Gender and Racial Inequality

Part 1: Introduction

Gender and racial inequality have been an ongoing topic of debate and controversy in the modern century (United Nations, 2021). Throughout history, there have been many instances of gender and racial inequality, such as women being barred from voting in national elections, and bathrooms segregated based on race. Humanity has made great progress fixing these inequalities; women can now vote, drive, work, and do many other things that they were not able to before, notwithstanding some extreme cases like the Taliban-ruled Afghanistan, under which women still struggle to achieve some of their basic human rights (United Nations, 2023). However, while many cases of extreme inequality have been eradicated, lesser forms of inequality are still present, among which the most heavily-discussed being salary-based inequality (Brito, 2022) (Lang & Spitzer, 2020, p. 81). There is debate as to whether women really earn less than men or whether one race earns less than another. In this report, we will evaluate some important variables that determine a person's income, and try to fit a model that allows us to examine the underlying causal relationship between income and other important variables, especially race and sex.

With the provided cross-sectional Australian economic data, we can begin to analyse our variables of interest. Specifically, we are interested in whether income differs across sexes and race, which in this case would be related to indigenous status. One thing to note: The variables in the dataset represent average values, which has implications for estimation.

Mean Average Income by Gender and Indigenous Status

Category	Mean Average Income
Male	786.5619
Female	672.3603
Indigenous	687.3370
Non-Indigenous	745.5046

fig.1

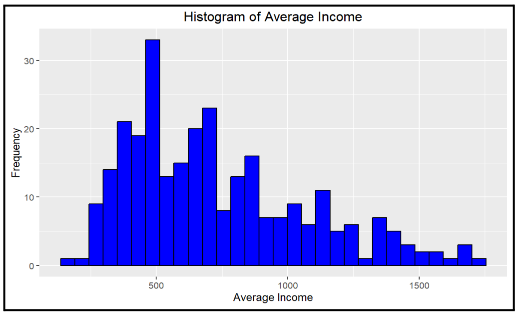


fig.2

Fig.1 shows the mean average income (MAI) achieved by the different sexes and races in Australia. We observe that males have a MAI that is more than AUD 100 higher than females. On the other hand, non-indigenous persons have a MAI that is more than AUD 50 more than indigenous persons. These statistics indicate a possibility of gender-based or racial-based discrimination in the workforce. However, these results alone are not conclusive as we have not yet taken into account variables and held them constant, so causality is not yet determined.

Fig.2 shows the distribution of our dependent variable, average income. We can see that it is positively skewed, with few observations having average incomes much higher than the rest. This aligns with the reality of income inequality across the world, in which certain demographics have a much higher income compared to others.

Part 2: Modelling

When creating our initial model, we assessed what factors have significant implications for model quality. We determined that logging the dependent variable, income, would improve the quality of our models by normalising its distribution **[fig.3]**. A non-normal distribution could violate several assumptions of the Ordinary Least Squares. A positively-skewed income variable may lead to heteroskedasticity as the variance of income increases with higher values, the model's ability to predict higher values becomes less accurate, leading to larger residuals at higher levels and vice versa. The assumption of normally-distributed residuals is also violated; if the dependent variable is highly non-normal, it often indicates that the relationship between Y and the X's is not well captured by a simple linear model. This misspecification can cause the residuals to deviate from normality.

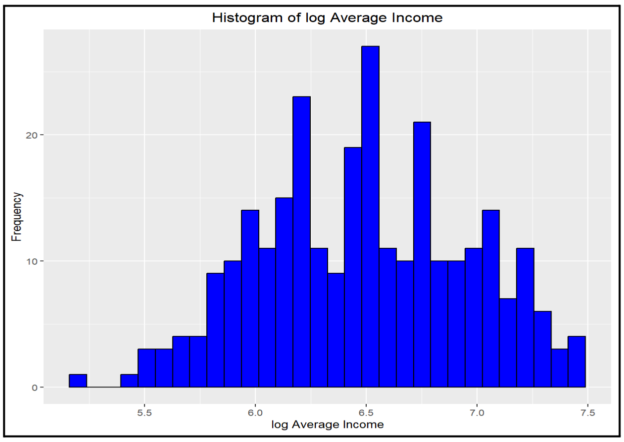


fig.3

At first, we did model specification before addressing HTSK which was incorrect **(See Appendix A)**. We removed "no_people" because its interpretation was nonsensical. Therefore, we overlooked using that as a weight in our model. Instead, we used "1 ÷ (age + no_years_school)" as these two variables hinted at HTSK from their scatter plots **(See Appendix B)**. Even though it passed the Breusch-Pagan test, "no_people" should have been the main cause of HTSK. After reconsideration, we used "no_people" as the weight.

Setting the weight as "1 ÷ no_people" means that both sides of the model equation are multiplied by that which applies the weightage while preserving the linear relationship between the dependent and independent variables. Since "no_people" causes HTSK, applying the weight corrects this by adjusting each observation based on its group size. In WLS, the weights ensure that smaller groups with more variable incomes are given more importance, and vice versa. This adjustment makes the betas BLUE without affecting the coefficients' interpretations.

$$w\left(\widehat{log\left(income\right)}\right) = w(\widehat{\beta_0} + \widehat{B} \cdot region + \widehat{\beta_{15}} fulltime + \widehat{\beta_{16}} engprofNotWell + \widehat{\beta_{17}} engprofWell + \widehat{\beta_{18}} engprofVeryWell + \widehat{\beta_{19}} Female + \widehat{\beta_{20}} NonIndigenous + \widehat{\beta_{21}} age + \widehat{\beta_{22}} no\_years\_school)$$

$Where: w = \frac{1}{no\_people}$; $\widehat{B} \cdot region = \left(\widehat{\beta_1} Adelaide, \widehat{\beta_2} Brisbane, ..., \widehat{\beta_{14}} WA\right)$ (Australian Capital Territory is base)

	Dependent variable:
	log(income)
regionGreater Adelaide	0.024
	(0.109)
regionGreater Brisbane	-0.146
	(0.109)
regionGreater Darwin	0.314***
	(0.103)
regionGreater Hobart	-0.022
	(0.126)
regionGreater Melbourne	-0.175
	(0.111)
regionGreater Perth	-0.036
	(0.120)
regionGreater Sydney	0.195
	(0.130)

fig.4a

regionRest of NSW	0.040
	(0.105)
regionRest of NT	-0.048
	(0.104)
regionRest of Qld	0.041
	(0.108)
regionRest of SA	-0.064
	(0.100)
regionRest of Tas.	-0.016
	(0.105)
regionRest of Vic.	0.193*
	(0.098)
regionRest of WA	-0.202*
	(0.111)
fulltime	0.600***
	(0.037)
englishproficiencyNot well	0.211***
	(0.059)

fig.4b

englishproficiencyVery well	0.784***
	(0.089)
englishproficiencyWell	0.307***
	(0.074)
Female	-0.085***
	(0.031)
indigenous_statusNon-Indigenous	0.335***
	(0.072)
age	0.013***
	(0.002)
no_years_school	0.074***
	(0.016)
Constant	4.342***
	(0.234)
Observations	282
R^2	0.808
Adjusted R^2	0.792
Residual Std. Error	0.046 (df = 259)
F Statistic	49.672*** (df = 22; 259)
Note:	*p<0.1; **p<0.05; ***p<0.01

fig.4c

Most of the Australian regions are insignificant with regards to income (Only Darwin is, at $a = 0.05$). When running a joint hypothesis test on all k regions, with all the regions being the restrictions, we get a significant p-value (p < 0.0001) **(See Appendix D)**. This leads to the rejection of the null hypothesis which states that none of the regions are significant in explaining mean income, and the acceptance of the alternative hypothesis that at least one of the regions is significant in this regard.

Running a Bruesch-Pagan test on this model, we can check if HTSK is present or not. The null and alternative hypotheses is as follows:

$H_o: Var(u_i) = \sigma^2$ (homoskedasticity) $H_1: Var(u_i) = \sigma_i^2$ (heteroskedasticity)

The p-value for the Breusch-Pagan test of this model is exactly 1, which suggests not enough evidence for heteroskedasticity in the model. However, a p-value of exactly 1 is abnormal, which may suggest multicollinearity or overfitting.

When running a Bruesch-Pagan test for this model (as well as all subsequent models, without having incorporated the weightage of 1/no_people) we get significant p-values, even at the 0.001% significance level **(See Appendix E)**, concluding the presence of heteroskedasticity. However, incorporating the specified weightage results in insignificant p-values for the Breusch-Pagan tests of all subsequent models **(See Appendix E)**.

To further determine if region is a significant variable in this model, we can simplify our model and increase our results' statistical power. Looking at fig.4, there are many variables, representing each region. Having many regions means each region has a small sample size and larger variability. Hence, the power of the test may be low which increases the Type II error risk—failing to detect a true effect of some regions on income. We can simplify the model by grouping the regions into "urban" and "rural" which helps address these issues in addition to parsimony. Additionally, the urban-rural distinction is more meaningful for global interpretation, focusing on well-understood economic differences, enhancing the relevance of the results for policy analysis. The enhanced model is as follows:

$$w\left(\widehat{log\left(income\right)}\right) = w(\widehat{\beta_0} + \widehat{\beta_1}urban + \widehat{\beta_2}fulltime + \widehat{\beta_3}engprofNotWell + \widehat{\beta_4}engprofWell + \widehat{\beta_5}engprofVeryWell + \widehat{\beta_6}Female + \widehat{\beta_7}NonIndigenous + \widehat{\beta_8}age + \widehat{\beta_9}no\_years\_school)$$

$$Where: w = 1 \div no\_people$$

	Dependent variable: log(income)
urban	0.045
	(0.039)
fulltime	0.652***
	(0.037)
englishproficiencyNot well	0.177***
	(0.051)
englishproficiencyVery well	0.821***
	(0.086)
englishproficiencyWell	0.256***
	(0.073)
Female	-0.119***
	(0.033)

fig. 5

indigenous_statusNon-Indigenous	0.395***
	(0.061)
age	0.010***
	(0.002)
no_years_school	0.047***
	(0.017)
Constant	4.682***
	(0.211)
Observations	282
R^2	0.735
Adjusted R^2	0.726
Residual Std. Error	0.053 (df = 272)
F Statistic	83.800*** (df = 9; 272)
Note:	*p<0.1; **p<0.05; ***p<0.01

fig. 5 (continued)

The results show that the dummy variable for urban turned out to be insignificant, meaning that there is no statistically significant difference in income between urban and rural areas, ceteris paribus. Given that this variable was assumed to have a higher statistical power compared to all individual regions for reasons explained previously, we have decided to remove all regions entirely, grouped or not, from our final model.

The AIC and BIC values, metrics that measure a model's goodness-of-fit, of this second model are slightly higher than the first model. This shows that the first model may have been a better fit. However, given the various benefits of the urban/rural grouping (parsimony, simplicity, ease of interpretation, stronger statistical power), we can disregard this.

In constructing our final model, nuance was key. Given the focus on gender and racial inequality, we decided to create three interaction terms relating to gender and race (***Female*** and ***indigenous_status)***: ***fulltime * Female, fulltime * indigenous_status***, and ***Female * indigenous_status***.

These interaction terms have important implications for policy assessment. For example, if the income discrepancy is largest for full-time indigenous female workers, then policies should be aimed primarily at indigenous females in the full-time employment sector, where the income inequality is most pronounced. Our final model is as follows:

$$w\left(\widehat{log\left(income\right)}\right) = w(\widehat{\beta_0} + \widehat{\beta_1}fulltime + \widehat{\beta_2}Female + \widehat{\beta_3}engprofNotWell + \widehat{\beta_4}engprofVeryWell + \widehat{\beta_5}engprofWell + \widehat{\beta_6}NonIndigenous + \widehat{\beta_7}age + \widehat{\beta_8}no\_years\_school + \widehat{\beta_9}Female \cdot fulltime + \widehat{\beta_{10}}NonIndigenous \cdot fulltime + \widehat{\beta_{11}}Female \cdot NonIndigenous)$$

$$Where: w = \frac{1}{no\_people}$$

	Dependent variable:
	log(income)
fulltime	0.911^{***}
	(0.058)
Female	0.054
	(0.054)
englishproficiencyNot well	0.149^{***}
	(0.052)
englishproficiencyVery well	0.809^{***}
	(0.078)
englishproficiencyWell	0.156^{**}
	(0.074)
indigenous_statusNon-Indigenous	0.547^{***}
	(0.075)
age	0.008^{***}
	(0.002)

fig.6

no_years_school	0.041^{**}
	(0.016)
fulltime:Female	-0.275^{***}
	(0.064)
fulltime:indigenous_statusNon-Indigenous	-0.265^{***}
	(0.067)
Female:indigenous_statusNon-Indigenous	-0.013
	(0.066)
Constant	4.705^{***}
	(0.198)
Observations	282
R^2	0.762
Adjusted R^2	0.753
Residual Std. Error	0.050 (df = 270)
F Statistic	78.761^{***} (df = 11; 270)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

fig.6 (continued)

Interpretation of Significant Coefficients and Hypothesis Tests

Intercept	For a 0 years old indigenous male with 0 years of schooling working part-time and who is not at all proficient in English, the mean income is $e^{4.705}$ ~ AUD 110.5. However, this is unrealistic as there is nobody with these characteristics.
$\hat{\beta}_{fulltime}$	Full-time workers have a mean income that is $(e^{0.911} - 1) \times 100$ ~ 148.6% higher than part-time workers, on average, ceteris paribus. With $p < 0.05$, there is sufficient evidence to conclude that full-timers have a higher mean income than part-timers. This is unsurprising as full-timers work more hours than part-timers and generally have more responsibilities as well.
$\hat{\beta}_{engprofNotWell}$	Individuals with 'Not Well' English proficiency have a mean income that is $(e^{0.149} - 1) \times 100$ ~ 16.07% higher than individuals with no English proficiency, on average, ceteris paribus. Since $p < 0.05$, there is sufficient evidence to conclude that individuals with 'Not Well' English proficiency have a higher mean income than individuals with no English proficiency, at $a = 0.05$.
$\hat{\beta}_{engprofVeryWell}$	People with 'Very Well' English proficiency have a mean income that is $(e^{0.809} - 1) \times 100$ ~ 124.57% higher than individuals with no English proficiency, on average, ceteris paribus. This result is statistically significant at the 5% level ($p < 0.05$), meaning there is sufficient evidence to conclude, at $a = 0.05$ that individuals with 'Very Well' English proficiency have a mean income that is higher than individuals with no English proficiency.
$\hat{\beta}_{engprofWell}$	People with 'Well' English proficiency have a mean income that is $(e^{0.156} - 1) \times 100$ ~ 16.88% higher than individuals with no English

	proficiency, on average, ceteris paribus. This result is statistically significant at the 5% level ($p < 0.05$), meaning there is sufficient evidence to conclude that individuals with 'Well' English proficiency have a mean income that is higher than individuals with no English proficiency.
$\hat{\beta}_{NonIndigenous}$	Non-indigenous workers have a mean income that is $(e^{0.547} - 1) \times 100 \sim 72.8\%$ higher than indigenous workers, on average, ceteris paribus. With $p < 0.05$, there is sufficient evidence to conclude at $\alpha = 0.05$ that non-indigenous part-time male workers have a higher mean income than indigenous part-time male workers.
$\hat{\beta}_{age}$	If the mean age increased by 1 year, then the mean income would increase by $(e^{0.008} - 1) \times 100 \sim 0.8\%$, on average, ceteris paribus. This result is statistically significant at the 5% level ($p < 0.05$). There is enough evidence to conclude that an increase in the mean of ages surveyed increases the mean income for that group, at $\alpha = 0.05$.
$\hat{\beta}_{no\_years\_school}$	If the mean no_years_school increased by 1 year, the mean income would increase by $(e^{0.041} - 1) \times 100 \sim 4.19\%$, on average, ceteris paribus. Since $p < 0.05$, there is enough evidence to conclude that an increase in mean number of years of schooling increases mean income, at $\alpha = 0.05$.
$\hat{\beta}_{Female*fulltime}$	Female full-time workers have a mean income that is $(e^{0.275} - 1) \times 100 \sim 31.65\%$ lower than male full-time workers, on average, ceteris paribus, after accounting for the main effects of gender and employment status. With $p < 0.05$, there is enough evidence to conclude, at $\alpha = 0.05$, that female full-timers have lower mean incomes than male full-time workers.
$\hat{\beta}_{NonIndigenous*fulltime}$	Non-indigenous full-time workers have a mean income that is $(e^{0.265} - 1) \times 100 \sim 30.34\%$ lower than indigenous full-time workers, on average, ceteris paribus, after accounting for the main effects of *fulltime* and *NonIndigenous*. Since $p < 0.05$, there is enough evidence to conclude, at $\alpha = 0.05$, that non-indigenous full-time workers have lower mean incomes than indigenous full-time workers.

We can summarise the income differences among races and gender, taking into account employment type (part-time or full-time) in the following tables:

Gender

Category	Log Mean Income (AUD)
Female * full-time	(4.7 - 0.275) + 0.054 + 0.911 = 5.39
Female * part-time	4.7 + 0.054 = 4.754
Male * full-time	4.7 + 0.911 = 5.611
Male * part-time	4.7

Race

Category	Log Mean Income (AUD)
Indigenous * full-time	4.7 + 0.911 = 5.611
Indigenous * part-time	4.7
NonIndigenous * full-time	4.7 + 0.911 + 0.547 - 0.265 = 5.893
NonIndigenous * part-time	4.7 + 0.547 = 5.247

In terms of model fit, both AIC and BIC values went down compared to the previous model **(See Appendix F)**. This means that all the changes we did, including the addition of the various interaction terms, proved to be good for model goodness-of-fit. In addition, our model shows little sign of multicollinearity, with no Generalised Variance Inflation Factor exceeding 2.5 **(See Appendix G).** Heteroskedasticity remains eliminated with an insignificant p-value for the model's Breusch-Pagan test. An interesting visualisation of fitted versus residual values of this model can be found in **Appendix H**.

One interesting observation from the data is the income differences among people of different English proficiency levels. People who had an English proficiency of "Not well" or "Well" earned 16.07% and 16.88% more than people who had no English proficiency at all (base level), respectively. However, people who had an English proficiency of "Very well"

The *Female* variable was insignificant, meaning there is insufficient evidence of a pay gap among the genders. However, when looking at full-time workers and at the combined effect of being Female and working full-time, through the coefficient of interaction term *Female * Fulltime,* there was empirical evidence of a gender pay gap among the genders. **Full-time working women earned, on average, much lower than full-time working men.**

In terms of race, non-indigenous workers have a higher mean income than indigenous workers, both for full-time and part-time workers. To see if indigenous women are even more marginalised based on being both female and indigenous, we added our last interaction term

$\hat{\beta}_{Female*NonIndigenous}$. However, its hypothesis test proved it to be insignificant, meaning there wasn’t sufficient evidence to conclude that this demographic (indigenous, female) face additional marginalisation in average income on top of the regular effects of being female or indigenous.

A full comparison of all three of our models can be found in **Appendix I**.

Part 3: Conclusion

Overall, the evaluation has provided nuanced results. After taking into account many confounding variables, the discrepancy in pay among races and genders tends to vary significantly depending on the work being full-time or part-time. It is clear though that among full-time workers, women earn less than men, and non-indigenous workers earn less than indigenous workers. There was also insufficient evidence of any extra marginalisation due to a person being both female and indigenous. Some of the hypothesis tests done in the model gave expected results (such as significant effect of number of years of schooling or full-time work on average income). Others, however, gave surprising results. Indigenous people are often thought of as marginalised (Gale & Mills, 2015, para. 2) which may also mean a lower pay. The empirical evidence showed this to be true for part-time work, however, for full-time work (which is more important) indigenous workers actually earned almost 25% more than their non-indigenous counterparts.

While our final model was pretty robust, there were still some limitations. From a prescriptive perspective, it may be inaccurate to globalise the results due to the nature of the data being Australian-centric. For example, Indigenous Australians are very specific to Australia, and the issues they may face in Australia may be vastly different than what other ethnic minorities might face in other countries. In addition, Australia is a well-developed first world country. In lesser developed countries, the interaction between race, gender, and income may vary.

Another limitation of our model is not taking into account other important variables, first one being job type/role/title. Job sector and title play a significant role in determining one's salary. By not holding these factors constant, causality is harder to determine. Another variable not taken into account is years of working experience. Work experience is another significant factor in determining one's income. In general, more years of experience comes with a higher income; an experience premium is placed as a result of their accumulated experience and seniority. One more variable not taken into account is intelligence. Intelligent people may bring more value to the workforce, potentially driving up their value and their income. Intelligence is hard to measure directly, so some metrics can be used as a proxy for intelligence, such as IQ. By not taking these two critical variables into account, the final model's ability to determine causality diminishes.

Other potential considerations that may have been taken into account are other interaction terms such as *no_years_school* * *Female*. It may have shown potential discrepancies between the effects of additional years of schooling between males and females. Perhaps females felt discouraged to pursue higher education due to the perceived lesser marginal benefit of an extra year of schooling for them compared to a male, ceteris paribus. These considerations would significantly aid the prescriptive power of our model.

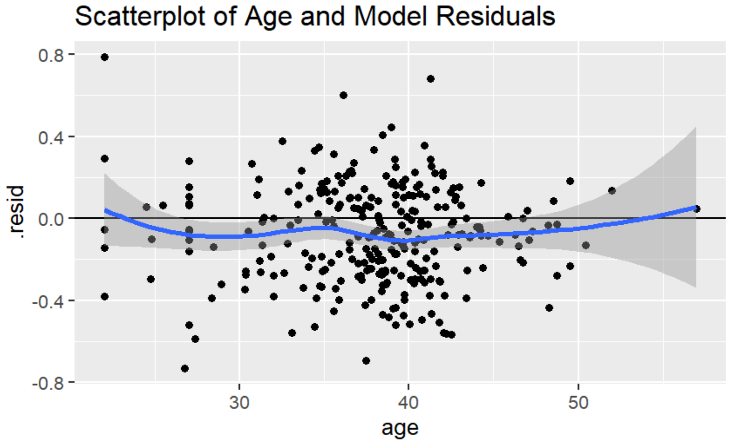
Appendix A

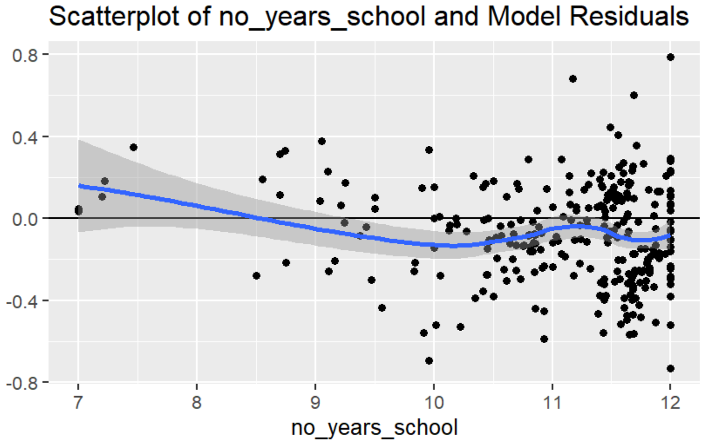
This was our model when we did model specification first before addressing HSK. We removed “no_people” because it represents the sample size for each observation. If we see it as an independent variable, it is like saying: “The more people are like me, the higher/lower my income will be.” which makes no sense. Instead, it is a major contributor to heteroskedasticity because variance empirically varies when sample size varies across observations. Thus, it should be the model’s weightage.

Summary of our Final Initial Model

	Dependent variable:
	log(income)
employstatusEmployed, worked part-time	-0.652^{***}
	(0.022)
englishproficiencyNot well	-0.713^{*}
	(0.384)
englishproficiencyVery well	-1.777^{***}
	(0.544)
englishproficiencyWell	-0.885^{**}
	(0.409)
sexMale	0.132^{***}
	(0.022)
indigenousNon-Indigenous	0.154^{***}
	(0.038)
age	0.014^{***}
	(0.002)
no_years_school	0.367^{**}
	(0.169)
log(no_years_school)	-3.393^{**}
	(1.648)
englishproficiencyNot well:no_years_school	0.071^{**}
	(0.036)
englishproficiencyVery well:no_years_school	0.199^{***}
	(0.049)
englishproficiencyWell:no_years_school	0.097^{**}
	(0.038)
Constant	9.925^{***}
	(2.114)
Observations	282
R^2	0.855
Adjusted R^2	0.849
Residual Std. Error	0.181 (df = 269)
F Statistic	132.622^{***} (df = 12; 269)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Appendix B





The scatterplot for age vs model residuals shows mild heteroskedasticity, with lower age values having more variance compared to older age values.

The scatterplot for no_years_school and model residuals shows major heteroskedasticity, with larger values of no_years_school showing significantly more variability compared to smaller

Appendix C

After model specification, we then addressed HTSK which led to this model.

Summary of our Final Model with Weighted Least Squares

	Dependent variable:
	log(income)
employstatusEmployed, worked part-time	-0.650^{***}
	(0.023)
englishproficiencyNot well	-0.661^{*}
	(0.392)
englishproficiencyVery well	-1.702^{***}
	(0.553)
englishproficiencyWell	-0.814^{**}
	(0.412)
sexMale	0.128^{***}
	(0.023)
indigenousNon-Indigenous	0.163^{***}
	(0.040)
age	0.014^{***}
	(0.003)
no_years_school	0.400^{**}
	(0.175)
log(no_years_school)	-3.672^{**}
	(1.702)
englishproficiencyNot well:no_years_school	0.065^{*}
	(0.037)
englishproficiencyVery well:no_years_school	0.191^{***}
	(0.049)
englishproficiencyWell:no_years_school	0.089^{**}
	(0.038)
Constant	10.232^{***}
	(2.178)
Observations	282
R^2	0.846
Adjusted R^2	0.839
Residual Std. Error	0.027 (df = 269)
F Statistic	123.100^{***} (df = 12; 269)
Note:	*p<0.1; **p<0.05; ***p<0.01

Appendix D

Linear Hypothesis Test of each Region in “m1.0”

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
273	0.7709256	NA	NA	NA	NA
259	0.5546187	14	0.2163069	7.215186	**1.277285e-12**

Appendix E

Bruesch-Pagan Test Results of All models (Weighted and Unweighted)

Model	p-value with weightage	p-value with no weightage
m1.0	1.0000	4.510e-09
m2.0	0.5482	9.277e-05
final_model	0.8795	8.894e-05

Appendix F

AIC & BIC of all Models

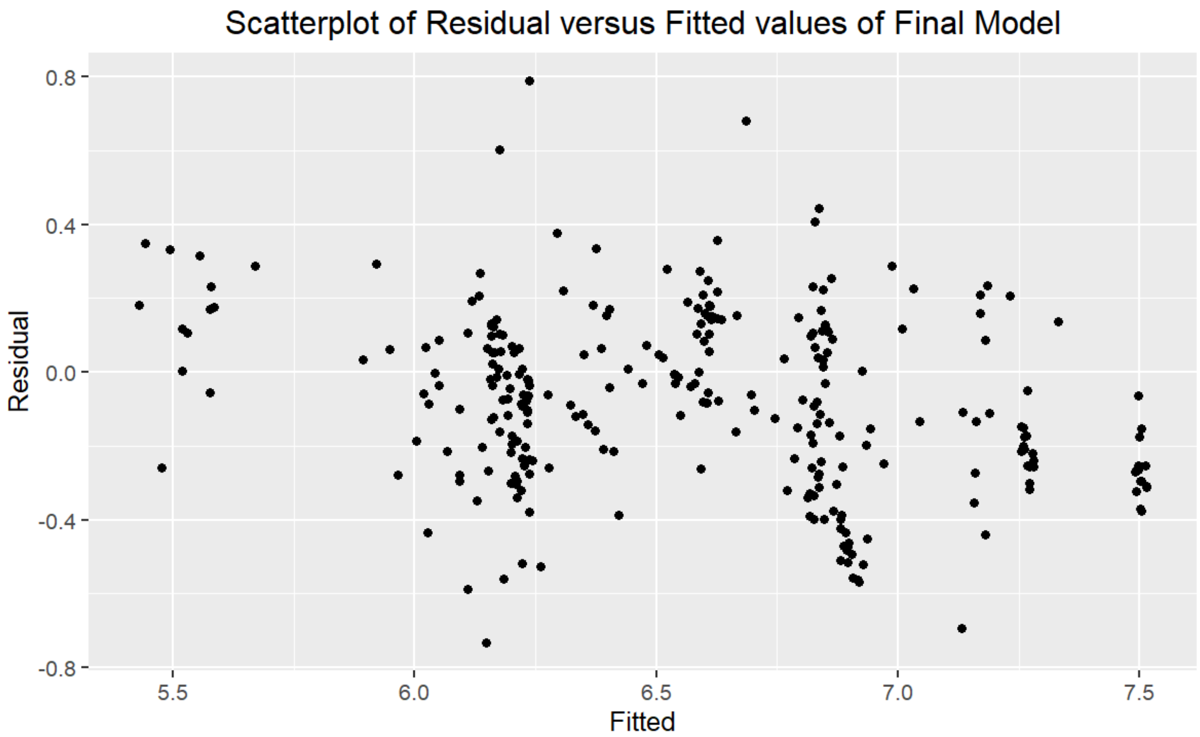
AIC and BIC of all Models		
Model	AIC	BIC
m1.0	699.1184	786.5241
m2.0	764.6401	804.7010
final_model	737.8012	785.1460

Appendix G

Final Model's Generalised Variance Inflation Factor for all Independent Variables

Variable	GVIF	Df	GVIF^(1/(2&Df))
fulltime	3.697517	1	1.922893
Female	3.177238	1	1.782481
englishproficiency	6.193229	3	1.355146
indigenous_status	6.125263	1	2.474927
age	1.829841	1	1.352716
no_years_school	2.110850	1	1.452876
fulltime:Female	2.772418	1	1.665058
fulltime:indigenous_status	3.592610	1	1.895418
Female:indigenous_status	2.911442	1	1.706295

Appendix H



Although the Breusch-Pagan test showed that we eliminated heteroskedasticity well before our final model, there still appears to be non-constant variance among residual values in our final model.

Appendix I

Comparison of all Three of Models

	Dependent variable:		
		log(income)	
	(1)	(2)	(3)
regionGreater Adelaide	0.024		
	(0.109)		
regionGreater Brisbane	-0.146		
	(0.109)		
regionGreater Darwin	0.314***		
	(0.103)		
regionGreater Hobart	-0.022		
	(0.126)		
regionGreater Melbourne	-0.175		
	(0.111)		
regionGreater Perth	-0.036		
	(0.120)		
regionGreater Sydney	0.195		
	(0.130)		
regionRest of NSW	0.040		
	(0.105)		
regionRest of NT	-0.048		
	(0.104)		
regionRest of Qld	0.041		
	(0.108)		
regionRest of SA	-0.064		
	(0.100)		
regionRest of Tas.	-0.016		
	(0.105)		
regionRest of Vic.	0.193*		
	(0.098)		
regionRest of WA	-0.202*		
	(0.111)		
urban		0.045	
		(0.039)	
fulltime	0.600***	0.652***	0.916***
	(0.037)	(0.037)	(0.058)
englishproficiencyNot well	0.211***	0.177***	0.130**
	(0.059)	(0.051)	(0.053)
englishproficiencyVery well	0.784***	0.821***	0.805***
	(0.089)	(0.086)	(0.078)
englishproficiencyWell	0.307***	0.256***	0.150**
	(0.074)	(0.073)	(0.074)
Female	-0.085***	-0.119***	-0.361
	(0.031)	(0.033)	(0.274)
indigenous_statusNon-Indigenous	0.335***	0.395***	0.572***
	(0.072)	(0.061)	(0.076)
age	0.013***	0.010***	0.008***
	(0.002)	(0.002)	(0.002)
no_years_school	0.074***	0.047***	0.025
	(0.016)	(0.017)	(0.019)
fulltime:Female			-0.265***
			(0.064)
fulltime:indigenous_statusNon-Indigenous			-0.286***
			(0.068)
Female:indigenous_statusNon-Indigenous			-0.022
			(0.066)
Female:no_years_school			0.038
			(0.024)
Constant	4.342***	4.682***	4.880***
	(0.234)	(0.211)	(0.227)
Observations	282	282	282
R^2	0.808	0.735	0.764
Adjusted R^2	0.792	0.726	0.754
Residual Std. Error	0.046 (df = 259)	0.053 (df = 272)	0.050 (df = 269)
F Statistic	49.672*** (df = 22; 259)	83.800*** (df = 9; 272)	72.770*** (df = 12; 269)
Note:			*p<0.1; **p<0.05; ***p<0.01

References

Brito, C. (2022, February 11). *Serena Williams says "it takes time" to address gender pay equality in the sports world*. CBS News. Retrieved September 17, 2024, from https://www.cbsnews.com/news/serena-williams-sports-gender-inequality/

Gale, T., & Mills, C. (2015, December 15). Creating Spaces in Higher Education for Marginalised Australians: Principles for Socially Inclusive Pedagogies. *Enhancing Learning in the Social Sciences*, *5*(2), 7-19. https://doi.org/10.11120/elss.2013.00008

Lang, K., & Spitzer, A. K.-L. (2020, Spring). Race Discrimination: An Economic Perspective. *Journal of Economic Perspectives*, *34*(2), 68-89. https://doi.org/10.1257/jep.34.2.68

United Nations. (2021, March 29). *UN News*. Gender equality, the ‘unfinished human rights struggle of this century’: UN chief. Retrieved September 17, 2024, from https://news.un.org/en/story/2021/03/1088512

United Nations. (2023, June 19). *OHCHR*. Taliban edicts suffocating women and girls in Afghanistan: UN experts. Retrieved September 17, 2024, from https://www.ohchr.org/en/press-releases/2023/06/taliban-edicts-suffocating-women-and-girls-afghanistan-un-experts