Exploring the Gap: An Exploration of the Difference in Income Received from Wages and Salaries by Women and Men in Full-Time Employment

Heather Kirkwood

Statistics New Zealand

Abstract

The difference between the income received by employed women and men in New Zealand is widely recognised. Often referred to as the "gender wage gap"[1], this income differential has prevailed despite significant improvements in the past 50 years including changes in society's attitudes towards women in employment and increasing numbers of women in paid work. What then can this difference be attributed to? Is it structural factors such as age and occupation, or are there other un-measurable factors at work?

Confining itself to the wages and salaries of women and men in full-time employment, this paper presents some of the results of recent analysis conducted by Statistics New Zealand's Analytical Support Division. Using data from the 1997 New Zealand Income Survey and several different multivariate techniques, this analysis attempts to explore some of the factors behind this "earnings gap".

Originally conducted for an upcoming Statistics New Zealand analytical publication, this research first explores the factors responsible for prescribing the level of earnings for all people then applies these results in a process of standardisation. The results leave much scope for both discussion and a continuation of analysis.

Keywords: *gender, earnings, New Zealand Income Survey*

Access to an adequate level of income enables women and men to buy the goods and services they require to maintain an acceptable standard of living for themselves and their family. Although access to income is not confined to that received by an individual (wealth, family and household income need also to be considered), personal income is an important measure of economic independence, enabling comparison between different groups in the population. Historically, no analysis of women's personal income has been complete without mention of the higher income received by men. Results from the 1996 Census of Population and Dwellings highlight this discrepancy with the median personal income of women, at $12,600, equal to 57 percent of that of men ($22,000). However, when the longer life expectancy of women is considered (meaning more years at lower incomes), along with the impact of child-bearing and child-rearing responsibilities on labour force participation (including reduced hours or withdrawal from the labour force), eliciting meaning from such a statistic is difficult. What is of more significance is a comparison of income received by men and women in similar circumstances.

In June 1997 Statistics New Zealand carried out the first New Zealand Income Survey. The survey collects recent gross income data including that from wages and salaries, self-employment, government transfers and private pensions, for people aged 15 years or over. The collection excludes income from investments or interest. Conducted as a supplement to the Household Labour Force Survey, the New Zealand Income Survey can be directly related to respondents' labour force characteristics. It is planned to repeat the survey on an annual basis.

Issues relating to income and income distribution are ones often examined by Statistics New Zealand through its series of census analytical publications. Entitled *New Zealand Now,* Statistics New Zealand aims in these publications to provide readers with an awareness of some of the issues and trends current in New Zealand today. Targeted at a reading age of approximately 15, the reports generally centre around information available from current and past censuses, although other data sources are drawn upon where appropriate. In the 1998 edition of *Women,* an attempt was made to explore some of the factors responsible for the observed difference in earnings of women and men in full-time employment. This paper presents some of the background to the research, a summary of methodologies used and discusses some of the results of that analysis.

Background

According to the New Zealand Income Survey, in June 1997 the gross average weekly income of women (aged 15 years

and over) in New Zealand was $311 - equal to 60 percent of that of men ($516). Meaningful conclusions cannot be drawn from this result as women are less likely than men to be employed. In June 1997, 54 percent of women were employed compared with 69 percent of men. Women are also likely to spend longer outside the workforce (ie. unemployed or not in the labour force) than men. Although lower participation in the paid workforce is likely to be a reflection of life-cycle factors such as child-bearing and child-rearing, women continue to be less likely than men to be employed at age 65 years and over. This may be attributable to two main factors: firstly, societal aspects meaning women in this cohort were less likely to ever have been employed and, secondly, women have an older age structure than men even within this older age group. The life expectancy of women aged 65 years in 1995/97 was a further 19.2 years compared with 15.6 years for men of the same age.

Equally, it is difficult to gain conclusions from a comparison of the income of all people in paid employment. Within this group, women have an average weekly income equal to 67 percent of that received by men ($448 compared with $671). Much of this discrepancy can be attributed to the higher proportions of women in part-time work. A total of 36 percent of all women in paid employment work less than 30 hours per week compared with 13 percent of employed men.

As such, to meaningfully explore the disparity between the incomes of women and men it is necessary to restrict the analysis to a smaller group of the population assumed to have similar characteristics. This current analysis assumes a proxy for this to be full-time employment. The scope has been further refined to exclude people with income from self-employment. The reasons for this are two-fold. Firstly, the factors responsible for determining income from self-employment are assumed to be vastly different to those associated with wages and salaries. Secondly, there are questions surrounding the quality of self-employment data. There is a relatively low response rate associated with people in self-employment with many either not sure of their income or supplying information that is difficult to reconcile. With the exclusion of both self-employment income and that from interest and investments, the term "income" seems no longer appropriate for this analysis. For the purposes of this report, income from wages and salaries will hereafter be referred to as "earnings" and the issue under analysis referred to as the earnings gap between men and women in full-time employment. The following analysis is based on average weekly earnings rather than hourly rates. There is no methodological reason for this decision, but rather an awareness of:

1. The audience of the original report for which this was intended. With the large number of women now in salaried work, it was expected that weekly earnings would be more meaningful than hourly rates.

2. The desire to clearly highlight the impact of number of hours worked and the likelihood of women working fewer paid hours than men, even amongst those in full-time employment. Comparisons based upon the hourly rate under-state the actual gap in the take-home earnings of women and men.

According to the New Zealand Income Survey, in June 1997 the average weekly earnings of men in full-time paid employment was $698. The comparable earnings for women was $554 - equivalent to 79 percent of that of men. This 21 percent gap (or $144 in June 1997) will be explored in the following work.

The data

Data used in this analysis are from the June 1997 New Zealand Income Survey and corresponding quarter of the Household Labour Force Survey (HLFS). There are a number of technical issues associated with the use of this survey information. They include the following:

Population

Only fully-responding HLFS respondents were asked to complete the 1997 New Zealand Income Survey (see Appendix 1). The original dataset had 28,292 observations. Restricting the analysis to people in full-time employment, excluding the self-employed, reduced the final dataset to 9,985 observations. Of these, 80 were proxy responses completed on behalf of people who were unable to answer the survey themselves for health or language reasons. Although proxy responses were not expected to adversely affect the analysis, the implications associated with including these responses was tested.

Imputation

In the HLFS, imputation is carried out for age and sex. Missing values for all other variables are left as "missing". The 1997 New Zealand Income Survey had a response rate of 80 percent of HLFS respondents. For the remaining 20 percent, the values from the HLFS were used but the income information was imputed using a hot deck method (whereby a record was replaced with another with similar HLFS variables from a donor pool of records). Leaving imputed records in the dataset means that some people from the donor pool are included twice. If omitted, however, there would be a significant impact upon the representativeness of the sample weighting factors which have been designed for the total population. In particular, there is a higher level of imputation amongst those in full-time employment. As such, imputed records have been included in this analysis.

Response

Valid responses to the 1997 New Zealand Income Survey, calculated as a percentage of those who responded to the HLFS (un-weighted), are presented below:

Sex	%
Male	77
Female	83

Ethnicity	
European/Pakeha	80
NZ Maori	78
Pacific Island	79
Other	82
Labour Force Status	
Full-time employed	77
Part-time employed	81
Not in the labour force	83
Unemployed	87
Total	80

Weights

The 1997 New Zealand Income Survey is weighted up to the total usually resident population. A post-stratification adjustment is applied to the survey weights to ensure the final estimates for sex and five year age groups match demographic estimates (based upon census totals). To ensure that a representative picture of New Zealand was achieved, it was decided to carry out this analysis on the weighted results. Anticipating that the inclusion of these weights may have a spurious effect on results (the exclusion of people not in full-time employment and those who are self-employed will impact upon the validity of the weights), weight was included and tested as a possible "explanatory" factor in this analysis.

Hours worked

The 1997 New Zealand Income Survey collected information from respondents on up to three paid jobs. In the current analysis, the usual hours in all jobs are taken for people with an HLFS labour force status of "employed". As such, overtime hours have been excluded, ensuring consistency with the dependent variable which is taken to be usual earnings per week. In a small number of cases, however, the usual hours for the reference week were not reported and so the actual hours worked were used instead.

Exploration

Multivariate analysis includes any one of several methods for examining more than two variables simultaneously. An important component of this kind of analysis is the ability to examine the relationship between two variables while controlling for how each of these may be influenced by other variables. Given a continuous dependent variable (earnings), standardisation seemed an appropriate method of correcting for other variables. However, to correct for all variables mentioned (including combinations of two or more variables) would have resulted in too many categories to analyse practically. Additionally, not all of the variables would prove to be relevant. Hence, to ensure the analysis was parsimonious, it was only necessary to analyse the factors of most importance in determining the earnings of all people in full-time employment. It was also important to correct for possible spurious correlations with other variables such as *qualifications, hours worked* etc. With this in mind, tree analyses seemed to be an ideal method to begin the exploration. This kind of analysis not only indicates relevant variables, but also lends itself to even more parsimonious analysis by selecting homogeneous subgroups defined by combinations of the relevant variables.

Methodology

Tree analyses

A review of previous research in this area led to the following variables being selected as possible explanatory factors:

Age	Household type
Sex	Family code
Ethnicity	Household income
Highest educational qualification	Age of youngest child
Hours worked	Region (Regional Council)
Occupation	Imputation
Industry	Proxy
Marital status	Weight

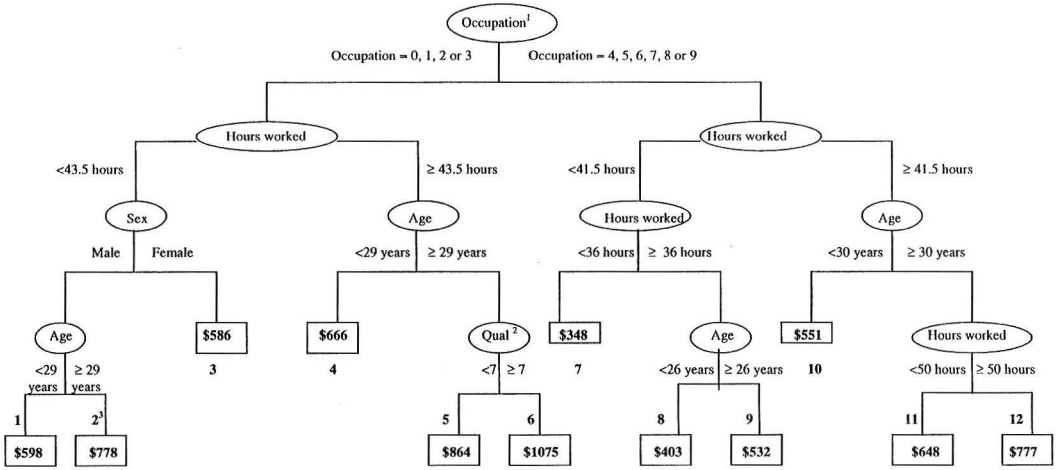
Acknowledging that the difference between the earnings received by women and men could be due to spurious (statistically related but not causally linked) relationships with other variables, attempts were made first to find a number of relatively homogeneous income sub-groups. By identifying these distinct sub-groups, each with their own unique and quantifiable characteristics, the main explanatory variables of different levels of average weekly earnings could be determined.

Using the 9,985 observations, a 'flat' file which was created as input into the computer package SPLUS - a statistical package capable of analysing and graphically presenting complex data. The classification and regression program TREE was then applied (MathSoft, 1993). The result is a visual depiction (a "classification tree") which identified 115 terminal nodes (or sub-groups), the attributes of which 'explained' 45 percent of the earnings of women and men in full-time employment. Each sub-group is relatively homogeneous in terms of the variables prescribed and the splits can be traced backwards to determine the variables of most significance in determining earnings. All splits lead back to the factor the analysis found to be the original determinant of earnings - occupation.

Figure 1 is an illustration of this and while truncated to improve readability, it still explains 29 percent of average earnings as defined by 12 terminal nodes (or sub-groups). As such, the remaining 103 terminal nodes can appear to each prescribe only a relatively small amount of average weekly earnings. The identification of these main terminal nodes allows for the identification of the most significant factors responsible for earnings.

As previously mentioned, the variables *weight, proxy* and *imputation* were included in the tree analysis as a check of their relationship with earnings. The *proxy* and *imputation* variables did not emerge as significant and while the *weight*

Figure 1. Tree analysis of main explanatory factors of average weekly earnings



Source: Statistics New Zealand, New Zealand Income Survey, 1997

[1] Occupation classification based upon NZSCO90: 0 = Armed forces, 1 = Legislators, administrators, managers, 2 = Professionals, 3 = Technicians and associate professionals, 4 = Clerks, 5 = Service and sales workers, 6 = Agriculture and fisheries workers, 7 = Trades workers, 8 = Plant and machinery operators, 9 = Elementary occupations.

[2] Qualification classified is based upon highest educational qualification: 1 = no qualification, 2 - 5 = school qualification, 6 - 8 = post-school qualification.

= explanatory factor = sub-group

[3] Numbers 1,2,3 etc. have been assigned to each sub-group. See Appendix 1 for details of sample size of each group.

variable did emerge, it was not in the main part of the tree (that displayed in Figure 1). The impact of this *weight* variable was tested further and found, overall, to be not statistically significant in this analysis. What did emerge from the results, and became the main variables of interest, were *occupation, hours worked, age, highest qualification* and, to a lesser extent, *sex*. *Sex* only appears in one subgroup of the tree, as shown in Figure 1, where it emerged as important amongst those employed in one of the four highest paid occupations, and who usually worked less than 44 hours per week.

None of the other variables, including *marital status, household type, family code* and *household income*, emerged as important in determining earnings. This is likely to be a reflection of the fact that these variables are more likely to have an impact in determining whether or not an individual is in full-time employment rather than their earnings within that employment. *Partner's income* may be of more significance. This variable, although explored, was not included here as it further narrowed the dataset to include only people who could be allocated to couples. However, analysis of this factor may prove useful.

It should be noted that often tree analysis is carried out for a random sub-sample and that the analysis is then continued by applying the resulting tree 'rules' on the rest of the sample. This is mainly done to obtain unbiased estimates of the so-called misclassification rate. We did not do this here because the misclassifications were not considered.

Standardisation

Using the results from the tree analyses to identify the most significant factors affecting earnings, the next step involved using these to standardise women's and men's earnings. In view of the continuous nature of the dependent variable (average weekly earnings), enabling easier application of the sampling weighting factors, analysing the data by standardisation made more sense than other multivariate techniques. Standardisation was restricted to the following variables:

Occupation
Hours worked
Age
Highest qualification
Ethnicity

This last variable, *ethnicity*, although not prominent in the tree analyses is an often considered factor in income analysis in New Zealand. As such, it was decided to include it to illustrate its effect. Standardisation was carried out for each factor individually; then each factor was, in turn, removed from the standardisation to indicate its impact on results. Finally, the average weekly earnings of women and men were standardised for the final sub-groups. Standardisation was based upon all data under consideration, ie. for both women and men, standardisation was based upon the sum of male and female data. An alternative method would have been to standardise women's earnings for the age, ethnic etc, structure of the male population rather than for the total population. This was briefly tested and the results gained were marginally different. However, this analysis attempts to explore the earnings gap by asking what the gap would be if women and men had the same structure, rather than "what if women were like men?".

Results

Overall, by standardising average weekly income for the 5 main variables of interest (*hours worked, age, ethnicity, qualification* and *occupation*), the earnings gap narrowed, from 21 percent to 17 percent. That is, the ratio of women's earnings to men's increased from 0.79 to 0.83 as women's earnings increased from $554 (non-standardised) to $572 per week - an increase of 3.2%. At the same time, men's average earnings fell from $698 to $686 (a 1.7 percent decline), as shown in Table1.

Extending beyond these five factors and standardising for all 12 final sub-groups in Figure 1 increased, slightly, the average weekly earnings of women (from $572 to $577). At the same time, the earnings of men declined from $686 to $681. The result of these two opposing shifts meant a further small "decrease" of the earnings gap from a ratio of 0.83 to 0.85. The difference in earnings between men and women was reduced to $104.

Standardisation for each variable individually

As shown in Table 2, *hours worked* had the biggest impact upon average weekly earnings. Standardising for *hours worked* increases women's income from $554 to $572 while decreasing men's from $698 to $684. This improves the position of women's earnings, relative to men's, from 0.79 to 0.84. While the observation that the more paid hours

Table 1. Standardisation for all factors

	Men	Women	Difference	Ratio
Not standardised	$698	$554	$144	0.79
Standardised for:				
All 5 variables	$686	$572	$114	0.83
Standardised for all 12 final sub-groups (Figure 1)	$681	$577	$104	0.85

Table 2. Standardisation for one factor at a time

	Men	Women	Difference	Ratio
Not standardised	$698	$554	$144	0.79
Standardised for:				
Hours worked (3 classes <37 hrs ,37-<40 hrs ,40+ hrs)	$691	$568	$123	0.82
,, (2 classes <40 hrs, 40+ hrs)	$690	$568	$122	0.82
,, (2 classes <42 hrs, 42+ hrs)	$684	$572	$112	0.84
Age (3 classes 15 - 25, 26 - 29, 30+ years)	$694	$556	$138	0.80
" (continuous)	$696	$557	$139	0.80
Highest educational qualification	$699	$553	$146	0.79
Ethnicity (2 classes)	$698	$554	$144	0.79
Occupation (2 classes, 0 - 3, 4+)	$703	$550	$153	0.78
" (10 classes, 1 digit)	$705	$539	$166	0.76

worked, the greater the earnings, seems intuitive, it is important to remember that this is usual hours worked and does not include overtime hours. Additionally, all people under consideration are already working at least 30 hours per week. As such, the results are not expected to be greatly skewed by people working longer hours and receiving higher earnings as a direct result of this. Rather, one interpretation of this distinction between "hours beyond 30 hours per week" and "overtime hours" could be that usual hours worked may be some kind of indicator of level of responsibility or seniority within an occupation. However this suggests some kind of relationship with a variable such as "*seniority*" which was not explicitly collected in the data and not available for interpretation here.

The effect of hours worked was looked at in detail, with several different classes or "splits" of hours tried in the standardisation process. The results show that the higher the split (in terms of number of hours), the smaller the earnings gap between men and women.

The next variable standardised for was *age*; first broken into 3 broad age bands identified by SPLUS (15 - 25, 26 - 29 and 30 years and over), then applied as a continuous single-year variable. The result was the same for both methods - if men and women in full-time employment had the same age structure, the average weekly earnings of women would increase to 80 percent of men's earnings. This seems to support known differences in the employment patterns of many women (such as interruptions to participation in employment brought about by child-bearing and child-rearing responsibilities).

Highest qualification emerged as relatively important in the tree analyses. When this variable was standardised for (grouping qualifications into two broad classes in the classification - see Appendix 2), the average weekly earnings of women improved to 80 percent of men's. It is not surprising that standardisation of this variable did not greatly enhance the relative position of women's earnings. Although pivotal in determining the earnings of the aggregate of people in full-time employment, women in full-time employment are now more likely to have an educational qualification than their male counterparts. However, women's superior educational attainment has been a relatively recent phenomenon and the slight improvement made by this standardisation is likely to be a result of the lower educational attainment of older people in full-time employment. *Ethnicity* appears to have no impact in standardisation. This is consistent with its non-emergence in the higher parts of the tree analyses.

An unexpected result from the standardisation for individual factors was that *occupation* did not appear to have any impact on reducing the earnings gap. In fact, it actually increased from 21 percent to 22 percent using 2 classes of occupation (0-3 and 4+) and to 24 percent (using all ten classes). Given that *occupation* emerged as the most important determinant of earnings for the total population, the result is somewhat surprising. However, it is consistent with findings from other research, including that conducted by the New Zealand Planning Council which concluded that "differences in occupational distribution, at least when measured at the minor group level, do not account for any of the remaining difference (in the income of men and women)"[2].

Some possible explanations for this are as follows:

1. Although *occupation* is the most important determinant of the earnings of all people in full-time employment, the difference between the occupational structures of women and men is, itself, not significant in explaining the earnings gap between men and women. Table 3 illustrates the differing occupational structure of men and women in full-time employment along with the average weekly earnings associated with each occupation at the time of the 1997 New Zealand Income Survey. The results demonstrate that while women are less likely than men to be employed in higher paid occupations, this is offset by their lower likelihood of being employed

in the lowest paid occupations such as elementary workers or plant and machinery operators and assemblers. Additionally, women are more likely than men to be working as professionals, or technicians and associate professionals - occupations with higher average weekly earnings than occupations they are less likely to be employed in (such as agriculture and fishery workers or trades workers). As a result, there is no great disparity between the distribution of men and women in higher and lower paid occupations, meaning that standardising for this has little or no effect.

2. Standardisation was carried out at both the 1-digit and 2-digit levels of the occupational classification, with the same results. It is possible that these levels are too broad and that standardising at the 5-digit level may improve the results.

3. Standardising for the different occupational structures of women and men (regardless of the level of the classification) cannot standardise for the fact that women may be employed at a lower level within occupations than their male counterparts.

that hours worked had the greatest impact for standardisation, it is possibly not surprising that altering the occupational structure without also changing hours worked failed to improve the earnings disparity. The following results from the 1996 Census highlight the fact that women in full-time employment tend to work fewer hours than their male counterparts, in all occupations. Overall, women were nearly four times more likely than men to work between 30 and 39 hours but only half as likely to work 50 or more hours per week (Table 4).

Returning to the analysis, and now progressively standardising for all 5 variables bar one, the results in Table 4 further highlight the importance of hours worked in determining the earnings of men and women in full-time employment. Standardising for *ethnicity, age, qualification* and *occupation* but not *hours worked* produced the lowest earnings ratio, of 0.80.

Time and resource constraints did not allow for standardisation for all 115 sub-groups found in the preliminary tree analysis. However, given that 29 percent of total earnings could be attributed to only 12 sub-groups, the impact of furthering the standardisation to include these remaining 103 sub-groups (bringing the amount explained up to 45 percent) is likely to be small.

Table 3. Average weekly earnings of full-time employed men and women by occupation (NZSCO90)

	Proportion in each occupation			(%)Average weekly earnings		
	Men	Women	Total	Men	Women	Total
Armed forces	0.2	0.0	0.1	$923	$642	$885
Legislators, administrators and managers	14.6	9.3	12.5	$973	$699	$891
Professionals	12.0	19.3	14.9	$884	$687	$782
Technicians and associate professionals	12.3	14.3	13.1	$775	$581	$690
Clerks	5.9	28.2	14.8	$634	$525	$551
Service and sales workers	9.0	14.3	11.2	$582	$440	$509
Agriculture and fishery workers	5.9	2.3	4.5	$511	$414	$491
Trades workers	16.5	2.0	10.7	$611	$467	$600
Plant and machinery operators and assemblers	14.9	5.3	11.0	$585	$419	$553
Elementary occupations	8.7	4.9	7.2	$516	$422	$490
Total	100	100	100	$698	$554	$640

Due to rounding, columns may not sum to stated totals.
Source: Statistics New Zealand, New Zealand Income Survey, 1997

For example, at the 5-digit level of the classification it was found that there is one code for "mathematician and/or statistician" but unless employed at management level, there is no distinction made between senior and less senior members of staff. The occupational classification gives little indication of status within a job description. The New Zealand Planning Council research referred to this as "ranking" within an occupation and considered that it is "probably an important reason for the earnings gap" (ibid, 1989).

4. Related to this, standardising for occupation alone does not reflect the fact that full-time employed women work less hours than men in every occupation (1-digit). Given

Conclusion

Around half of the earnings of both women and men in full-time employment can be explained by the variables selected for exploration. When the characteristics of 12 of the sub-groups identified are applied in a process of standardisation, the identification of these variables improves the average weekly earnings of women (relative to men)

Table 4. Hours worked (per week) by occupation (NZSC090) for men and women in full-time employment

Occupation	30 - 39		40 - 49		50 – 59		60 or more		Total	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
	(%)		(%)		(%)		(%)		(%)	
0 Armed forces	2.6	1.8	54.2	60.8	20.0	15.0	23.2	22.4	100	100
1 Legislators, administrators and managers	3.6	16.4	36.7	48.1	32.3	18.5	27.3	17.1	100	100
2 Professionals	9.6	21.3	48.1	47.9	25.7	19.2	16.7	11.7	100	100
3 Technicians and associate professionals	11.4	25.9	55.5	57.1	19.7	10.4	13.4	6.6	100	100
4 Clerks	18.1	38.3	61.6	53.9	13.0	4.7	7.3	3.1	100	100
5 Service and sales workers	9.8	32.6	58.9	53.5	17.7	7.4	13.6	6.4	100	100
6 Agriculture and fishery workers	6.2	20.9	32.5	39.6	21.4	15.0	39.8	24.5	100	100
7 Trades workers	4.3	18.6	67.5	67.2	17.9	8.3	10.3	5.9	100	100
8 Plant and machinery operators and assemblers	3.8	14.1	56.9	73.6	19.7	6.8	19.6	5.5	100	100
9 Elementary occupations	7.6	2 1.3	65.1	64.7	15.1	7.4	12.2	6.6	100	100
Total	7.1	26.8	52.7	53.3	21.3	10.9	18.9	8.9	100	100

Due to rounding, rows may not sum to stated totals.
Source: Statistics New Zealand, Census of Population and Dwellings, 1996

from 0.79 to 0.85, indicating that at least some of the earnings gap can be attributed to structural differences within the full-time workforce. The majority of this improvement occurred when standardised for *hours worked*. What may be of more interest here is not further quantification of the difference in hours worked between the sexes but rather attempts to explore the reasons *why* women work less hours than men in all occupations. The reasons are likely to be varied and reflect women's greater levels of responsibility for childcare, care for adults, and unpaid work in general, and differing societal roles and expectations. Information from the upcoming 1998/99 Time Use Survey as well as the 1998 Childcare Survey may shed more light on the matter. Alternatively, the information could be used to answer the equally valid question - "why do men work more hours than women?".

Overall, the results of this work indicate that personal characteristics such as hours worked, level of highest qualification and age are of some importance in determining earnings while household and family characteristics are not. Consistent with other research in this area, the role of occupation appears not to be a main factor. The findings indicate that there are indeed some structural factors associated with the earnings gap (particularly hours worked), but that the remaining 15 percent could not be attributable to the main variables under consideration here.

Future research

This work is just the beginning of an area open to further exploration. The analysis was greatly restricted by both time and resource constraints, and by the scope of the initial research proposal. Standardising for all 115 sub-groups found in the preliminary tree analysis would be advantageous, perhaps also extending the number of variables considered to evaluate such factors as partner's income or years in employment. Revisiting the results of the occupational

Table 5. Standardisation for five factors less one factor at a time

	Men	Women	Difference	Ratio
Not standardised	$698	$554	$144	0.79
Standardised for:				
4 variables, not occupation	$687	$574	$113	0.84
4 variables, not qualification	$688	$567	$121	0.82
4 variables, not ethnicity	$686	$571	$115	0.83
4 variables, not age	$690	$565	$125	0.82
4 variables, not hours worked	$694	$558	$136	0.80

standardisation, including a trial of the 5-digit classification of occupation, may prove beneficial while attempts to explore "ranking" within occupations would be useful. Finally, an exploration of time-use and childcare data in an attempt to understand why women work less hours than men in all occupations (or, alternatively, why men work more hours than women) would add weight and illumination to this discussion.

Notes

1. Cook, Diana and Briggs, Phil (January 1997), *Gender Wage Gap - Scenarios of the Gender Wage Gap*, Report for the Ministry of Womens Affairs, NZ Institute of Economic Research, Wellington.

2. New Zealand Planning Council (1989), *Work Today*, New Zealand Planning Council, Wellington.

References

Cook, Diana and Briggs, Phil (1997) *Gender Wage Gap - Report for the Ministry of Womens Affairs*, New Zealand Institute of Economic Research (Inc), Wellington.

Horsfield, Anne (1988) *Women in the Economy*, Ministry of Women's Affairs, Wellington.

MathSoft (1993) *S-Plus Guide to Statistical and Mathematical Analysis*, MathSoft

New Zealand Planning Council (1989) *Work Today*, New Zealand Planning Council, Wellington.

Statistics New Zealand (1998) *Women*, Statistics New Zealand, Wellington (draft only).

Statistics New Zealand (1993) *All About Women in New Zealand*, Statistics New Zealand, Wellington.

Appendix 1

Technical Notes on Income Survey

The New Zealand Income Survey was run for the first time in the June 1997 Household Labour Force Survey (HLFS) quarter (April to June) as a supplement to the HLFS. The survey will be run in subsequent June HLFS quarters.

The New Zealand Income Survey was asked of all respondents to the HLFS. No data from proxies was accepted in the survey with the exception of those people who were unable to answer the survey on health or language grounds. All questions with the exception of the annual income question related to the respondent's most recent pay period. The following items were collected:

Actual and usual gross wages and salaries by:

Ordinary time
Overtime
Other income

For main job and all other jobs:

- Weeks and hours corresponding to earn the components of wages and salaries as above
- Sources of latest, actual, gross government transfers received
- Total latest, actual gross, government transfers received, in total and from each agency (NZISS, IRD and ACC)
- Total latest, actual gross private superannuation payment(s) received
- Total latest, gross, actual, 'all other private transfers including pensions, annuities' received
- Weeks covered for the transfer payments definedabove
- Approximate weekly, gross income received from self-employment including overtime
- Total annual gross income (range only as per Census categories) from all income sources

There were some forms of income that were not asked for in the New Zealand Income Survey for the recent income component of the survey. The largest of these is interest and investment income. It is important to note, therefore, that the income reported may not represent the entire income of an individual or household.

The collected data was cash only (pre-tax income was collected wherever possible) and did not include any non-cash fringe benefits.

Scope and coverage

People excluded from the HLFS survey sample (and, correspondingly, the New Zealand Income Survey sample) are:

- those living in non-private dwellings
- long-term residents of old people's homes
- hospitals and psychiatric institutions
- inmates of penal institutions
- members of the permanent armed forces
- members of the non-New Zealand armed forces
- overseas diplomats
- overseas visitors who expect to be resident in New Zealand for less than 12 months
- those aged under 15 years of age
- people living on offshore islands (except for Waiheke Island)

Response rate

Of the 90 percent of eligible households that responded to the HLFS, 80 percent of eligible people gave a valid response to the New Zealand Income Survey. The most common reasons for non-response to the New Zealand Income Survey were that a respondent was not able to be contacted or that a respondent was not able to provide the relevant information about their income when asked. Although little is known about the characteristics of the population that did not respond to the HLFS, there is a uniformity in responses to the New Zealand Income Survey across the various sub-populations analysed. Furthermore, there is no

under-representation of those likely to be in the lowest income brackets (those unemployed or not in the labour force).

Imputation

For those records where there was not a valid and usable response, a form of imputation known as hot-deck imputation was used. In this imputation method, a donor record replaces the non-usable record in the dataset. The donor record is chosen randomly from an imputation pool of records which have similar characteristics to that of the record to be imputed. The imputation pools were constructed on the basis of the following HLFS variables; age group, ethnicity, highest qualification, labour force status (modified to include whether they were wage and salary earners or self-employed), full time/part time work status and region. Given the possible combinations of these variables, these were in theory 720 possible imputation classes although not all of these were used. Imputation classes were combined in priority order when there were an insufficient number of donor records in any particular class.

Hot deck imputation was chosen for three reasons:

1 To maintain, or construct, an accurate income distribution of the New Zealand population
2 To allow the calculation and analysis of household income
3 To account for likely biases in household income analyses against larger households

There was little change in the income distribution or average income figures in the dataset due to imputation.

Reliability of estimates

The HLFS sample comprises approximately 15,000 private households, sampled on a statistically representative basis from rural and urban areas throughout New Zealand. The final New Zealand Income Survey dataset consists of approximately 22,500 valid records and 5,700 imputed records. The dataset used in this analysis consisted of 9,985 people in full-time employment (excluding those who were self-employed).

Creating sample errors associated with this analysis is no easy matter and, as this work is seen as exploratory only, has not been attempted. However, the reliability of the estimates is regarded as high, particularly when confined to the main part of the tree and the 12 sub-groups in Figure 1.

The numbers of respondents for each group are all regarded as reasonable and are as follows:

1.	204	7.	931
2.	953	8.	698
3.	1145	9.	2331
4.	269	10.	743
5.	770	11.	822
6.	469	12.	650

Total = 9,985 respondents

Appendix 2

Classifications

Ethnicity

1 = Maori and Pacific Islands
2 = European and Other

Occupation (NZSCO90 1-digit)

0 = Armed forces
1 = Legislators, administrators and managers
2 = Professionals
3 = Technicians and associate professionals
4 = Clerks
5 = Service and sales workers
6 = Agricultural and fishery workers
7 = Trades workers
8 = Machine operators etc.
9 = Elementary occupations

Occupation (NZSCO90 2-digit)

01 = Armed forces
11 = Legislators and administrators
12 = Corporate managers
21 = Physical, mathematical and engineering science professionals
22 = Life science and health professionals
23 = Teaching professionals
24 = Other professionals
31 = Physical science and engineering associate professionals
32 = Life science and health associate professionals
33 = Other associate professionals
41 = Office clerks
42 = Customer services clerks
51 = Personal and protective service workers
52 = Salespersons, demonstrators and models
61 = Market oriented agricultural and fishery workers
71 = Building trade workers
72 = Metal and machinery trades workers
73 = Precision trades workers
74 = Other craft and related trades workers
81 = Industrial plant operators
82 = Stationary machine operators and assemblers
83 = Drivers and mobile machinery operators
84 = Building and related workers
91 = Labourers and related service workers

Highest educational qualification

1 = No qualification
2 - 5 School Qualification
  2 = School certificate qualification
  3 = Sixth form qualification
  4 = Higher school qualification
  5 = Other school qualification
6-8 Post School Qualification
  6 = Vocational or trade qualification
  7 = Bachelor or higher degree
  8 = Other post-school qualification

Industry (NZSIC87 1-digit)

1 = Agriculture, hunting, forestry and fishing
2 = Mining and quarrying
3 = Manufacturing
4 = Electricity, gas and water
5 = Construction
6 = Wholesale, retail trade and restaurants
7 = Transport, storage and communication
8 = Business and financial services
9 = Community, social and personal services

Marital status

Married
Separated
Divorced
Widowed
Never married

Household type

Couple only
Couple with children
One parent family
Family type not known
Non-family household
One person household

Family code

Parent, first family
Child, first family
Parent, second family
Child, second family
Parent, third family
Child, third family
Member of fourth or subsequent family
Non-family member
One person household
Not applicable, including visitor

Author

Heather Kirkwood is a Social Statistics Analyst at
Statistics New Zealand,
PO Box 2922, Wellington.
E-mail: heather_kirkwood@stats.govt.nz