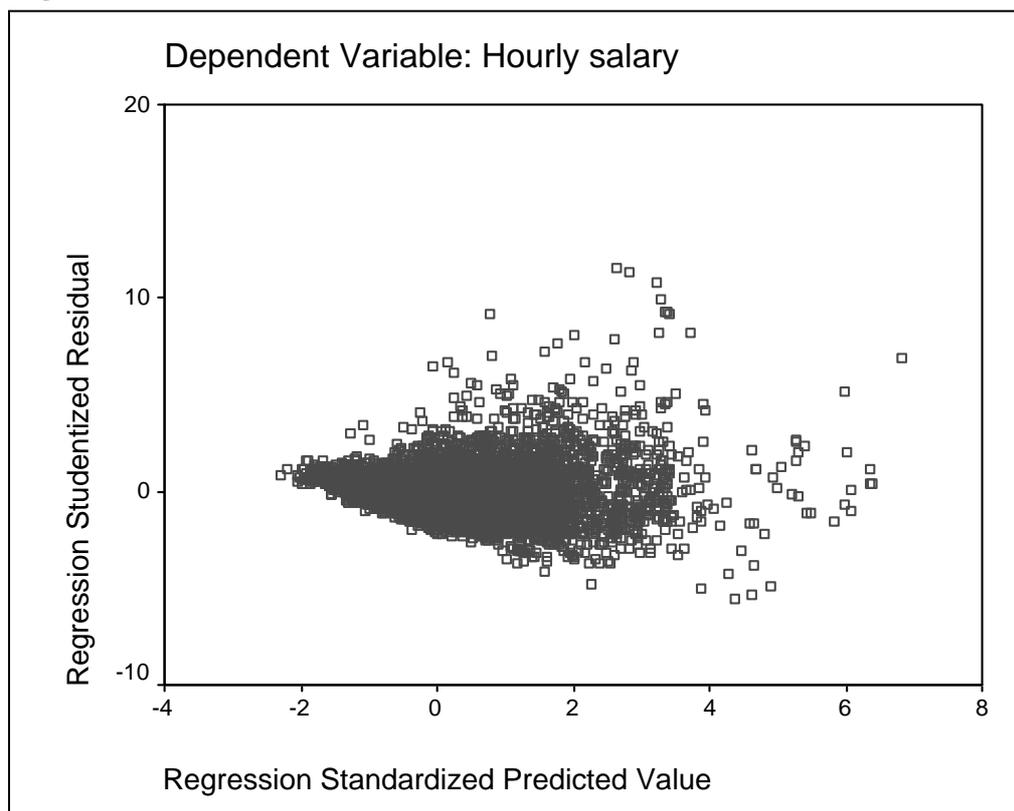
The adjusted  $R^2$  of the untransformed model was 66.8%, compared to 71.1% in the reported model. This result indicates that the natural log transformation of the dependent variable is associated with an increase in explanatory power of the model. The intercept term is negative (-7.497) in the untransformed model, compared to a positive intercept (1.75895) in the reported model.

The untransformed model meets fewer regression assumptions compared to the reported model. The normal probability plot (Figure E1 below) of the standardised residuals suggests that hourly salary does not meet the assumption of coming from a normal distribution. The scatter plot of studentised residuals versus standardised predicted values (Figure E2) suggests that the assumption of a linear relationship between hourly wages and the matrix of predictor variables, and the assumption that hourly wages has a constant variance, were not met. Heteroscedasticity is evident in Figure E2, with the residuals fanning out as the standardised, predicted value of hourly salary increases.

**Figure E1. Normal Probability Plot of the Untransformed Model**



**Figure E2. Residuals Plot of the Untransformed Model**



The one-sample Kolmogorov-Smirnov test for goodness-of-fit produced a test statistic of 11.852 ( $p < .001$ ). This test statistic is over twice the size of that produced for the reported model (5.577,  $p < .001$ ) and strongly suggests that the residuals from the untransformed model have an even worse fit with the normal distribution. The largest Cook's D statistic was 0.29 compared to only 0.24 for the reported model, and the Durbin-Watson statistic was 2.0 (the same as for the reported model).

## **E2 Recommendations**

The regression diagnostics of the alternative untransformed model, using hourly salary as the dependent variable, show noticeable deviation from the regression assumptions of a normal distribution of hourly salary, of linearity of hourly salary and the predictor variables, and of constant variance of hourly salary. There was much less deviation using the transformed predictor variable. Accordingly, the natural log of hourly salary has been used in both decompositions.

<sup>1</sup> This is an illustrative rather than an exhaustive list of protected minorities/groups.

<sup>2</sup> The Act also prohibits discrimination on these grounds in the areas of: education and vocational training; access to places, vehicles and facilities; provision of goods and services; and provision of land, housing and other accommodation.

<sup>3</sup> Defined as at least 35 hours per week, whereas most other jurisdictions define full-time employment as 30 hours or more.

<sup>4</sup> This Act recognised the rights of unions to wage bargain and to hold assets, and was heavily based in English law.

<sup>5</sup> A two-stage process where a national occupational award for wages and conditions was set and then "above award" supplementary conditions were negotiated in larger organisations.

<sup>6</sup> There are eight occupation categories, excluding "Not specified". These categories are: "Associate Professionals", "Customer Services Clerks", "Managers", "Office Clerks", "Personal and Protective Services

---

Workers”, “Professionals”, “Science/Technical”, and “Trades and Production Workers”. While these categories are constructed from the 2-digit NZSCO codes, they are not directly comparable to the 1-digit NZSCO classifications for all categories.

<sup>7</sup> Defined as the Mincer proxy for experience.

<sup>8</sup> Pay directly contingent upon job performance, e.g., bonus, commission, and profit sharing.

<sup>9</sup> At least 60% of the full-time employees for that occupation are male.

<sup>10</sup> Four models were run in a 2x2 design (two experience calculations x two coefficient weighting methods). The range represents the minimum and maximum values from the four models.

<sup>11</sup> The core Government departments defined in Schedule 1 of the State Sector Act 1988. The June 2000 dataset contains information on 37 out of 38 departments as one department did not provide information. The June 2001 dataset contains information on 39 departments as Archives New Zealand was separated from the Department of Internal Affairs during this year.

<sup>12</sup> These employees have no ongoing expectation of employment.

<sup>13</sup> Where a firm is engaged rather than a person. Employees on fixed-term agreements, who may be regarded in some sense as “contractors”, are included if they are paid wages or salary from the department’s payroll. Typically, the firm of a contractor is paid from the Finance/Accounting section of departments and is not entered on payroll.

<sup>14</sup> Employees who terminated their employment between 1 July and 30 June of the analysis year, and employees on secondment, leave without pay, or parental leave as at 30 June. It is assumed that the salary information provided for current employees reflects the actual salary as at 30 June, whereas the salary information for non-current employees is influenced by additional factors (e.g., no incremental pay increase in the current year).

<sup>15</sup> Due to the small number of part-time employees, these people were excluded from the analysis rather than incorporating an indicator variable indicating full-time status into the model.

<sup>16</sup> These are occupations that comprise solely women or solely men. Refer Appendix A.

<sup>17</sup> The HRC dataset contains up to two ethnicities per employee. The standard ethnicity prioritisation was used so that the ethnicity indicator variables in the model consisted of mutually exclusive groups. The priority system, in descending priority order, is: NZ Māori, Pacific peoples, Asian, Non-NZ European, Other, NZ European. The only alteration that has been made to the normal use of the system is that a NZ European/Other European combination has been coded to NZ European instead of Non-NZ European.

<sup>18</sup> A reviewer suggested that occupation and earnings are simultaneously determined within the labour market. In the case of simultaneous determination (i.e., occupation is endogenous to wage determination), occupation should be excluded from the decompositions. I decomposed the 2000 dataset, excluding the occupation indicator variables, to test this hypothesis. The reduced versions of the male and female models produced lower adjusted  $R^2$  values (52.3% and 37.9% respectively, compared to 71.1% and 65.7% in the full models), and the explained gender pay gap was reduced to 50.9% compared to 71.4% in the full decomposition. These results are the basis for retaining occupation in the decompositions. A discussion of indirect versus direct discrimination in occupation is beyond the scope of this study.

<sup>19</sup> A fixed-term contract has a specified end date to the employment, whereas an open-term contract does not.

<sup>20</sup> An individual agreement is made between an employee and the employer. A collective contract has multiple employees and one employer as parties.

<sup>21</sup> In SPSS version 10, the Lilliefors significance for testing normality is used automatically. The reason for this correction is that the mean and standard deviation of the hypothesised normal distribution have been estimated from the sample data.

<sup>22</sup> While ethnicity is arguably a human capital variable, it was represented in the decompositions by five indicator variables so is shown separately in the table.

<sup>23</sup> All four employment agreement coefficients are positive and statistically significant at  $p < .001$ .

<sup>24</sup> The NZSCO categories do not incorporate measures of seniority. For example, a policy analyst and a senior policy analyst will both be coded in the same 5-digit occupation.

<sup>25</sup> The use of  $j-1$  indicator variable categories is due to the requirements of the linear regression model that presumes the absence of perfect collinearity among the predictor variables (Hardy, 1993). This means that none of the predictor variables can be expressed as a perfect linear combination of the other predictor variables.

<sup>26</sup> Categorical level variables such as part-time work status are included as a series of indicator variables in a regression model. The regression model includes  $k-1$  indicator variables, as inclusion of all the categories would be overspecification in the model. The omitted category is sometimes referred to as the “reference” category for that variable. In the example of part-time work status, one variable taking a value of “1” for part-time work and “0” for full-time work would be included in the model examining the data of employees.

<sup>27</sup> To illustrate using age and an indicator variable for part-time employment, the model would contain a mean for age ( $X_1$ ) and a mean for part-time employment ( $X_2$ ).  $X_1$  would be the mean age of the sample, and  $X_2$  – because this variable is an indicator – would be the proportion of the sample in part-time employment. Thus, the “mean” for any indicator variable must contain a value between 0 and 1.

- 
- <sup>28</sup> There was a 1972 Federal decision of “equal pay for work of equal value” but most of the translation of this decision into employee wages occurred post-1973. Thus, the expectation was that the pay gap should have decreased between the two periods due to the implementation of this interpretation of equal pay after 1973.
- <sup>29</sup> The differences in hourly wages based on the levels of each variable (e.g., the number of years of work experience) are taken into account by the means of the variables in the model. Therefore, the coefficients are assumed to show pure discrimination.
- <sup>30</sup> Experience is typically defined as [age – (years of education + 6)]. The value of six accounts for the pre-school years, and can be altered to suit the school age of the country where the data analysis is being performed. This definition of experience is from Mincer.
- <sup>31</sup> These four groups are composed from the  $2^2$  combinations of the indicator variables. The fourth group – represented by the values 0,0 – is included in the model. With three indicator variables there would be  $2^3$  combinations, and so forth.
- <sup>32</sup> I suggest that the reason for this is that with the proxy variable the unexplained residual is capturing measurement error associated with that proxy. In this sense, the unexplained residual is at least partially acting as an error term in the model. However, a reviewer has suggested that a proxy measure could reduce the unexplained residual if the “real” variable is unavailable to employers and the proxy is a reliable (and valid) substitute.