

# Gender differences in tax filing behavior: Is there a gender tax deduction gap?

Presentation at ACE – 19 July 2017.

Peter Varela

# Research Question

- Do men and women in similar economic circumstances claim different amounts of deductions on their tax return (main focus of paper)?
  - How large is this difference? (About 12%)
  - How robust is this difference to different specifications?
  - Which types of deductions are the largest contributors?
- And if so, what is driving this difference? (Lesser focus of paper, some questions left unanswered with this research design.)

# Overview of presentation/paper

- Introduction to the Australian tax system.
  - Types of deductions.
  - 2 per cent sample file of tax returns.
- Visual representation of the main results.
- Oaxaca-Blinder Decompositions.
- Different specifications. (Not discussed today)
  - Family tax planning.
  - Different types of income.
- DiNardo, Fortin and Lemieux Decomposition. (Not discussed today)

# Data

- ATO 2 per cent sample of personal income tax returns
  - Data for 2013/14
  - 257,639 unique records
- Data on:
  - Income
  - Each type of deduction
  - Partner status
  - Age
  - Occupation (at the one digit level).
  - Whether filed with a tax agent
  - Partner income

<u>Raw Differences</u>	Mean - Men	Mean - Women
<b>Work Related Expenses</b>		
Car	\$877.3606	\$433.2816
Travel	\$242.1494	\$80.3612
Uniform	\$149.3518	\$111.1986
Self-Education	\$87.59023	\$87.90331
Other WRE	\$672.4164	\$443.8699
<b>Total WRE</b>	<b>\$2024.869</b>	<b>\$1152.615</b>
<b>Other deductions</b>		
Dividend Deduction	\$117.8514	\$56.24083
Interest Deduction	\$71.72165	\$41.19763
Charitable Gift	\$252.0678	\$209.4201
Voluntary Superannuation Contribution	\$322.0257	\$239.7711
Cost of Tax Affairs	\$211.6603	\$136.6033
Other Deductions	\$223.5952	\$86.59049
<b>Total Deductions</b>	<b>\$3217.791</b>	<b>\$1916.438</b>

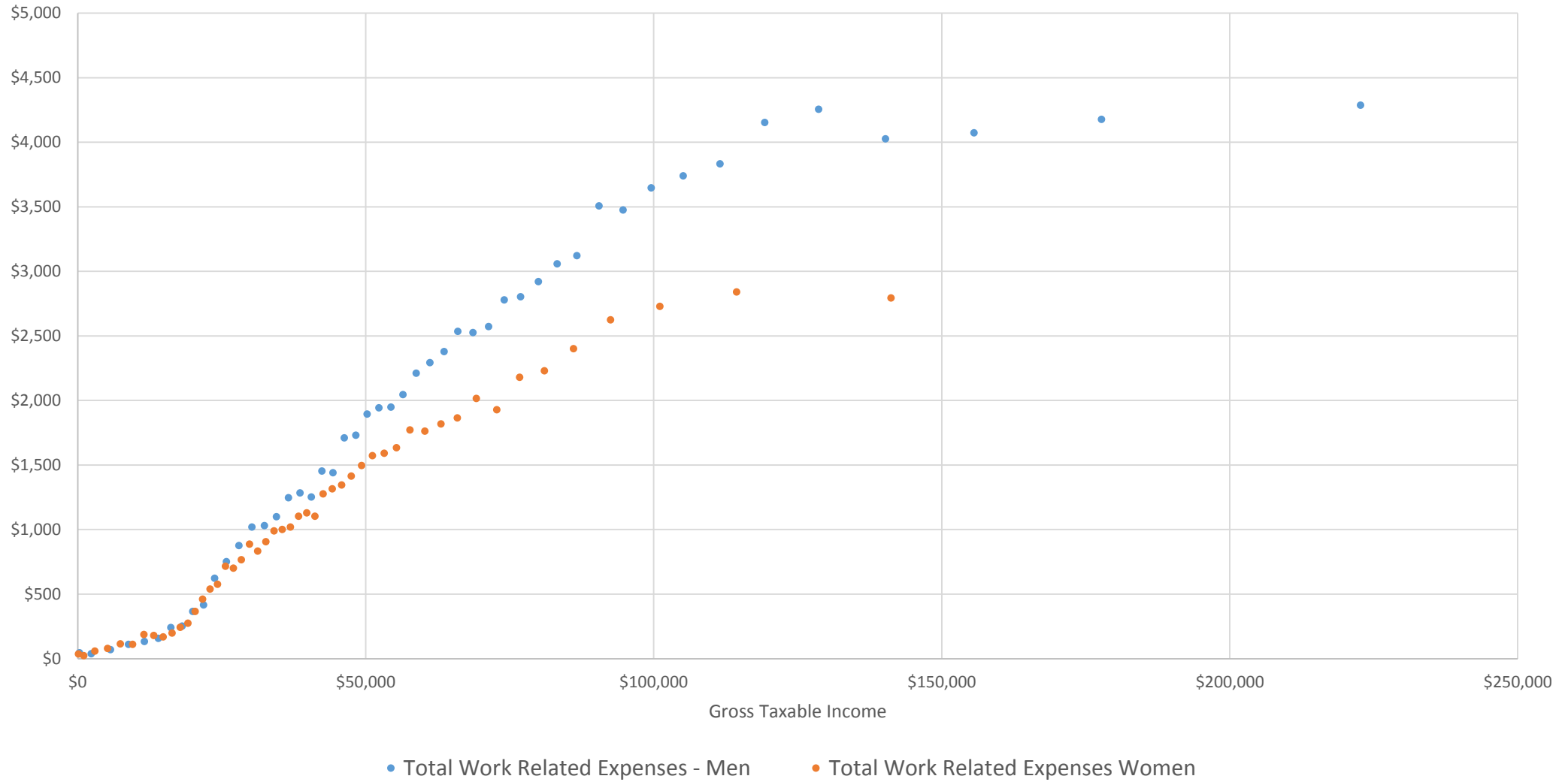
# Are women less likely to claim deductions?

## Supporting evidence

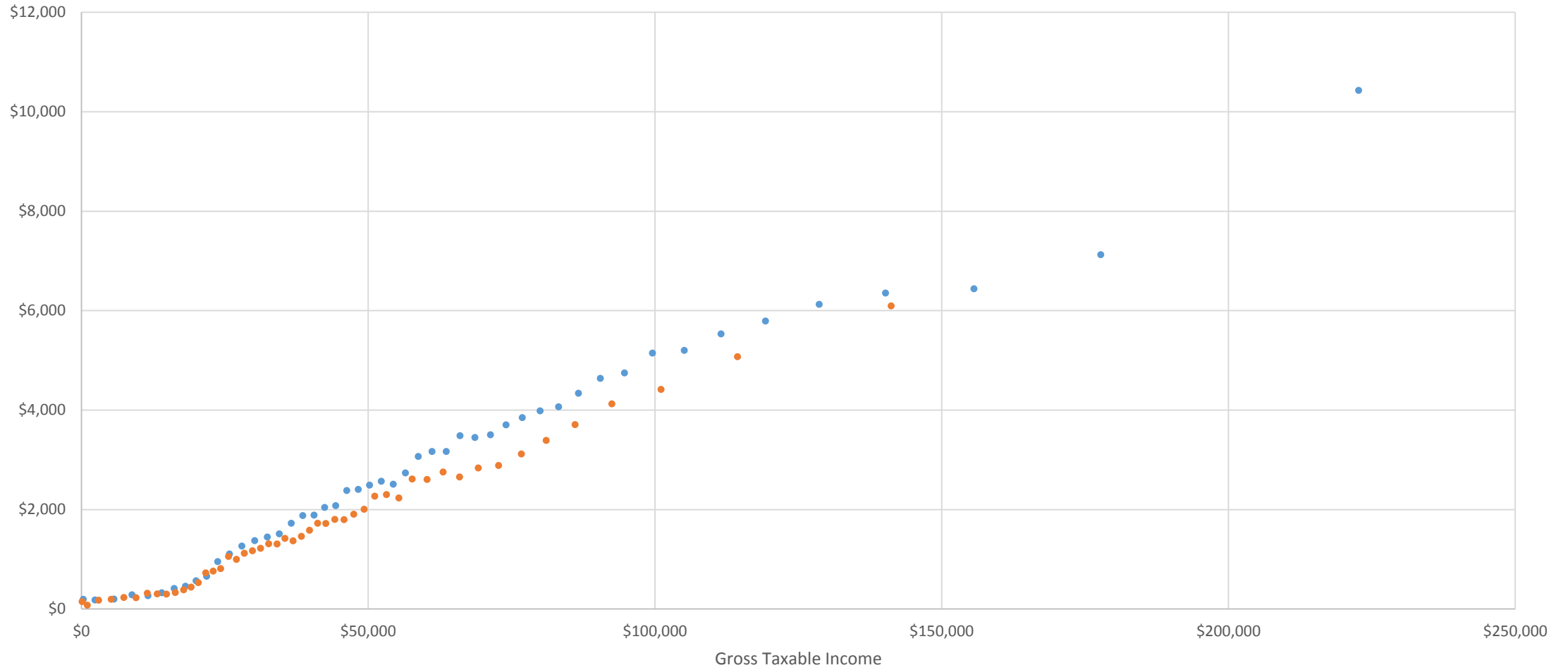
- Randomised audits find men are more likely to require ‘adjustments’ than women (Kleven et al. (2011), Paetzold and Winner (2014)).
- The World Values Survey has a question ‘is it ever justifiable to cheat on taxes if you have the opportunity to do so’. Women were more opposed, and this was the case in the vast majority of countries (McGee 2014).
- A variety of possible explanations of the gender pay gap are also relevant in this setting:
  - Men take weaker stances on ethical behavior (Glover et al. (1997)).
  - Are less risk averse (Croson and Gneezy (2009), Eckel and Grossman (2008)).
  - Generally more likely to commit crime (Schwartz and Seffenmeier 2008).

# Graphical representations

# Total Work Related Expenses

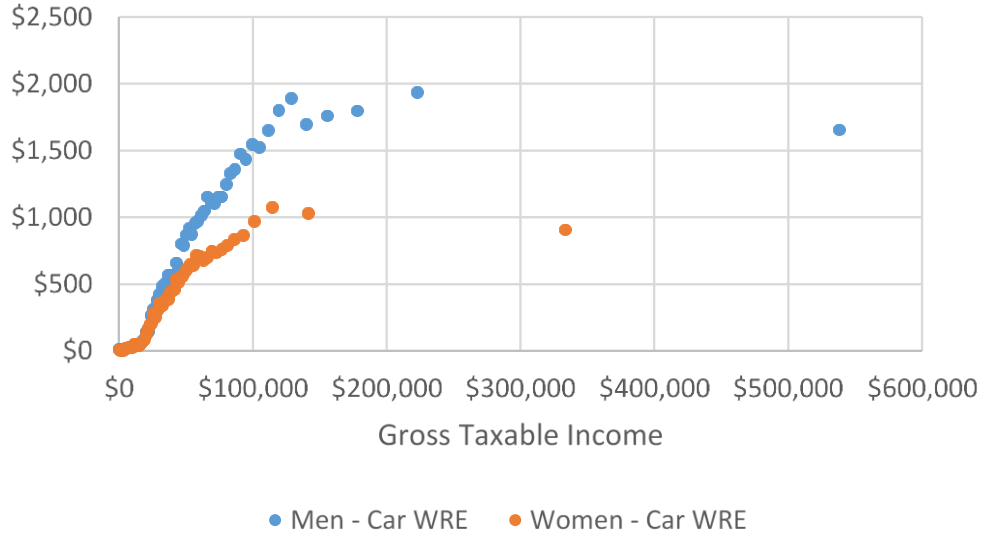


# Total Deduction Amount



• Total Deduction Amount - Men      • Total Deduction Amount - Women

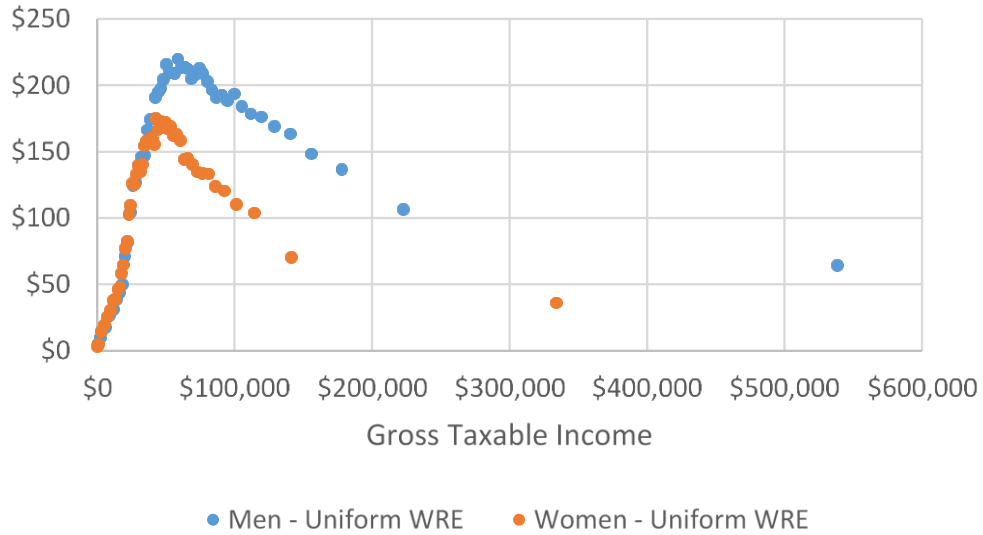
### Car WRE



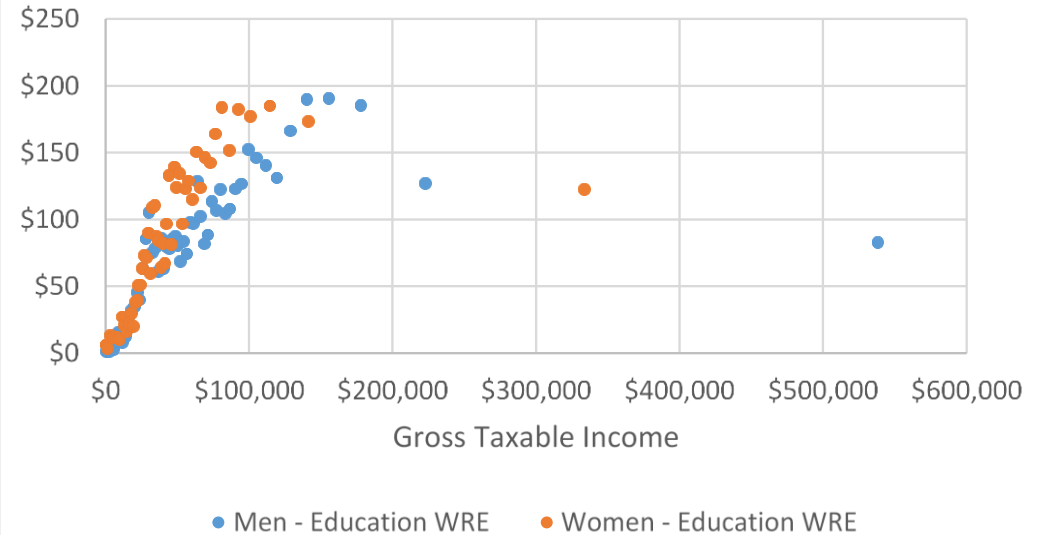
### Travel WRE

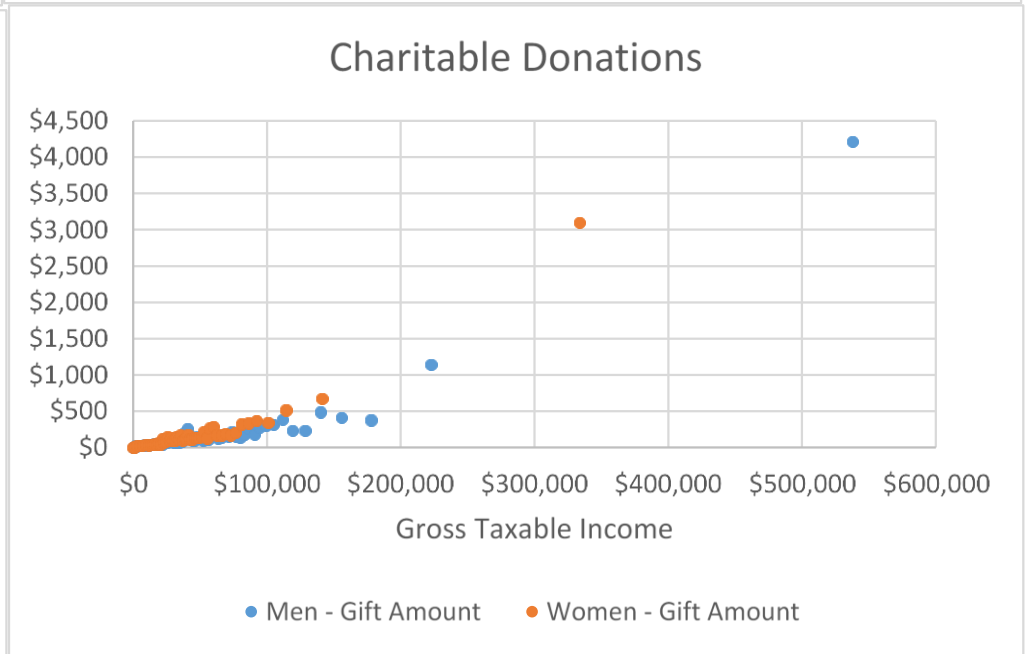
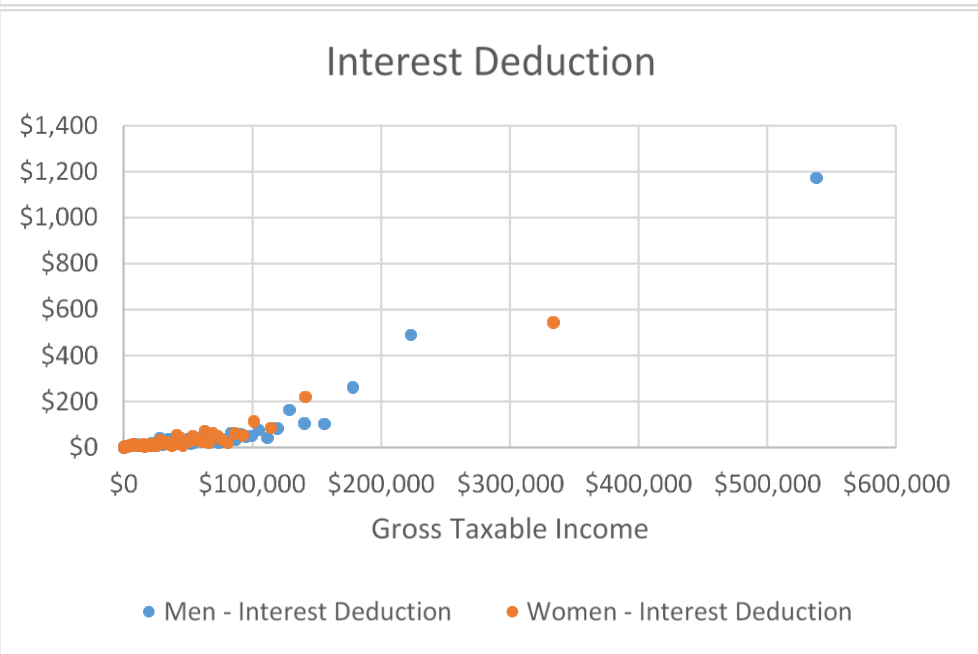
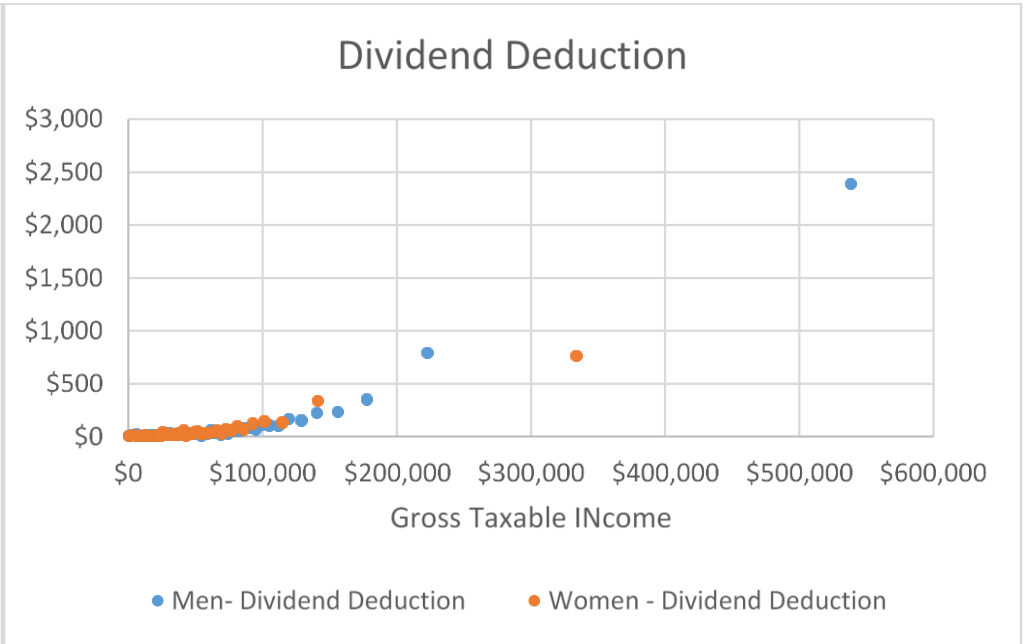
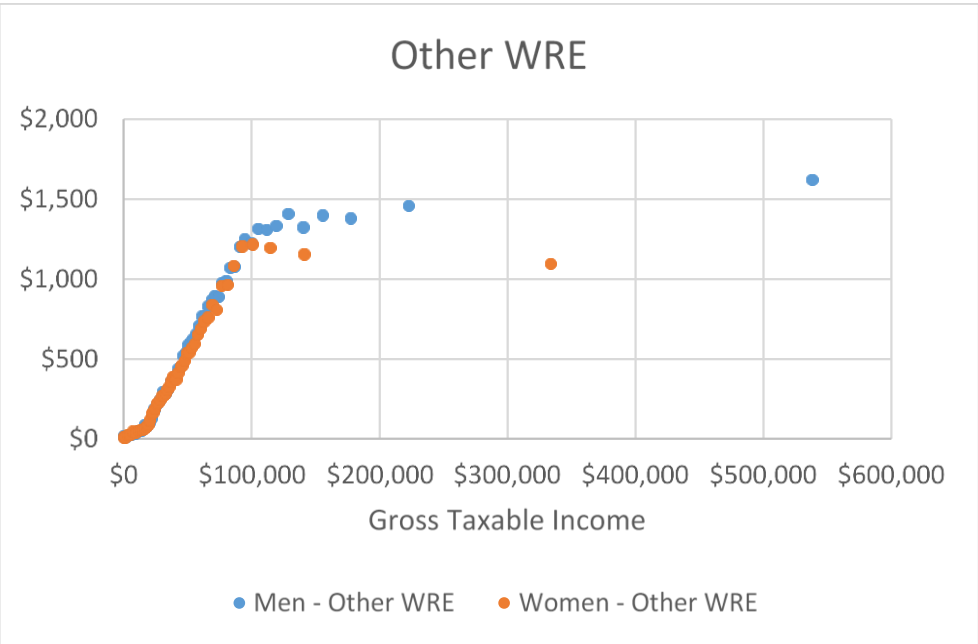


### Uniform WRE

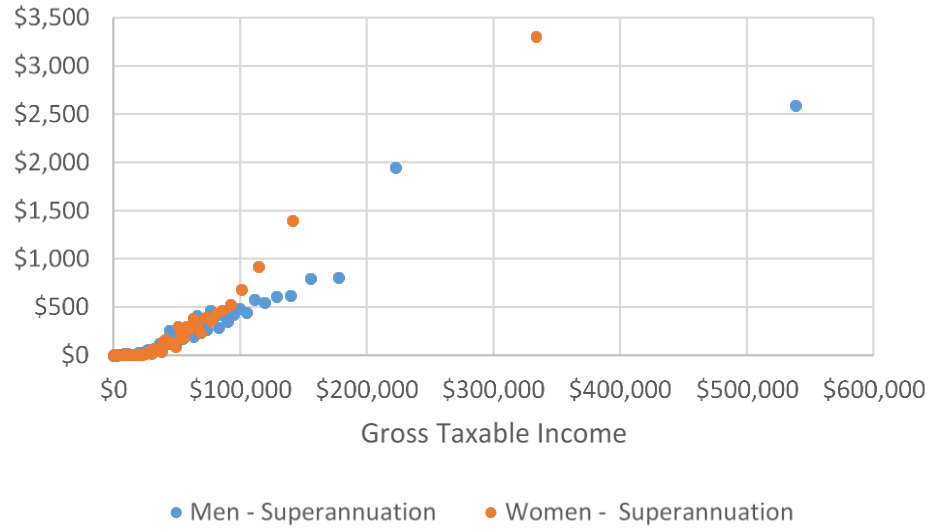


### Self Education Expenses WRE

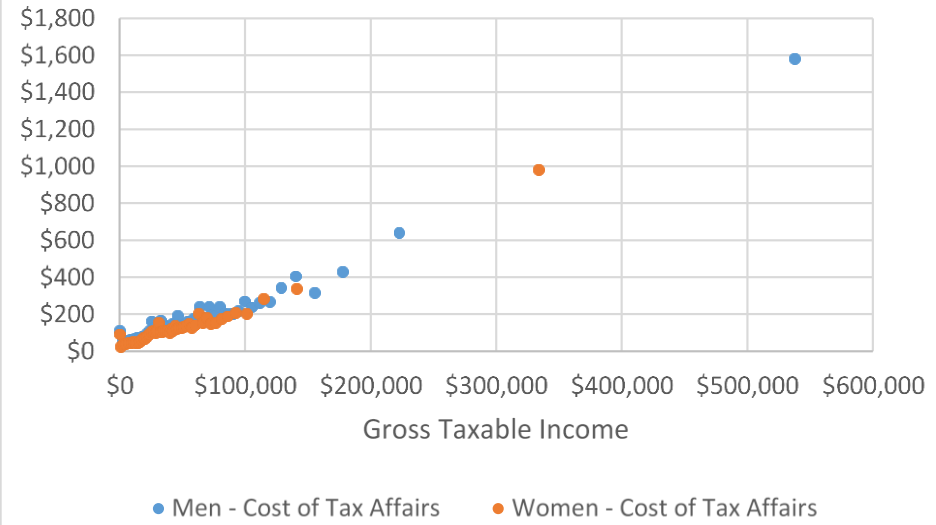




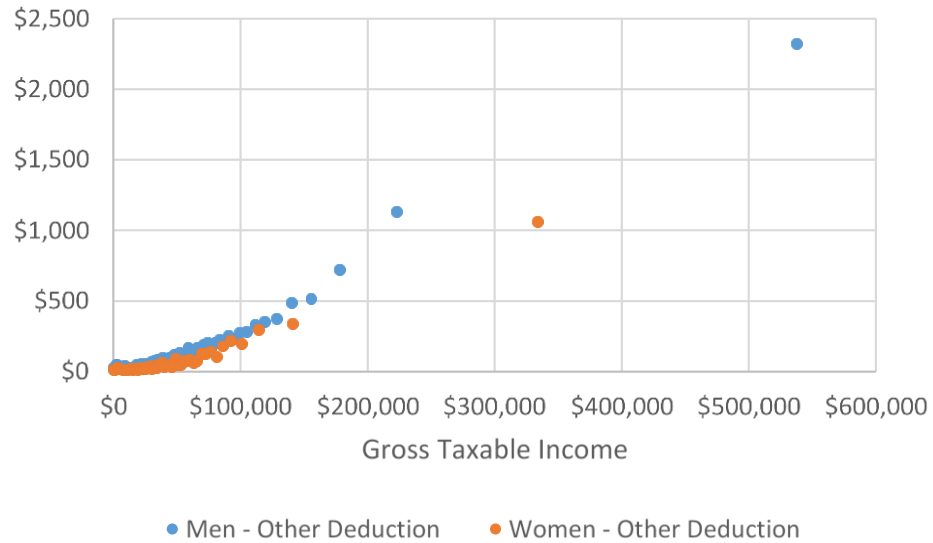
### Non-Employer Superannuation



### Cost of Tax Affairs



### Other Deductions



# Oaxaca-Blinder

- Decomposition technique based on Blinder (1973) and Oaxaca (1973).
- This paper uses the so called ‘two-fold’ decomposition
- Run two separate regressions for men and women:
  - $$\ln Deduction_i = Constant + \beta_1 \ln (Gross\ Taxable\ Income)_i + \beta_2 Occupation_i + \beta_3 Age_i + \beta_4 Partner\ Status_i + \beta_5 Lodgment\ Method_i + \varepsilon_i$$
- Then, combine the X variables of the women and the  $\beta$ s of the men to see how much of the gap between men and women is explained by the different X characteristics. The rest is an unexplained gap.
  - Difference =  $\beta_M(X_M - X_F) + X_F(\beta_M - \beta_F)$   
“explained” + “unexplained”

Oaxaca-Blinder Main Specification	Total WRE	Total Deduction Amount
Total difference (log percentage points)	52.0 (1.4)	57.1 (1.2)
Explained	33.8 (1.2)	45.5 (1.0)
Unexplained	18.2 (1.1)	11.6 (1.1)
Explained by:		
Ln (Total Income)	19.3 (0.3)	33.8 (0.5)
Occupation	11.9 (1.0)	7.9 (0.7)
Lodgment method	2.3 (0.1)	3.3 (0.1)
Age Range	0.2 (0.1)	0.0 (0.1)
Partner Status	0.0 (0.0)	0.3 (0.0)
Number of observations	257,639	257,639

	Difference (percentage points)	Explained (percentage points)	Unexplained (percentage points)
<b>Aggregates</b>			
Total WRE	52.0	33.8	18.2
Total Deduction Amount	57.1	45.5	11.6
<b>By Deduction</b>			
Car WRE	43.9	27.0	16.9
Travel WRE	22.3	10.0	12.3
Uniform WRE	49.0	42.2	6.8
Self-Education WRE	-3.0	-0.7	-2.3
Other WRE	47.0	24.7	22.4
Dividend Deduction	3.2	0.8	2.3
Interest Deduction	0.7	0.6	0.1
Charitable Gifts	-10.7	4.6	-15.3
Non-Employer Superannuation	5.5	7.3	-1.8
Cost of Tax Affairs	43.6	26.1	17.6
Other Deductions	32.1	8.9	23.1

# Summary of the rest of the paper

- Family tax planning
- Different sources of income
- DFL decomposition
- Negative gearing
- Fringe benefit taxation exemptions

Thank You

# Family tax planning – income shifting

- Through this analysis, individuals are treated as independent, but in reality many couples will file taxes in a coordinated manner.
  - Some deductions are transferrable between family members. For instance charitable donations.
  - There is an incentive to claim a deduction against the income of the person with the higher marginal tax rate.
  - To the extent that men are more likely to be in a higher tax bracket than their partner, this may be the cause of an observed gender deduction gap.

# Just look at singles?

- When just looking at singles, the 'unexplained portion' become small for Total Work Related Expenses (3.1 per cent), and negative for Total Deductions (-5.0 per cent)
  - However, it could just be that single men and single women are more similar than men and women generally.

# Including whether your partner pays a higher tax rate in the regression

- 2013/14 tax records include partner's income
  - Use this to construct a variable of whether you are in a higher tax bracket, same tax bracket and lower tax bracket.
- Add this to the baseline model:
  - $\ln Ded_i = Constant + \beta_1 \ln Gross\ Taxable\ Income_i + \beta_2 Occupation_i + \beta_3 Age_i + \beta_4 Partner\ status_i + \beta_5 Lodgment\ Method_i + \beta_6 (Higher\ tax\ bracket)_i + \beta_7 (lower\ tax\ bracket)_i + \varepsilon_i$
  - *Excludes singles from regression*

	Difference (log based percentage)	Explained	Unexplained	Explained by being in a different tax bracket
<b>Aggregates</b>				
<b>Total WRE</b>	52.0	42.5	9.5	10.6
<b>Total Deduction Amount</b>	57.1	46.8	10.3	1.9
<b>By Deduction</b>				
<b>Car WRE</b>	43.9	22.4	21.5	6.8
<b>Travel WRE</b>	22.3	12.0	10.2	2.5
<b>Uniform WRE</b>	49.0	48.7	0.3	7.5
<b>Self-Education WRE</b>	-3.0	-2.0	-1.0	-1.5
<b>Other WRE</b>	47.0	35.1	11.9	12.7
<b>Dividend Deduction</b>	3.2	0.1	3.0	-0.8
<b>Interest Deduction</b>	0.7	0.0	0.7	-0.8
<b>Charitable Gifts</b>	-10.7	12.5	-23.2	9.5
<b>Non-Employer Superannuation</b>	5.5	2.4	3.1	-5.7
<b>Cost of Tax Affairs</b>	43.6	36.7	6.9	13.0
<b>Other Deductions</b>	32.1	8.9	23.1	0.0

# Income source

- Does the different composition of income between men and women effect the results?
  - Use wage and salary income in the Oaxaca framework.
  - Remove groups that might have a strong influence on the results such as:
    - Those with very high “cost of tax affairs”.
    - Those with high business income.
    - Those with high “other income” (a proxy for self employed).
    - Those with a high unearned income.
    - Those outside of the 25-60 age range.
  - Include the proportion of different income types in the regression.

# Income Composition by Gender

	Men	Women	Difference	p value (difference = 0)
Salary/wage amount	71.3%	71.0%	0.3%	0.063
Allowances amount	1.4%	0.8%	0.6%	0.000
Employment termination payments taxable component	0.3%	0.1%	0.2%	0.000
Gross interest amount	2.6%	4.5%	<b>-1.9%</b>	0.000
Government pensions or allowances amount	3.9%	5.4%	<b>-1.6%</b>	0.000
Unfranked dividends	0.2%	0.2%	0.0%	0.002
Franked dividends (not including the credit)	2.4%	3.1%	-0.6%	0.000
Dividends franking credit	1.1%	1.4%	-0.3%	0.000
Net rental income	1.8%	2.4%	-0.5%	0.000
Net farm management deposits	0.0%	0.0%	0.0%	0.002
Net primary production business income	0.2%	0.1%	0.1%	0.000
Net non-primary production business income	5.7%	2.8%	<b>2.9%</b>	0.000
Net capital gains	0.7%	0.8%	-0.1%	0.000
Australian annuity or superannuation income stream - taxed	0.3%	0.3%	0.0%	0.899
Australian annuity or superannuation income stream - untaxed	1.2%	0.7%	<b>0.5%</b>	0.000
Other net foreign source income	0.1%	0.2%	0.0%	0.004
Other income not separately listed. Eg total income less listed income	2.8%	2.3%	0.5%	0.000
Net partnership and trusts primary production distributions	0.6%	0.5%	0.1%	0.001
Net partnership and trusts non-primary production distributions	5.0%	5.2%	-0.2%	0.002

# Using wage and salary income to predict total deductions

	Total WRE	Total Deductions
<b>Percentage point difference</b>	52.0	57.1
<b>Explained</b>	31.9	25.9
<b>Unexplained</b>	20.1	31.2
<b>Explained by</b>		
<b>Log (Salary and Wages)</b>	13.2	10.4
<b>Occupation</b>	15.8	10.6
<b>Age</b>	2.5	4.1
<b>Lodgment Method</b>	0.3	0.1
<b>Partner Status</b>	0.1	0.6

	Baseline (from section 3)	Removing the highest “cost of tax affairs”	Excluding high “business income”	Excluding high “other income”	Only including prime age	Excluding a high unearned income	Removing all groups
n (men)	134961	132693	117962	126446	94618	121069	88230
n (women)	123812	122377	112056	118402	86834	106799	83872
Percentage point difference	57.1	55.3	68.4	55.2	64.4	50.3	65.2
Explained	45.5	44.2	54.8	45.9	43.5	38.1	57.0
Unexplained	11.6	11.1	13.6	9.3	20.9	12.2	8.2
Explained by							
Ln (Total Income)	33.8	32.7	37.3	34.6	38.5	31.8	40.9
Occupation	7.9	8.1	13.9	8.1	2.3	2.7	13.3
Age	0.1	0.0	-0.1	0.0	0.0	0.4	0.1
Lodgment Method	3.3	3.1	3.3	3.1	3.0	3.2	2.8
Partner Status	0.3	0.3	0.3	0.2	0.2	0.1	-0.2

# Including types of income in the decomposition

- Add a vector of variables with proportions of income earned from each source:
  - $\ln Ded_i = Constant + \beta_1 \ln Gross\ Taxable\ Income_i + \beta_2 Occupation_i + \beta_3 Age_i + \beta_4 Partner\ status_i + \beta_5 Lodgment\ Method_i + \psi Prop_i + \varepsilon_i$
- This is equivalent to estimating a model in which each type of income has a different proportion of deductions associated with it.

# Results with proportions of income types

	Total WRE	Total WRE	Total Deductions	Total Deduction
<b>Includes income types</b>	Yes	No	Yes	No
<b>Percentage point difference</b>	52.0	52.0	57.1	57.1
<b>Explained</b>	36.0	33.8	47.8	45.5
<b>Unexplained</b>	16.0	18.2	9.3	11.6
<b>Explained by:</b>				
<b>Log Total income</b>	18.7	19.3	34.5	33.8
<b>Occupation</b>	12.6	11.9	8.5	7.9
<b>Age</b>	0.3	2.3	0.1	3.3
<b>Lodgment Method</b>	2.7	0.2	3.4	0.0
<b>Partner Status</b>	0.0	0.0	0.3	0.3
<b>Income types</b>	1.6		1.1	

# DFL (DiNardo, Fortin and Lemieux)

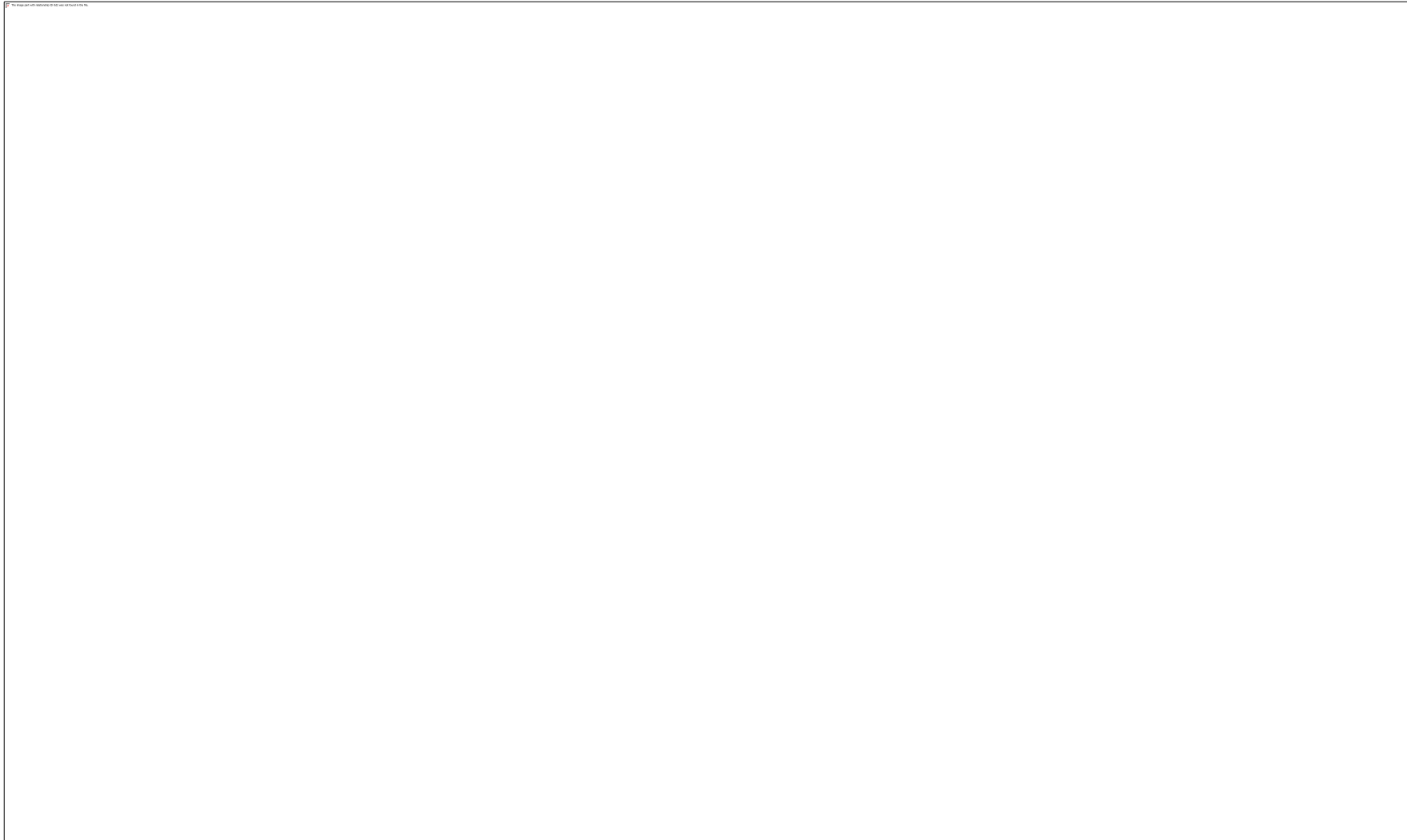
- An extension of the Oaxaca-Blinder approach
  - Creates a counterfactual density rather than a counterfactual mean
- Create a counterfactual distribution of females with the same observable characteristics as men.
  - Give higher weights to women who work in male dominated industries such as Machinery Operators and Drivers.
  - Give smaller weights to women who work in female dominated industries such as Community and Personal Service Workers

# Estimating weights

- $\psi (X) = \frac{\Pr (X \setminus \text{Gender} = \text{Men})}{\Pr(X \setminus \text{Gender} = \text{Women})}$
- In order to generate these probabilities, both the numerator and denominator are expanded using Bayes rule. ie
  - $\Pr(X \setminus \text{Gender} = \text{Men}) = \Pr \frac{(\text{Gender} = \text{Men} \setminus X) \cdot \Pr(X)}{\Pr(\text{Men})}$
  - $\psi (X) = \frac{\Pr(\text{Gender} = \text{Men} \setminus X) / \Pr(\text{Men})}{\Pr(\text{Gender} = \text{Women} \setminus X) / \Pr(\text{Women})}$
- Where the probabilities are obtained with a logit estimation that predicts gender using covariates (X) (but not level of deduction).
  - The logit predicts gender with Log Gross Income, Occupation, Age, Partner Status, Lodgment Method, and a dummy for having zero income.

	Men	Women	Women (With DFL weights)
<b>Gross Taxable Income</b>	71556.48	48275.14	65495.9
<b>Proportion with Zero Income</b>	0.46%	0.58%	0.48%
<b>Proportion with partner</b>	56.29%	54.62%	54.04%
<b>Managers</b>	11.49%	7.70%	11.35%
<b>Professionals</b>	14.37%	19.61%	14.19%
<b>Technicians and Trades Workers</b>	15.21%	2.51%	14.10%
<b>Community and Personal Service Workers</b>	4.72%	11.53%	4.72%
<b>Clerical and Administrative Workers</b>	4.93%	18.56%	4.97%
<b>Sales workers</b>	4.14%	8.17%	4.16%
<b>Machinery operators and drivers</b>	8.41%	0.54%	8.19%
<b>Labourers</b>	10.35%	5.24%	9.86%
<b>Consultants, apprentices and Not Specified</b>	7.12%	7.66%	7.15%
<b>Occupation not listed/ specified</b>	19.26%	18.48%	21.31%

# DFL – Total Work Related Expenses



# DFL – Total Deductions



# What can we conclude?

- On average, men claim more deductions than women, and this difference can't be explained by different levels of income or different occupations.
  - This is true across various categories of deductions.
  - This finding is robust to a wide variety of specifications.
  - Equivalent to around \$240 in deduction or around \$75 in take home income.
- However, while this paper has looked at some potential explanations for this difference, further research is required to determine the relative importance of these factors.

# What can we do (policy options)?

- The correct response depends on what is causing the difference
  - If it is a story of information, then providing this information may help.
  - If deductions are allowed for items used commonly by men, but not for similar items used by women, then the correct response is to change the tax code (either allowing items for women or disallowing items used by men).
  - If men are more willing to risk audit, then the tax office, should include gender in the process of determining who to audit.
  - Automating more parts of the tax system could be useful.

Additional assorted slides

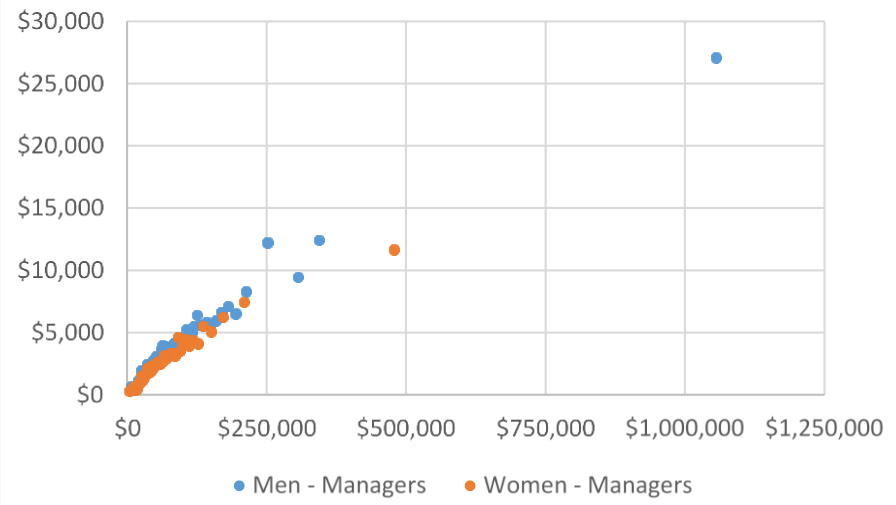
# Results for different years

	Total WRE Difference (percentage points)	Explained (percentage points)	Unexplained (percentage points)	Total Deduction Difference (percentage points)	Explained (percentage points)	Unexplained (percentage points)
2013/14	52.0 (1.4)	33.8 (1.2)	18.2 (1.1)	57.1 (1.2)	45.5 (1.0)	11.6 (1.1)
2012/13	53.8 (1.4)	34.7 (1.2)	19.1 (1.1)	58.9 (1.2)	45.4 (1.0)	13.5 (1.1)
2011/12	55.8 (1.3)	32.7 (1.2)	23.1 (1.1)	59.5 (1.2)	42.9 (0.9)	16.6 (1.1)
2010/11	62.0 (1.9)	52.6 (1.7)	9.3 (1.8)	67.6 (1.7)	65.5 (1.5)	2.2 (1.7)
2009/10	57.9 (1.9)	49.2 (1.7)	8.7 (1.8)	61.9 (1.7)	62.9 (1.5)	-1.0 (1.7)
2008/09	62.5 (1.9)	54.0 (1.7)	8.5 (1.8)	67.4 (1.7)	65.9 (1.5)	1.4 (1.7)
2007/08	60.4 (1.8)	50.0 (1.7)	10.4 (1.8)	63.6 (1.7)	60.6 (1.5)	3.0 (1.7)
2006/07	63.7 (1.9)	55.2 (1.8)	8.5 (1.9)	68.4 (1.8)	70.4 (1.6)	-2.0 (1.9)
2005/06	67.2 (2.0)	61.5 (1.8)	5.6 (1.9)	69.4 (1.8)	77.1 (1.7)	-7.7 (1.9)
2004/05	71.6 (2.0)	66.8 (1.9)	4.8 (2.0)	76.3 (1.8)	83.1 (1.7)	-6.7 (1.9)
2003/04	68.7 (2.0)	62.7 (1.9)	6.0 (2.0)	74.4 (1.8)	78.0 (1.8)	-3.5 (2.0)

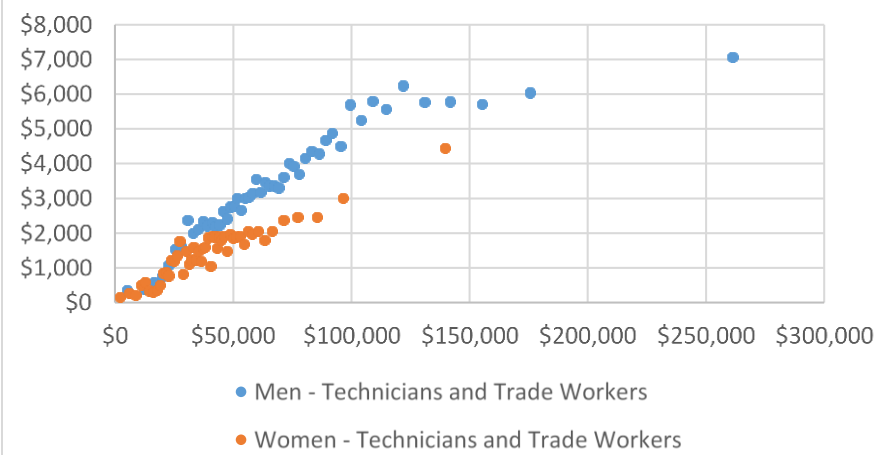
# Estimates for WRE with different groups excluded

	Baseline (from section 3)	Excluding the highest "cost of tax affairs"	Excluding high "business income"	Excluding high "other income"	Including only prime age workers	Excluding a high unearned income	Excluding all groups
n (men)	134961	132693	117962	126446	94618	121069	88230
n (women)	123812	122377	112056	118402	86834	106799	83872
Percentage point difference	52.0 (1.4)	51.9 (1.4)	74.5 (1.4)	52.0 (1.4)	56.1 (1.5)	41.5 (1.4)	72.2 (1.4)
Explained	33.8 (1.2)	33.9 (1.2)	54.4 (1.1)	35.0 (1.2)	31.1 (1.3)	23.7 (1.1)	61.2 (1.1)
Unexplained	18.2 (1.1)	17.9 (1.1)	20.0 (1.2)	16.9 (1.2)	25.0 (1.4)	17.7 (1.2)	11.0 (1.4)
Explained by							
Log (Total Income)	19.3 (0.3)	19.4 (0.4)	25.9 (0.4)	20.8 (0.4)	24.9 (0.5)	21.2 (0.4)	38.6 (0.6)
Occupation	11.9 (1.0)	12.1 (1.0)	25.8 (0.9)	11.8 (1.0)	3.5 (1.2)	-0.1 (1.0)	19.3 (0.8)
Age	2.3 (0.1)	2.2 (0.1)	0.1 (0.2)	0.2 (0.1)	0.5 (0.1)	0.0 (0.2)	0.7 (0.1)
Lodgment Method	1.2 (0.1)	0.2 (0.1)	2.5 (0.2)	2.2 (0.1)	2.3 (0.2)	2.6 (0.1)	2.6 (0.2)
Partner Status	0.0 (0.0)	0.0 (0.0)	0.1 (0.0)	0.0 (0.0)	-0.1 (0.0)	0.0 (0.0)	-0.1 (0.0)

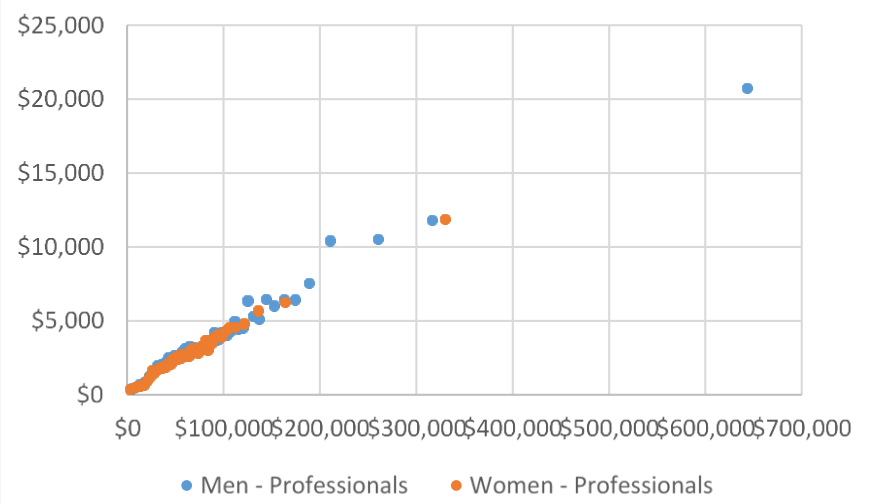
Total Deductions: Managers



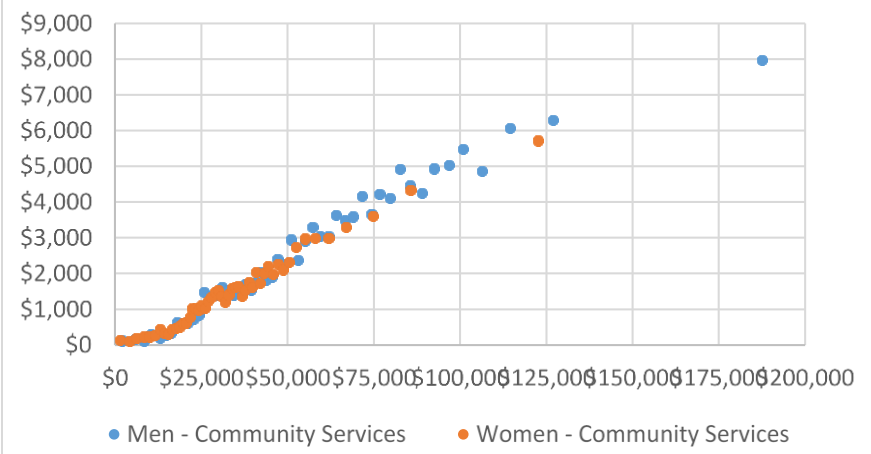
Total Deductions: Technicians and Trade Workers



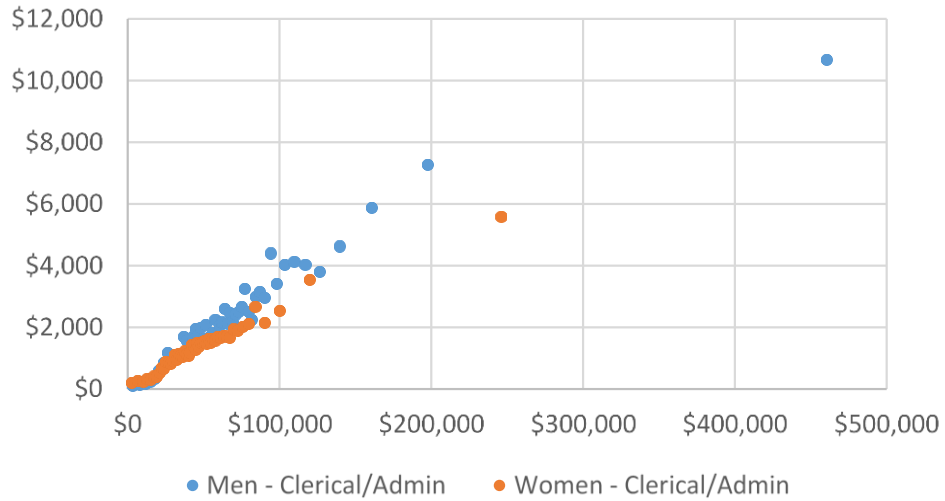
Total Deductions: Professionals



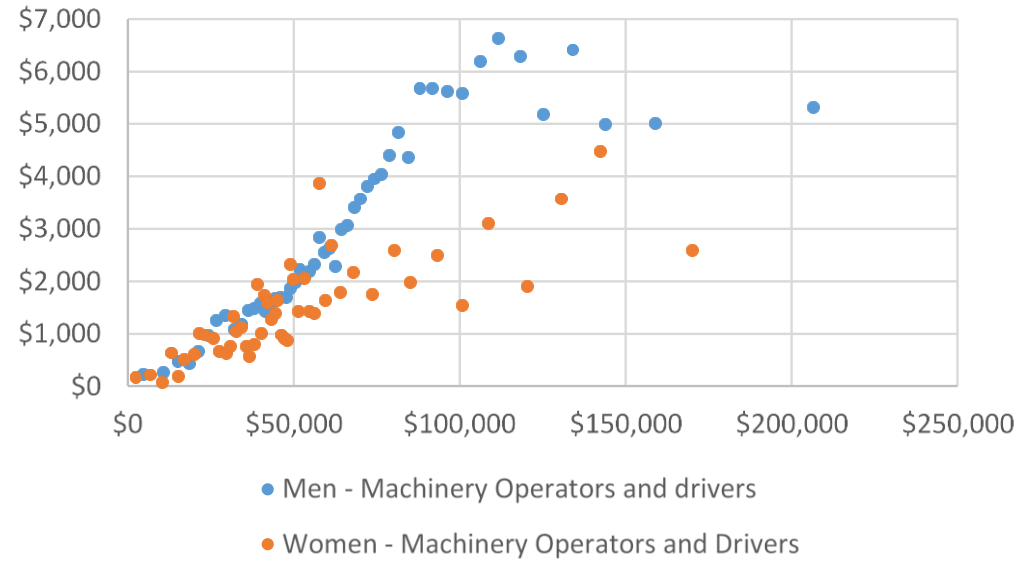
Total Deductions: Community Service Workers



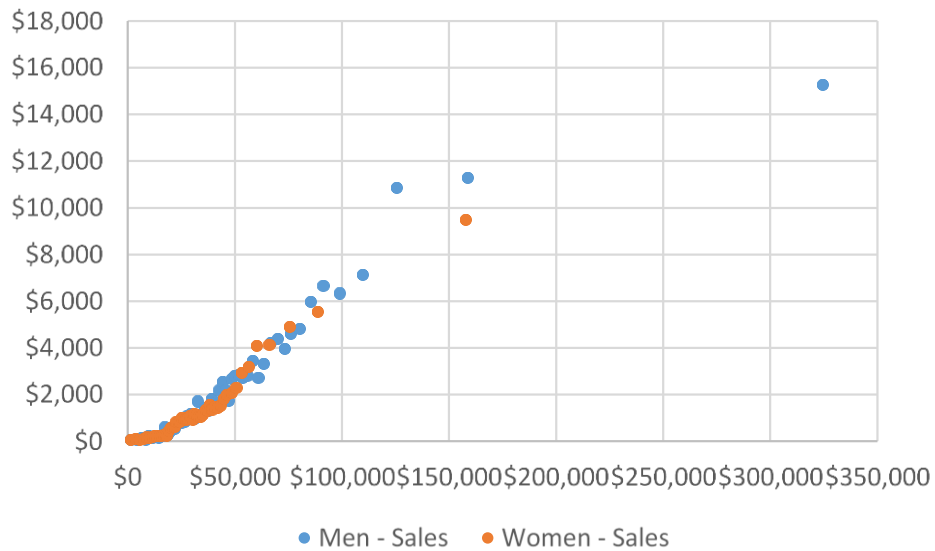
### Total Deductions: Clerical and Administrative workers



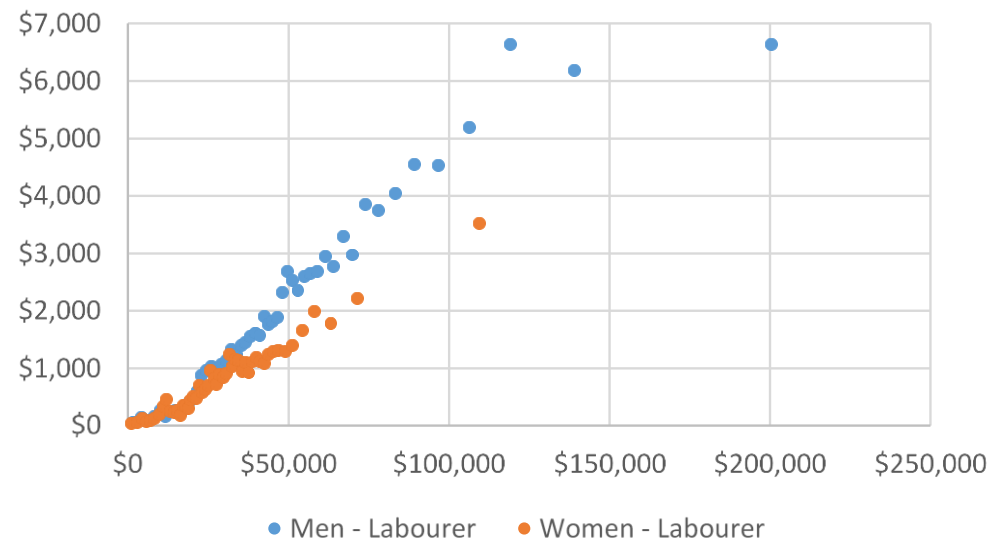
### Total Deductions: Machinery Operators and Drivers



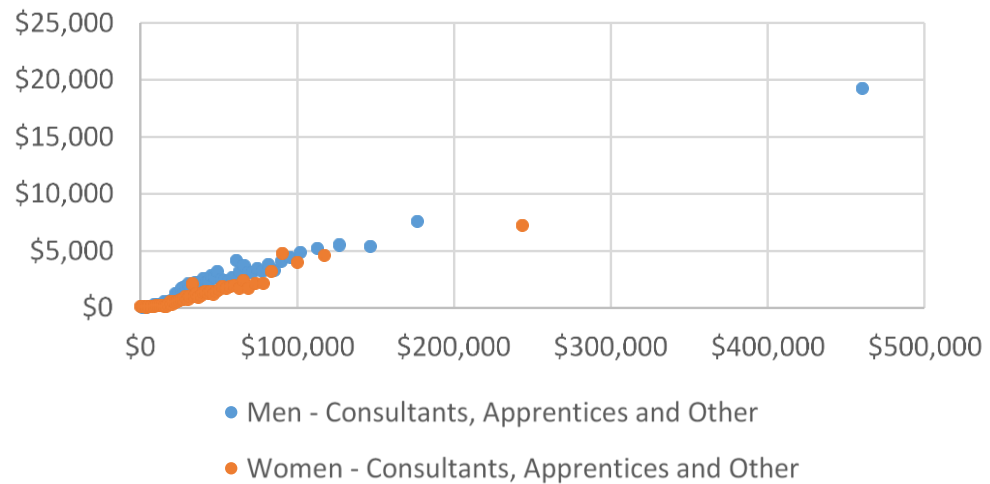
### Total Deductions: Sales Workers



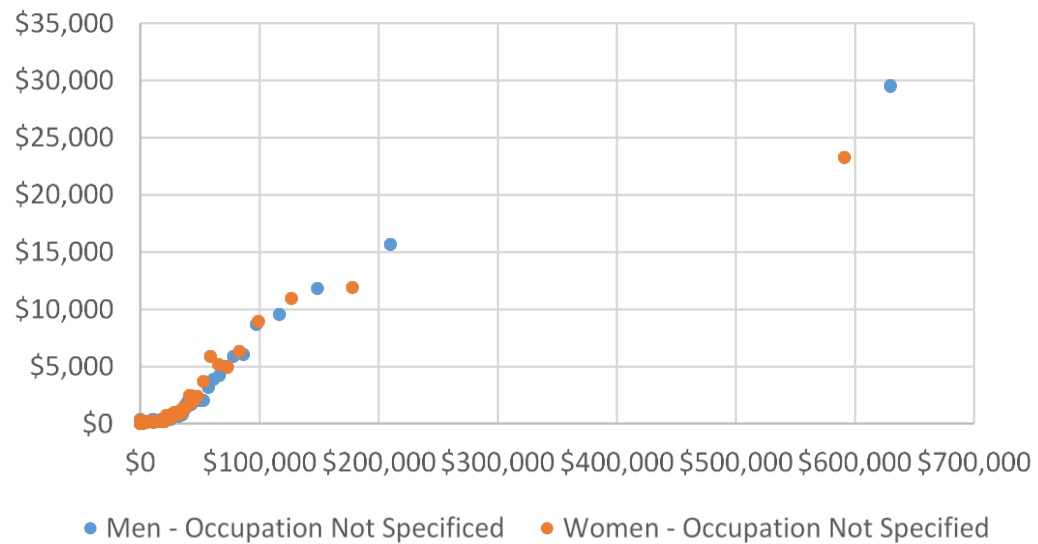
### Total Deductions: Labourers



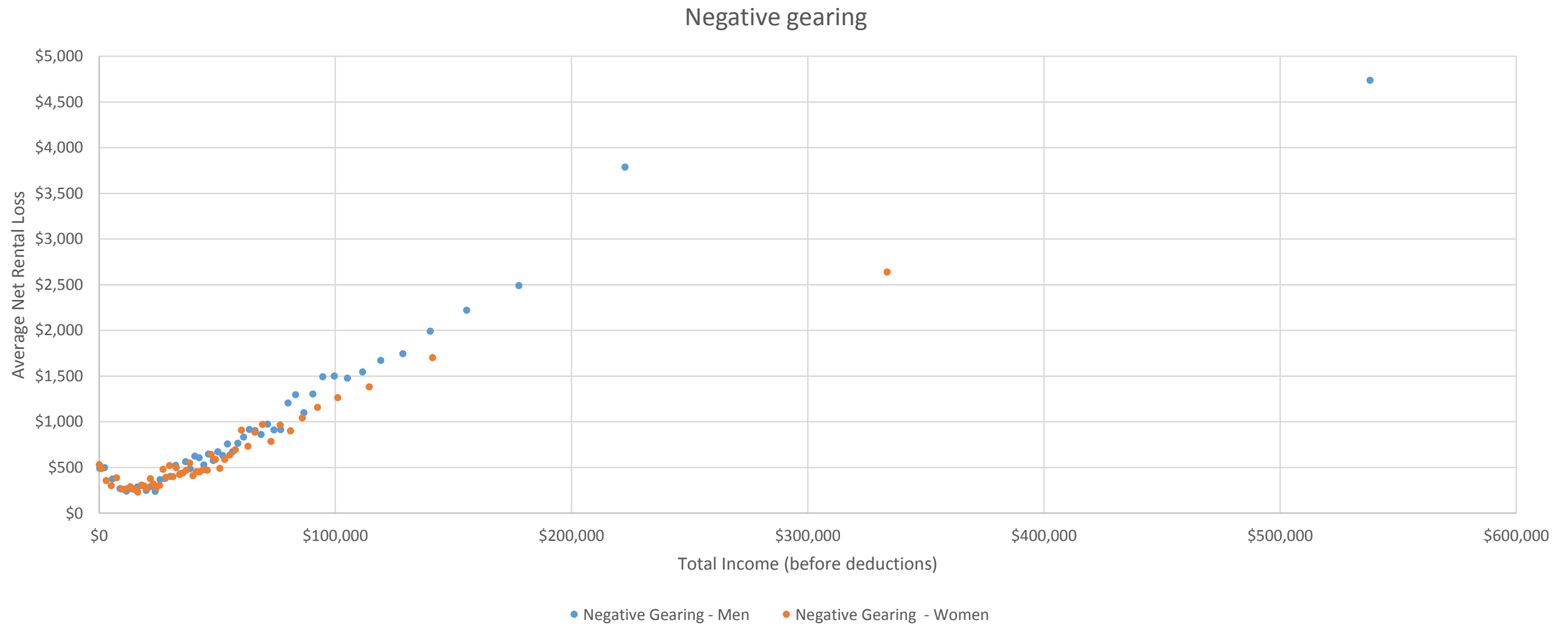
### Total deductions: Consultants, Apprentices and Other



### Total Deductions: Occupation Not Specified



# Negative Gearing



# Negative Gearing

	Negative Gearing (log percentage points)
<b>Total difference</b>	12.9
<b>Explained</b>	0.4
<b>Unexplained</b>	12.5
<b>Explained by:</b>	
<b>Ln (Total Income)</b>	4.0
<b>Occupation</b>	-6.2
<b>Lodgment method</b>	2.0
<b>Age Range</b>	-0.1
<b>Partner Status</b>	0.6
<b>Number of observations</b>	257,639