writing sample: Undergraduate Research Project

DRAFT: WORKING PROGRESS

Decomposing the Gap: Tracing the Sources of Indigenous Income Inequality

Indigenous-non-Indigenous income gap Analysis Using 2016 and 2021 Census Data

Kai Song

Abstract

The 'Closing the Gap' plan aims to reduce inequality between Aboriginal and Torres Strait Islander people and non-Indigenous Australians in terms of real-life outcomes. The plan targets key areas such as health, education, and employment. This report focuses on average weekly income as an indicator of socioeconomic opportunity, analysing the income gap between Indigenous and non-Indigenous Australians and tracing the sources of this inequality using publicly available Australian Census data in *TableBuilder*. This report highlights the possibility of income inequality analysis without individual-level data.

Keywords: Indigenous-non-Indigenous income gap, census data, employment, education, Australia

Contents

1 Introduction

To paint a more constructed picture of the Indigenous-non-Indigenous income gap, I organised the Australian Census data provided by TableBuilder on labour market outcomes and education by Indigenous and non-Indigenous status and will present it in this report. The data can be accessed free of charge and relatively quickly, implying the potential improvement in the efficiency of research in such areas.

One major drawback with such source usage is the lack of individual data, which raises the difficulty of Oaxaca's decomposition on inequality. This report provided a counter approach: that is, assume the general function form without considering the regression form between variables, then use the *shift-share analysis* to estimate the size of each decomposed module, the *shift-share analysis* here refers to giving the Indigenous people the property of interest of the non-Indigenous people, such as income, education distribution, employment distribution, etc.

In section 2, I will provide a general background of the study. I will discuss the historical context of Australian *"Closing the Gap"* strategies and its focus in subsection 2.1, then I will summarise the previous work related to the Indigenous community income in subsection 2.2. The data will be described in section 3 to provide an overview of the structure and define each variable and value. After providing an overview of the income gap in section 4, I will formally describe how I approach this Oaxaca's decomposition without regression in section 5, by deducting the theoretical result for this decomposition in subsection 5.1, and how to measure each component in the decomposition form using shift-share analysis in subsection 5.2. In section 6, we address the existing gap in education and employment between the Indigenous and non-Indigenous people, and section 7 provides the result of the shift-share analysis described before. In section 8, I compare the decomposed result in 2016 with 2021 with similar data, aiming to describe the change in the source of inequality within 5 years. The later sections, **??** and **??** address some other concerns that are not included in the discussion, such as heterogeneity in the age distribution raised by self-identification of Indigenous, and the unique story about Indigenous women.

2 Background

2.1 Context

> *"The time has now come for the nation to turn a new page in Australia's history by righting the wrongs of the past and so moving forward with confidence to the future."*
>
> – Kevin Rudd, *The Apology*, 2008

Indigenous Australians' history of pursuing equality is a long and ongoing journey. The focus of the Indigenous equal rights movement has evolved over time, expanding from equal land and water rights to equality in other real-life outcomes. However, Indigenous Australians continue to face poorer labour market outcomes compared to non-Indigenous Australians, contributing to socioeconomic disadvantage, as well as challenges in public health, education, and access to institutions. The reasons behind it are tangled and complicated, including issues with geographical distribution, parental education, discrimination, and culture.

On 13 February 2008, the Australian Prime Minister at that time, Kevin Rudd made a formal apology to the 'stolen generation': the Indigenous children who were removed from their families by governmental or church agents. *The Apology* is a historical moment in Australia, not only it is when the mistakes are acknowledged, but also it brings the discussion of equality from the realm of rights and legislation to the lived experiences of individuals.

One year after *the Apology*, the first annual *Closing the Gap* report, carried out by National Indigenous Australians Agency (2009) (NIAA) was released, signalling the beginning of the national effort of *Close the Gap in real-life outcomes* between Indigenous and non-Indigenous Australians. From then, a report was carried out annually by the National Indigenous Australians Agency and delivered by the prime minister at that time. The first *Closing the gap* report set the following 6 targets:

- close the life expectancy gap within a generation
- halve the gap in mortality rates for Indigenous children under five within a decade
- ensure access to early childhood education for all Indigenous four years olds in remote communities within five years
- halve the gap in reading, writing and numeracy achievements for children within a decade
- halve the gap for Indigenous students in year 12 attainment or equivalent attainment rates by 2020, and
- halve the gap in employment outcomes between Indigenous and non-Indigenous Australians within a decade.

Achieving the target required halving the gap in employment outcomes by 2018, which was not met. The gap in year 12 or equivalent attainment was successfully halved by 2020. The fact that human capital increased via education failed to be revealed in employment outcomes has not been looked at closely.

Employment outcomes of an individual are revealed in the labour market, emphasising individuals' engagement and development in employment or actions to do so. Employment rate and full-time employment rate in data can capture such outcomes. However, relying on them only cannot capture the value endowed to employment outcomes, meaning it is harder to compare or aggregate employment outcomes across individuals. For such reason, it is crucial to introduce a universal variable to represent the value of

the labour market engagement, combined with the employment rate to represent employment outcome fully. This report uses average weekly income, the methods and reasoning will be discussed in the later sections.

2.2 Previous Work

Well-structured discussions on income differences between Indigenous and non-Indigenous Australians are limited by data availability and the ethics approval process. Consequently, there are not many studies accessible to the public or academics.

Biddle (2024) took a human capital approach and measured the outcome of education by lifetime income. This work systematically discussed the impact of education. A wide range of variables are discussed besides income, education and employment; including the field of study, participation in early education, student status maintenance and population growth. However, in the latest version, the employment variables only included employment and unemployment and did not consider full-time or part-time employment. He reestimates the socioeconomically underprivileged of the Indigenous people regarding low human capital, and high potential income increment by obtaining education in the Indigenous community, and the possibility of obtaining such results using only the census data. Biddle (2024)'s work has largely incentivised this work and the author has undertaken a similar classification system when defining the values in education variables, the chosen age group, presenting strategies etc.

One interesting thing Biddle (2024) pointed out is that there are fundamental differences between Indigenous and non-Indigenous research on human capital. While researchers treat education as a way to obtain human capital for non-Indigenous communities, which serves as a means of potential lifelong income and welfare improvement, the education obtained by Indigenous communities is considered as a means to preserve the culture, tradition, or to give back to the community for altruistic reasons. Such differences regarding treating education acquisition are reflected in the lack of studies on the Indigenous people's income or human capital, which emphasise the importance of Indigenous income or human capital studies.

Several studies have examined Indigenous income, primarily utilising micro-data, such as the Household, Income and Labour Dynamics in Australia (HILDA) survey. Using 2010 data, Howlett et al. (2015) discussed differences in income sources comparing Indigenous to non-Indigenous people, then found the main source of the Indigenous-non-Indigenous income gap is under those working full time, which can be interpreted as the lack of human capital accumulation or discrimination (Biddle et al., 2013). Birch and Marshall (2017) focused on the workers and discussed the income of the Indigenous people. The research used 2011 data accessed by ABS's Remote Access Data Library (RADL), suggesting little reduction in income inequality between Indigenous and non-Indigenous people from 2001 to 2011. It also indicates that more reasons are contributing to such disparity than discrimination regarding pay rates. One leading factor could be the low full-time employment rate in the Indigenous communities.

The work mentioned above emphasises the importance of employment and education in the discussion on income. Kalb et al. (2012)'s work used individual-level information, looked at non-Indigenous Australians based on HILDA wave 8 data, Indigenous people based on NATSISS 2008 data due to the under-sampling of Indigenous people in HILDA. Their findings show that lower education levels, larger family sizes and worse public health situations contribute to worse employment levels for Indigenous people, such as labour force participation and employment rate, and full-time employment rate.

3 Data

All data for this report is sourced from TableBuilder provided by the Australian Bureau of Statistics (2016, 2021) (ABS). By selecting variables as criteria, TableBuilder provides the number of people who meet those criteria. The following variables, along with their corresponding names in the original file were used as criteria:

- Sex: *'SEXP'*
- Age: *'AGE5P'*
- Indigenous Status: *'INGP'*
- Employment: *'LFSP'* (2021) and *'LFHRP'* (2016, due to the variable name update)
- Education: *'HSCP'* and *'HEAP'*
- Weekly Income: *'INCP'*

3.1 Variables and Values

In this section, I will outline the detailed information of each variable and the implication of each value.

3.1.1 Sex

Male or *Female*, exclude the unstated.

3.1.2 Age

Grouped by 5 years, from *15 to 19 years old* to *60 to 64 years old* and *65 years old and above*. However, when the discussion includes employment, we only look at those aged between 25 and 64, as they are most likely finished with education and in the labour force.

3.1.3 Indigenous Status

Non-indigenous and *Indigenous*, which includes those identified as Aboriginal people, Torres Strait Islander people, both Aboriginal and Torres Strait Islander people. Exclude unstated. What is worth highlighting is Indigenous status is a self-reporting status, therefore one identified as Indigenous in the previous census might not identify as Indigenous in the later census.

3.1.4 Education

The education variable referred to the framework Biddle (2016) provided, which is generated by 2 variables: HSCP (*Highest school year completed*), and HEAP (*Level of highest educational attainment*). HSCP can take on values on:

- *"Finish year 12 or equivalent"*; or
- *"Not finished year 12 or equivalent"*.

HEAP can take on 3 values:

- *"No post-school qualification"*,
- *"Have post-school qualification"*, or
- *"Have a Bachelor's degree or above"*.

The Detail of the HEAP variable categorisation is as follows:

Value	Description
No post-school qualification	Do not have any educational qualification (excluding school), or only have Cert I or II.
Have post-school qualification	Have one of the education qualifications in Cert III or IV, Associate Degree, Advanced Diploma, Diploma, Associate Diploma, Advanced Certificate, and do not have a degree above Bachelor degree.
Have Bachelor degree or above	Have one of the following qualifications: Doctorate, Master Degree, Graduate Diploma, Graduate Certificate, Bachelor Degree with Honours, Bachelor Degree

Table 1: HEAP classification definition

Using the re-categorised variables from HEAP and HSCP, we can now construct the education variables that take the value of *"No year 12, no post-school qual"*, *"No year 12, with post-school qual"*, *"With year 12, no post-school qual"*, *"With year 12, with post-school qual"*, or *"Bachelor degree or above"*. The table below outlines the details: the first column corresponds to the values from HEAP, the first row represents the values from HSCP, and the cells display the resulting value of the *education* variable for each combination of HSCP and HEAP.

	No year 12 or equivalent	*Year 12 or equivalent*
No post-school qualification	No year 12, no post-school qual	With year 12, no post-school qual
Have post-school qualification	No year 12, with post-school qual	With year 12, with post-school qual
Have Bachelor degree or above	Bachelor Degree or above	

Table 2: The definition for the values of *Education* variable

Those who only have a Certificate I or II will be considered to have "no post-school qualification", even though Certificate I or II is not a school qualification. ABS provided a ranked list of education (see appendix), in such a list, Certificate I or II is considered under the education of Year 10. That implies those who have only obtained Cert I or II have a lower level of education than those who have finished year 10. As a result, *the highest education attainment* of those finished year 12, and have a certificate II in year 12 instead of Certificate II (see appendix).

Certificates III and IV have a unique position in the Australian Education system. In ample contexts, Certificate III or IV are considered "equivalent" to year 12. However, the way ABS framed the question suggested the education level should include the "school level" education obtained in other institutions (such as TAFE). Regardless of the intention of the survey designer, those who file the survey only have the information on "school-level education", and might interpret the term themselves. A trade or culinary course is less likely to be interpreted as a "school level" qualification than the HSC courses provided by TAFE. Here in this report, we assume the "Year 12 or equivalent" value truly reflects the attainment of school-level education and does not include course completion purely on vocational education training (blue-collar industrial working, culinary etc.).

3.1.5 Employment

Employment captured the labour market outcomes one week before the census to determine whether the individual is employed, and the job-seeking action 4 weeks before the census to determine whether the individual is in the labour force. There are minor changes in terms of classification between the 2016 and 2021 censuses. Generally, we can break down the population who declare their employment status into 3 categories: employed, unemployed, and not in the labour force. In 2022, the detailed classification goes as follows, where the variable stared (*) indicated such value is included in my analysis:

Employed	Unemployed*	Not Looking to work
*Working full-time**	*Looking to work full time*	*Not in the labour force**
*Working part-time**	*Looking to work part time*	
*Away from work**		

Table 3: Employment definition for 2021 census

In 2016, the classification goes as follows:

Employed	Unemployed*	Not Looking to work
*Working full-time**	*Looking to work full time*	*Not in the labour force**
*Working part-time**	*Looking to work part time*	
*Away from work**		
*Not stated**		

Table 4: Employment definition for 2016 census

While full-time is defined as working for more than 35 hours per week in all jobs combined, working part-time is defined as working less than 35 hours for all the jobs combined. Using these values, we can compute the following variables if we treated there as 0 employed but hour non-stated people in 2021:

$$Employed = Working\ full\ time + Working\ part\ time + Away\ from\ work + Not\ stated$$

$$Employed\ Percentage = \frac{Employed}{Labour\ force + Not\ in\ Labour\ Force} = \frac{Employed}{Population}$$

$$Full\ time\ employed\ percentage = \frac{Working\ full\ time}{Population}$$

Full-time versus part-time comparison is a non-trivial part when discussing education outcome, or income more specifically. According to the 2021 census, full-time employed workers aged between 25 to 64 make an average of $1947.184 per week, while part-time employed workers make an average of $1062.984 per week, 54% less than full-time employed individuals as demonstrated in Table 5.

	Sex	Indigenous or not	Average income	Full-time income	Part-time income	Full time/Part time
1	Male	Indigenous	1556.92	1709.19	1025.69	1.67
2	Male	Non-Indigenous	1875.97	2051.33	1208.41	1.70
3	Female	Indigenous	1251.52	1572.16	876.38	1.79
4	Female	Non-Indigenous	1427.89	1798.53	1001.93	1.80

Table 5: Full-time/ Part-time Income comparison for different population

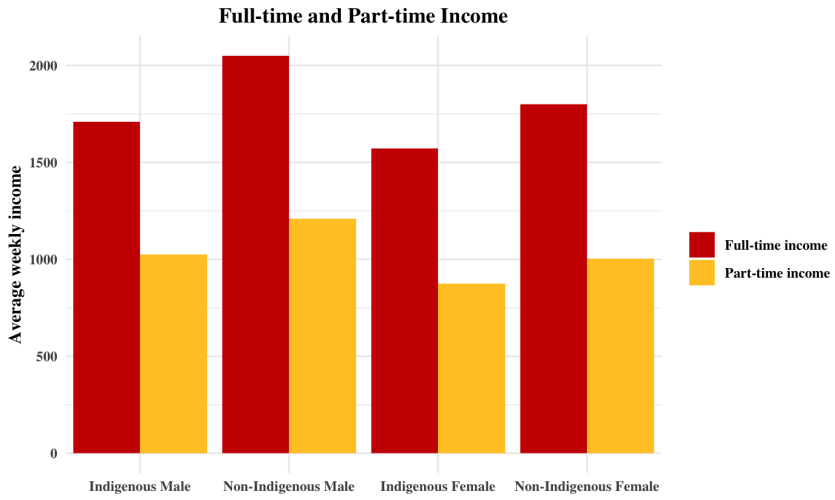

Figure 1: Full-time/ Part-time Income comparison for different population

3.1.6 Weekly Income

ABS does not collect individual detailed income in the census, though it has almost been a common practice to link census data with taxation data to obtain individual-level income for data sets like HILDA. This report uses the categorical income variable to estimate the average weekly income of a given population subset using the middle Riemann sum approximation.

Say there are m income intervals. For any population subgroup, say there are n_i people who fall into the income interval $\$a_{i-1}$ and $\$a_i$. We assume all these n_i people have the income $\$\frac{a_{i-1}+a_i}{2}$. For the m^{th} income interval, there is no upper-bound (e.g. \$3000 and above), I picked the estimator $\hat{a}$, such that $\hat{a} - a_{m-1} = a_{m-1} - a_{m-2}$. Hence, for a subset of the population, we can define the following:

$$Total\ income = \$ \sum_{i=1}^{m-1} n_i * \frac{a_{i-1} + a_i}{2} + n_m * (2a_{m-1} - a_{m-2})$$

$$Average\ income = \frac{Total\ income}{\sum_{i=1}^{m} n_i}$$

The income intervals are generally large due to privacy concerns. Hence such an estimate is not ideal and is flawed regarding the reflectivity on reality. However, this is the best estimator feasible without individual-level data.

The average weekly income does not necessarily represent human capital as it comes from different sources, including wage and salary, pension and benefits, and interest and profit etc. For those who are employed, the income is generated largely from wages or profits. In this case, income captures the marginal labour productivity and therefore captures the human capital for individuals obtained by education and training. For those unemployed or out of the labour force, pension and welfare payments are the major components of their income, and they do not necessarily capture human capital.

3.2 Data Structure

I will use the *'Sex', 'Age', 'Indigenous Status', 'Employment', 'Education'* as identification variables. These variables split the population into different subsets. For each subset, I used *TableBulider* to ob-

tain the number of people that fit in such a description and their *average income*. The final data frame contains 1100 rows, each row is a unique combination of the values of each variable, which leads to 2*(gender)* $\times 2$*(Indigenous status)* $\times 11$*(age group)* $\times 5$*(employment)* $\times 5$*(education)* $= 1100$ rows. Besides identification variable columns, the data frame has 2 more columns: *"Population"* and *"Average weekly income"*, leading to 7 columns in total.

This data structure is efficient for obtaining purposes, as it is similar to the data TableBuilder provided. This is an ideal data structure without individual-level data as it contains sufficient information to compare the variables of interest. The limitation of this type of data could be it assumes a high level of homogeneity within each type of identity created by identification variables.

ABS randomly adjusted the non-zero value in each cell of the TableBuilder outcome to prevent locating specific individuals in aggregated data. This may lead to data inconsistency, such as the value in each cell does not add up to the total amount in the TableBuilder outcome.
The impact of random perturbation is particularly non-trivial for smaller cells, which will be the case for the studies on Indigenous Australians considering how small the population is. According to the 2021 census, there are 547188 Indigenous people aged 15 or over who responded to the census, while 19215580 non-Indigenous people aged 15 or over responded to the census. Indigenous people made up 2.77% of the sample observed by the census.

	Sex	Indigenous or not	Population	Percentage
1	Male	Indigenous	267138	1.35%
2	Male	Non-Indigenous	9361512	47.37%
3	Female	Indigenous	280050	1.42%
4	Female	Non-Indigenous	9854068	49.86%

Table 6: Population and percentage of Indigenous and non-Indigenous people by gender

The strategy to minimise the impact of such perturbation is to introduce as few variables as possible for each analysis. For example, when analysing the relationship between education and employment rate, I did not include the *"average income"* variable. When providing a general picture of income inequality, I did not include the *"employment"* or *"education"* variable. I did include *"Age", "Sex", "Indigenous Status"* in all the data frames, while did not include some of the *"Employment", "Education" or "Average Weekly Income"* for some. Hence generated $2^3 = 8$ data frames.

4 The Income Gap: an Overview

In this report, I quantify the income gap using the *Income Ratio*, δ, which represents the ratio of average Indigenous income, Y_i, to the average non-Indigenous income, Y_n. This ratio is calculated as $\delta = \frac{Y_i}{Y_n}$. In cases where $\delta > 1$, Indigenous individuals on average earn more than their non-Indigenous counterparts. However, more commonly, $0 < \delta < 1$, indicates that Indigenous people earn less on average. The closer δ is to zero, the greater the income disparity, and hence, the larger the income gap.

Using the data described in the previous section, the average income for all Indigenous non-Indigenous people grouped by gender and age groups was obtained in Table 7 and visualised in Figure 2. Without looking at any classification on education level and labour market status, the income differences between Indigenous and non-Indigenous people were revealed largely in the 25 to 64 ages. For those between 25 to 64 years old, Indigenous people make on average $ 1043.42 weekly, while non-Indigenous people make 1479.83 $ weekly. An average Indigenous person earns 70% of an average non-Indigenous person.

	Identity	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	1047.61	1208.28	1282.62	1263.86	1252.64	1190.77	1106.39	974.30
2	Non-Indigenous Male	1316.90	1654.54	1853.28	1926.63	1914.84	1836.75	1687.77	1428.23
3	Indigenous Female	860.80	931.39	967.77	1006.35	1023.99	972.02	902.80	773.11
4	Non-Indigenous Female	1137.14	1285.20	1348.78	1400.62	1398.86	1329.71	1198.92	1001.38

Table 7: Average weekly income grouped by gender, age and Indigenous status (25 to 64 years old), 2021 cencus

	Identity	Income
1	Indigenous Male	1168.77
2	Non-Indigenous Male	1704.07
3	Indigenous Female	933.08
4	Non-Indigenous Female	1266.89
3	Indigenous	1043.42
4	Non-Indigenous	1479.83

Identity	Income Ratio δ
Male	0.68
Female	0.74
Total	0.70

Table 8: Income and income ratio δ for those aged between 25 to 64

	Income Ratio	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Male	0.80	0.73	0.69	0.66	0.65	0.65	0.66	0.68
2	Female	0.76	0.72	0.72	0.72	0.73	0.73	0.75	0.77

Table 9: Income ratio grouped by age group and gender

The Indigenous non-Indigenous income gap is different between males and females. From Table 7, we can calculate the income ratio for males and females separately for any age group. As shown in Table 9 and Figure 3, from the age group 35 to 39 years old, the Indigenous male is faced with a larger income gap (lower income ratio) than Indigenous female. As shown in Table 16, the average weekly income revived by Indigenous males is 68% of such compared to their non-Indigenous counterparts, while such ratio is 74% for the females.

Indigenous and non-Indigenous people have similar gender distribution (48.82% of Indigenous are male, and 48.72% of non-Indigenous are male), therefore the Indigenous-non-Indigenous income gap was po-

tentially not contributed by differences in gender. However, it is important to not combine Indigenous males and females when discussing inequality due to the differences in inequality magnitude.

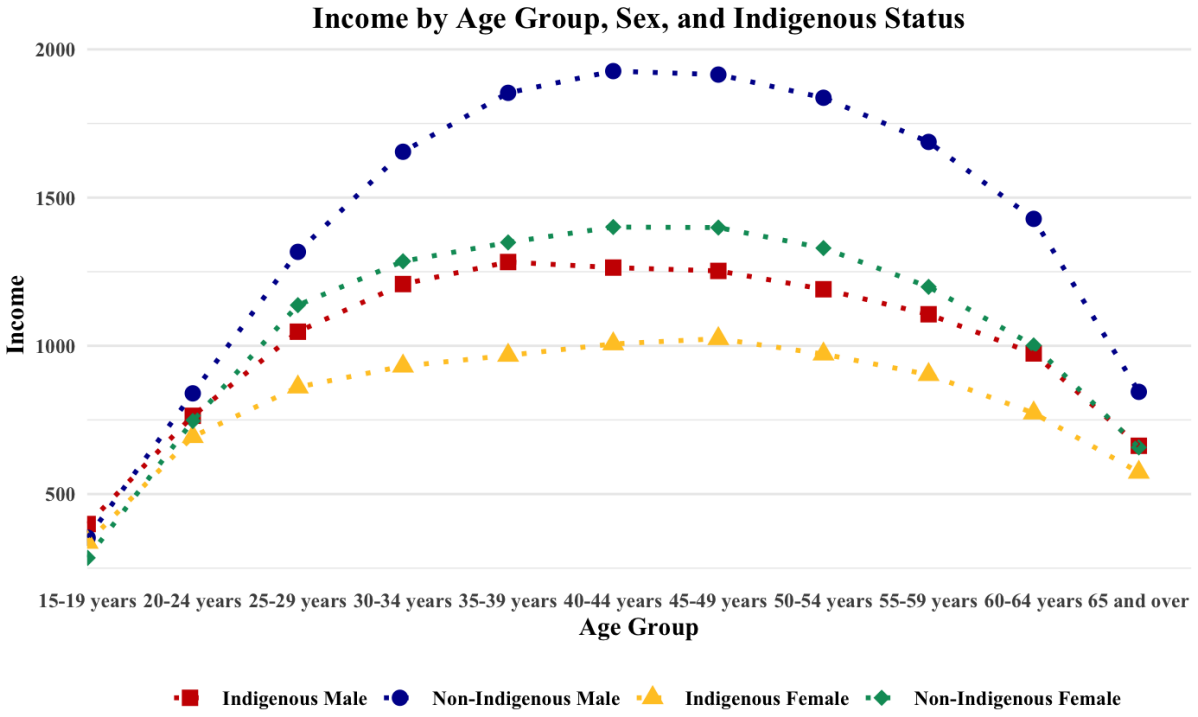

Figure 2: The average weekly income (2021)

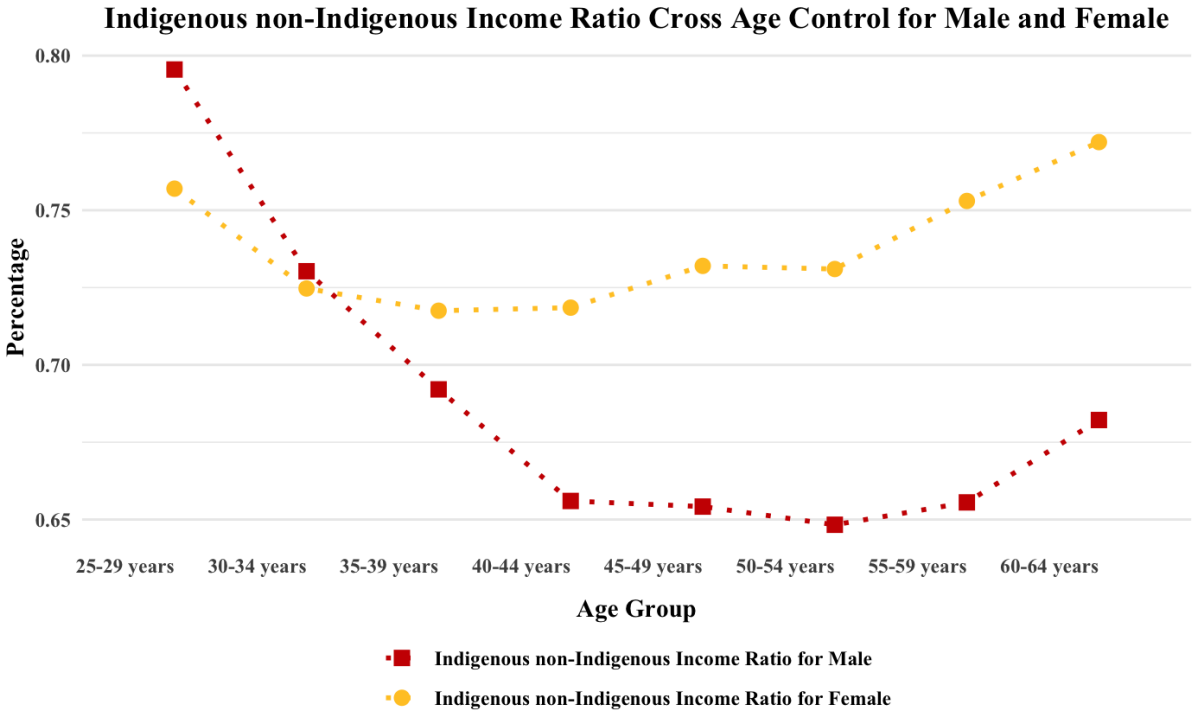

Figure 3: Indigenous non-Indigenous income for male and female

5 Methods

This section outlines the general method of analysing and decomposing the income gap between Indigenous and Non-Indigenous people with the data described before.

5.1 Theoretical Motivation

The methods adopted here is an adapted version of Oaxaca (1973) decomposition without regression, and with multiple variables. We first examine three variables:

- income (y), defined on real numbers $\mathbb{R}$;
- education level (edu), defined on EDU; and
- employment status ($empl$), defined on $EMPL$.

We assume *EDU, EMPL* are 2 complete continuums, with a partial order defined in each. We view income as a function of both education and employment status and employment status as a function of education, so we get the following set of equations.

$$\begin{aligned} y &:= F(edu, empl, X) \\ empl &:= g(edu, X) \\ \text{Hence: } y &= F(edu, g(edu, X), X) := f(edu, X) \end{aligned}$$

Here X indicates all other control factors. In this context, it is *age, gender, and Indigenous status*. We assume $\frac{\partial X}{\partial edu} = \frac{\partial edu}{\partial X} = \frac{\partial X}{\partial empl} = \frac{\partial empl}{\partial X} = 0$ and functions F, f, g are at least once differentiable regarding any independent variable. I address F as the *income generating function* and g as the *employment generating function* in the later text.

Consider a subset of the population whose income is systematically lower than their counterpart, in this case, Indigenous people. To improve their income, consider changing their education level, or employment level. The partial derivative of y over edu or $empl$ captures the direction and magnitude of the marginal improvement in income y as an outcome of changing education level or employment status. We obtain the following expressions.

$$\begin{aligned} \frac{\partial y}{\partial edu} = \frac{\partial F(edu, empl, X)}{\partial edu} &= \frac{\partial F}{\partial edu} \cdot \frac{\partial edu}{\partial edu} + \frac{\partial F}{\partial empl} \cdot \frac{\partial empl}{\partial edu} \\ &= \frac{\partial F}{\partial edu} + \frac{\partial F}{\partial empl} \cdot \frac{\delta empl}{\delta edu} \\ &:= F_{edu} + F_{empl} \cdot g_{edu} \end{aligned}$$

$$\begin{aligned} \frac{\partial y}{\partial empl} = \frac{\partial F(edu, empl, X)}{\partial empl} &= \frac{\partial F}{\partial empl} \cdot \frac{\partial empl}{\partial empl} + \frac{\partial F}{\partial edu} \cdot \frac{\partial edu}{\partial empl} \\ &= \frac{\partial F}{\partial empl} + \frac{\partial F}{\partial edu} \cdot \frac{\delta edu}{\delta empl} \\ &= \frac{\partial F}{\partial empl} \\ &:= F_{empl} \end{aligned}$$

The term $g'(edu)$ is defined as $\frac{\partial empl}{\partial edu}$, which measures the change in employment status for the marginal change in education level. We assume that $\frac{\partial edu}{\partial empl} = 0$, which implies by changing one's employment

status, their education level will not automatically change, which eliminates the cases for someone who was offered a position, and then went to an education institute to up-skill themselves.

Consider a representative individual in each of the Indigenous and non-indigenous communities, namely *i (Indigenous)* and *n (non-Indigenous)*. The *income gap* can be indicated by $y_n - y_i$. We can rewrite it into the following:

$$\begin{aligned} gap &= y_n - y_i \\ &= F_n(edu_n, empl_n, x) - F_i(edu_i, empl_i, x) \\ empl_n &= g_n(edu_n, x) \text{ and} \\ empl_i &= g_i(edu_i, x) \end{aligned}$$

Hence:

$$gap = F_n(edu_n, g_n(edu_n, x), x) - F_i(edu_i, g_i(edu_i, x), x)$$

Here x is all the control factors except for Indigenous status, in this context, it is *age and gender*. Furthermore, edu_n or edu_i indicates the education level of non-Indigenous and Indigenous people, defined in the same domain. $empl_n$ or $empl_i$ indicates the employment status of non-Indigenous and Indigenous people, defined on the same domain as well. To allow flexibility, I defined different income and employment-generating functions across Indigenous and non-Indigenous people, however, their differences should be demonstrated by empirical data.

If $F_i(\cdot, \cdot, x) = F_n(\cdot, \cdot, x)$, then conditioned on control factors, Indigenous and non-Indigenous people will yield identical income if they have identical combination of $(edu', empl'), for\ all\ edu' \in EDU,\ empl' \in EMPL$. If $g_i(\cdot, x) = g_n(\cdot, x)$, then conditioned on control factors, Indigenous and non-Indigenous people will yield identical employment status if they have identical edu', *for all* $edu' \in EDU$. The primes of such propositions are flawed, and this report will demonstrate their inaccuracies with empirical evidence from 2021 in the later analysis. Data reveals an income gap for those with the same $(edu, empl)$ combination, and a much more significant employment gap for those with identical edu.

For any control factor x, rewrite the variables regarding non-Indigenous people into the following:

$$\begin{aligned} edu_n &= edu_i + \delta_{edu} \\ empl_n &= empl_i + \delta_{empl} \end{aligned}$$

Expand the first term of *gap* by Taylor Expansion up to the first term. For each control factor x, we have:

$$\begin{aligned} F_n(edu_n, empl_n) &= F_n(edu_i + \delta_{edu}, empl_i + \delta_{empl}) \\ &\approx F_n(edu_i, empl_i) + \delta_{edu} \cdot \frac{\partial F_n}{\partial edu} + \delta_{empl} \cdot \frac{\partial F_n}{\partial empl} \\ &= F_n(edu_i, empl_i) + \delta_{edu} \cdot (F^n_{edu} + F^n_{empl} \cdot g^n_{edu}) + \delta_{empl} \cdot F^n_{empl} \\ &\textit{Hence:} \\ gap &= F_n(edu_i, empl_i) - F_i(edu_i, empl_i) + \delta_{edu} \cdot (F^n_{edu} + F^n_{empl} \cdot g^n_{edu}) + \delta_{empl} \cdot F^n_{empl} \\ &= [F_n(edu_i, empl_i) - F_i(edu_i, empl_i)] + \delta_{edu} \cdot F^n_{edu} + \delta_{empl} \cdot F^n_{empl} + \delta_{edu} \cdot g^n_{edu} \cdot F^n_{empl} \end{aligned}$$

The final equation provides a theoretical decompose of the Indigenous income gap by splitting the pay gap into 4 components. The first term $F_n(edu_i, empl_i) - F_i(edu_i, empl_i)$, namely *income function gap* describes the income differences that cannot be explained by education or employment. The rest of the terms are hard to examine due to the endogeneity within education and employment variables, hence I performed the later rearrangement.

We can expend the δ_{empl} in the third term $\delta_{empl} * F^n_{empl}$ into $g_n(edu_i + \delta_{edu}) - g_i(edu_i)$ since:

$$\begin{aligned} g_n(edu_n) &= g_n(edu_i + \delta_{edu}) \\ \textit{and: } g_n(edu_n) &= g_i(edu_i) + \delta_{empl} \end{aligned}$$

Then we can rewrite the *gap* as:

$$\begin{aligned} gap \approx\ & \textit{Income function gap} + \delta_{edu} \cdot F^n_{edu} + \delta_{edu} \cdot g^n_{edu} \cdot F^n_{empl} + \\ & (g_n(edu_i + \delta_{edu}) - g_i(edu_i)) \cdot F^n_{empl} \\ =\ & \textit{Income function gap} + \delta_{edu} \cdot F^n_{edu} + \\ & F^n_{empl} \cdot (\delta_{edu} \cdot g^n_{edu} + (g_n(edu_i + \delta_{edu}) - g_i(edu_i))) \end{aligned}$$

Consider the first-order Taylor expansion of function g_n:

$$\begin{aligned} g_n(edu_i + \delta_{edu}) &\approx g_n(edu_i) + \delta_{edu} * \frac{\partial g_n}{\partial edu} \\ &:= g_n(edu_i) + \delta_{edu} * g^n_{edu} \\ \textit{Hence:}& \\ g_n(edu_i + \delta_{edu}) &\approx \delta_{edu} * g^n_{edu} + g_n(edu_i) \end{aligned}$$

Now we can rewrite the *gap* into:

$$\begin{aligned} gap =\ & \textit{Income function gap} + \delta_{edu} \cdot F^n_{edu} + \\ & F^n_{empl} \cdot (2\delta_{edu} \cdot g^n_{edu} + (g_n(edu_i) - g_i(edu_i))) \\ =\ & \textit{Income function gap} + \delta_{edu} F^n_{edu} + F^n_{empl} \delta_{edu} g^n_{edu} + \\ & F^n_{empl}[g_n(edu_i) - g_i(edu_i)] + F^n_{empl} \delta_{edu} g^n_{edu} \\ =\ & \textit{Income function gap} + \delta_{edu} \frac{\partial F}{\partial edu} + \delta_{edu} g^n_{edu} \frac{\partial F}{\partial empl} + [g_n(edu_i) - g_i(edu_i)] \frac{\partial F}{\partial empl} \end{aligned}$$

The second term is the first-order approximation of the *education resulted income gap*, where the term δ_{edu} directly measures the *education gap*, and $\frac{\partial F_n}{\partial edu}$, estimate the change in income for non-Indigenous people yield by the marginal change in education.

The third and fourth terms both measure the income gap generated by employment as a result of education differences, while $\delta_{edu} g^n_{edu}$ and $[g_n(edu_i) - g_i(edu_i)]$ measure the difference in employment, $\frac{\partial F}{\partial empl}$ yields such differences to the income gap. To demonstrate the difference, assume the relationship between education and employment (the function g) is linear and monotonically increasing, demonstrated in Figure 4, while the magnitude of each component needs to be supported by empirical evidence.

This is indeed the case as $\delta_{edu} g^n_{edu}$ measures the employment gap yield by education gap, using the em-

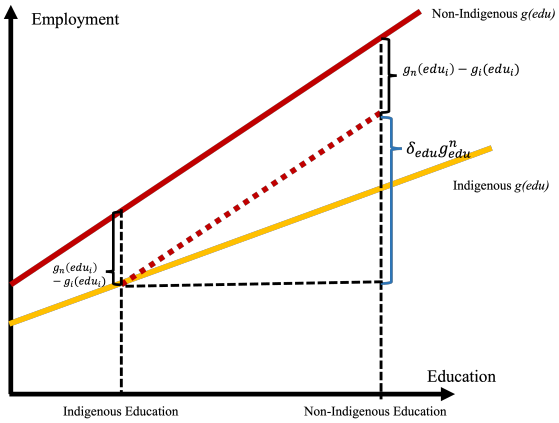

Figure 4: Differences between $\delta_{edu} g^n_{edu}$ and $[g_n(edu_i) - g_i(edu_i)]$

ployment generating function of non-Indigenous people, and $g_n(edu_i) - g_i(edu_i)$ measure the employment level differences for Indigenous people and non-Indigenous people at the Indigenous people's education level. Such decomposition breaks down the employment level gap into the *differences in employment generating function* and *difference due to education level*.

5.2 Shift Share Analysis

The previous decomposition holds without specifying the functional form of *income generating function* and *employment generating function*. To decompose the gap by measuring each component, an alternative approach should be taken. Shift share analysis gives Indigenous individuals some characteristics of non-Indigenous individuals hypothetically, which will result in the closure of the income gap. Using this approach, the impact of each variable is captured by the magnitude of the income gap closure, resulting in the giving of Indigenous people the variable of non-Indigenous people.

To measure the inequality captured by *Income function gap* $(F_n(edu_i, empl_i) - F_i(edu_i, empl_i))$, Indigenous people will be given the average weekly income of the non-Indigenous people if they have the identical employment status and education status and other identification (age group and gender). This variable captures the difference in income that is not a result of employment or education. Assume the Indigenous people have the identical average weekly income of non-Indigenous people if they have the same education and employment level, such shift share will increase the hypothetical Indigenous income with 0 magnitude, implying 0 income gap as a result of the income function gap.

This gap is sometimes referred to as *"discrimination effect"*, where Indigenous people are paid less than non-indigenous people on average only because they are Indigenous, and not because they have less education or are not employed at the same level. However, it is not true in this context for many reasons. For example, the income gap results from omitted variables such as urban-rural distribution, and language backgrounds. Education level can be captured in the so-called *"discrimination effect*; conversely, heterogeneity of the treatment in variables might contribute to such a gap. For example, the schools that white students went to offered high-quality education as a preparation for higher education, while the racial minorities' schools only provided low-quality education which served the purpose of regulation and discipline, and as a preparation for low-skilled position in the labour market. Even though the reason behind such heterogeneity can still be discrimination, it is not simply someone receiving less pay due to

their racial status, using the word *"Discrimination effect"* can be misleading.

To measure the inequality captured by *education resulted in income gap* ($\delta_{edu} * F^n_{edu}$), Indigenous people are given the identical education distribution of the non-Indigenous people for each age and gender group. This shift-share analysis does not require any employment information, therefore failing to include the employment level improvement from a higher education level. The hypothetical average weekly income of Indigenous people increase is only a result of a higher proportion of people having a higher level of education, and an individual with a higher level of education has a higher weekly income regardless of the reason behind it being human capital accumulation or better employment status. In this particular example, say for an arbitrary age and gender group, non-Indigenous people have $(e_1, e_2, e_3, e_4, e_5)^T$ people in each education category respectively, and therefore have $E = \sum_{i=1}^{5} e_i$ people in such age and gender group. Say there are in total E' Indigenous people in such age and gender groups, the hypothetical education distribution of Indigenous people $\hat{e'} \in \mathbb{R}^5$ is given by $E' * \frac{1}{E}(e_1, e_2, e_3, e_4, e_5)^T$.

To measure the *differences in employment generating function* ($g_n(edu_i) - g_i(edu_i)$), Indigenous people were given the same employment distribution as non-Indigenous people if they belong to the same age, gender group **and have the same level of education**. This is indeed the case as $\delta_{edu} * g^n_{edu}$ addressed the employment level differences as a result of education. To control it, Indigenous people should be given the same employment level if the education level is identical, namely have the identical *employment conditioned on education*.

For an arbitrary education, age, and gender group, non-Indigenous people have $(l_1, l_2, l_3, l_4, l_5)^T$ people in each education category respectively, and therefore have $L = \sum_{i=1}^{5} l_i$ people in such age and gender group. Say there are in total L' Indigenous people in such education, age and gender groups, the hypothetical education distribution of Indigenous people $\hat{l'} \in \mathbb{R}^5$ is given by $L' * \frac{1}{L}(l_1, l_2, l_3, l_4, l_5)^T$. The hypothetical average weekly income of Indigenous people increase is not caused by them "getting more education", as the education distribution of the Indigenous people remains unchanged; it increases because there is a higher proportion of the population receiving a higher employment outcome given their education level.

We can take a counter approach to measure the impact of employment on the income gap directly. To measure the inequality captured by *employment resulted in income gap*, Indigenous people were given the same employment distribution as non-Indigenous people if they belong to the same age and gender group. The calculation procedure is identical to the *"Education Resulted Income Gap"*. Similarly, this calculation does not require any information on education level. The hypothetical average weekly income of Indigenous people increase is only a result of a higher proportion of people having a better employment status, and an individual with a better employment status has a higher weekly income, this shift-share method captures both $\delta_{edu} g^n_{edu}$ and $[g_n(edu_i) - g_i(edu_i)]$.

If Indigenous people are given the same *level of education* and *employment conditioned on education*, then there are no differences between Indigenous and non-Indigenous people in terms of employment status distribution. But If Indigenous people are given the same *level of education* and *employment*, the *employment conditioned on education* will only be identical if employment and education are independent, which violates the assumption.

6 Descriptive Statistics: Gaps in Education and Employment

The decomposition is not trivial if and only if the Education gap and Employment gap are not 0. Empirical evidence suggests there is a non-trivial gap between Indigenous and non-Indigenous people regarding education level and employment outcome. The graph here is similar to and supports the results provided by Biddle (2024).

6.1 Education Gap

This report did not take a traditional approach by translating the level of education into the year of schooling due to the large difference in employment outcome for Indigenous people between finishing year 12 without pursuing a qualification (eg. Cert III or IV) and not finishing year 12 but pursuing a qualification that will be demonstrated in the later section. I calculated the share of the population with a certain education level grouped by age, gender and Indigenous status.

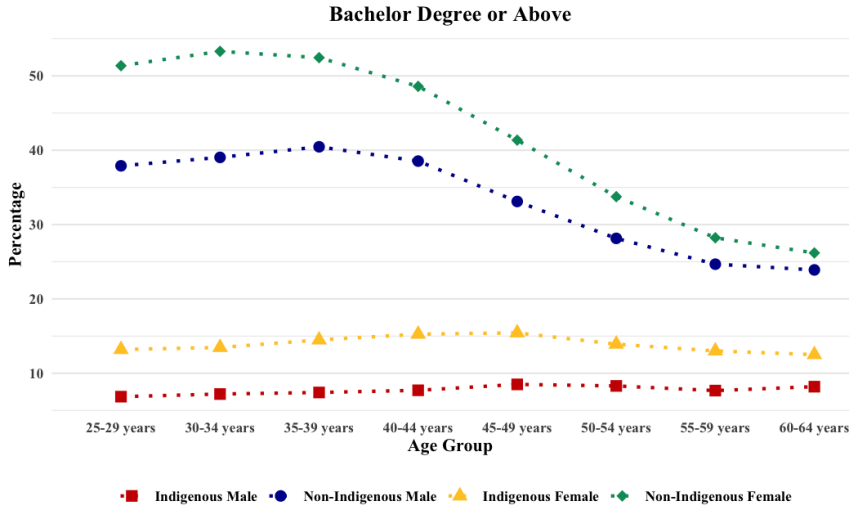

Figure 5: The share of the population that has a Bachelor or above

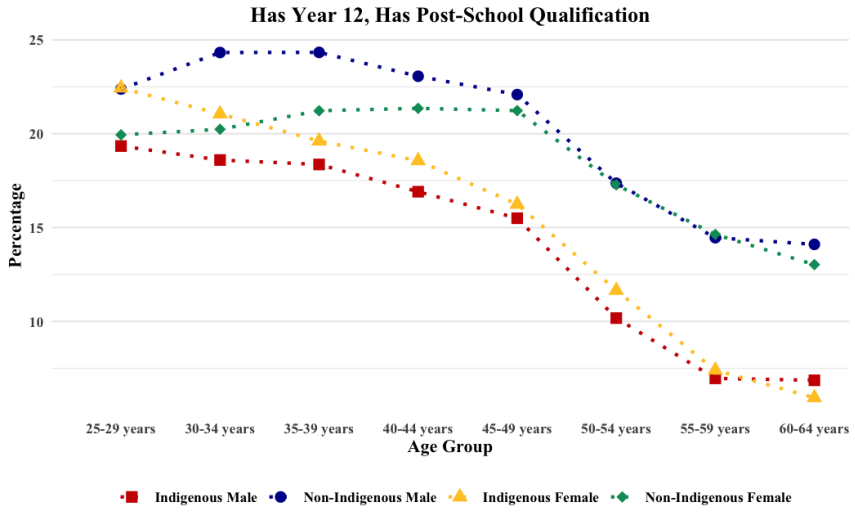

Figure 6: The share of the population that finished year 12, have a post-school qualification

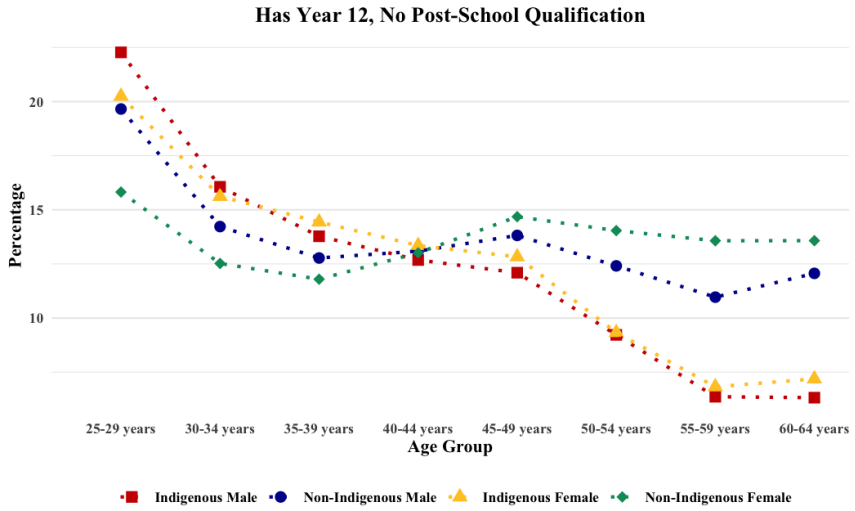

Figure 7: The share of the population that finished year 12, does not have post-school qualification

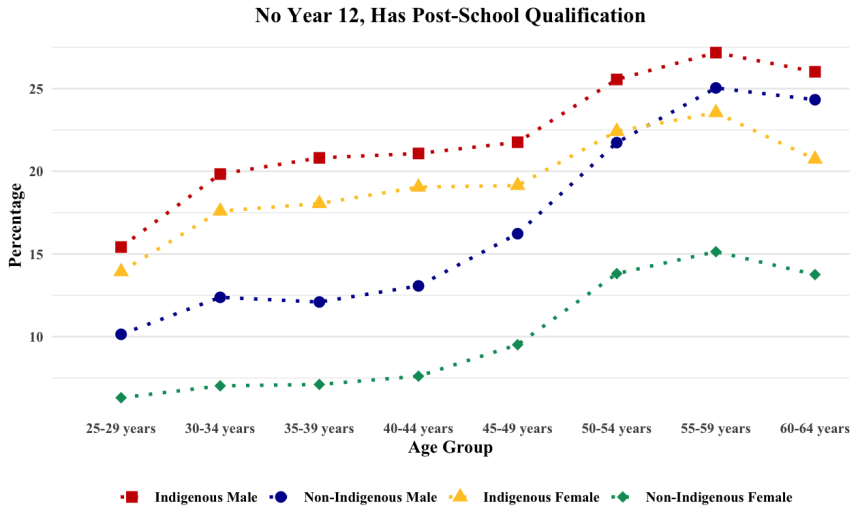

Figure 8: The share of the population that do not have year 12, have post-school qualification

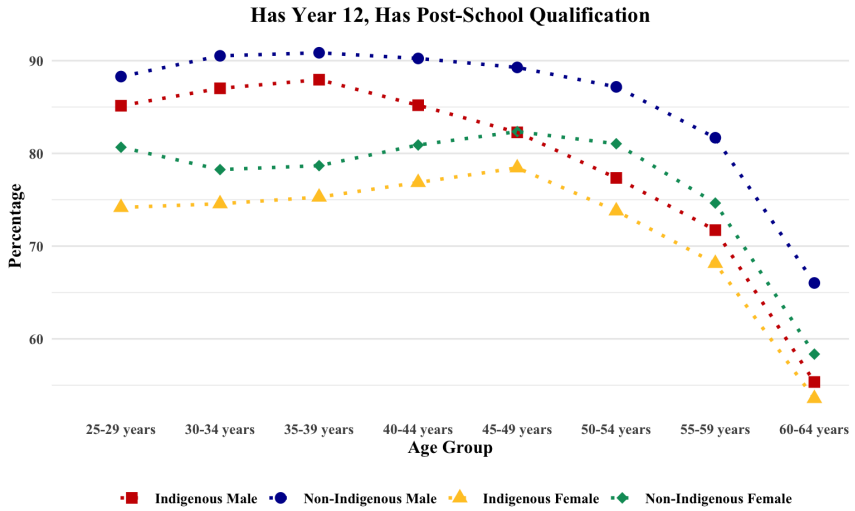

Figure 9: The employed percentage for those who finished year 12, and have a post-school qualification

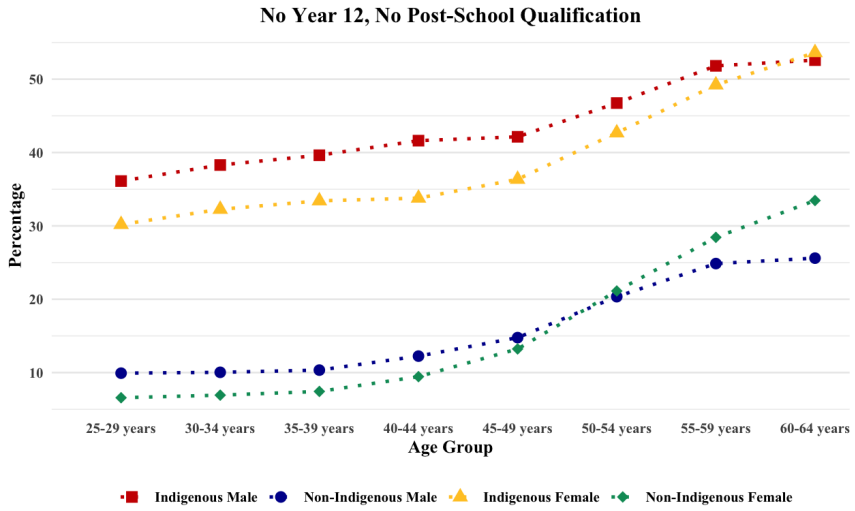

Figure 10: the share of the population that do not have year 12, does not have post-school qualification

The improvement in education for all Australians generally seems to have a limited impact on Indigenous people. Compared to the older cohort, the younger non-Indigenous people are likelier to have a bachelor's degree or above; while the share of the population that has a bachelor's degree did not change for the younger Indigenous people. Indigenous people are less likely to have no year 12 education since the share of those without year 12 decreases as the age decreases. This improvement is impressive, and it aligns with the target set by *Closing the Gap* plan. However, there is a clear difference between Indigenous and non-Indigenous people in each education category, particularly for the younger generation. Specifically, Indigenous people are more likely to have no year 12 or post-school qualification, and more likely to have a bachelor's degree, particularly for the younger Indigenous people.

6.2 Employment Gap

Despite having 5 variables representing employment level, this section only presents 2 main ratios; *employed percentage* and *full-time employed percentage*.

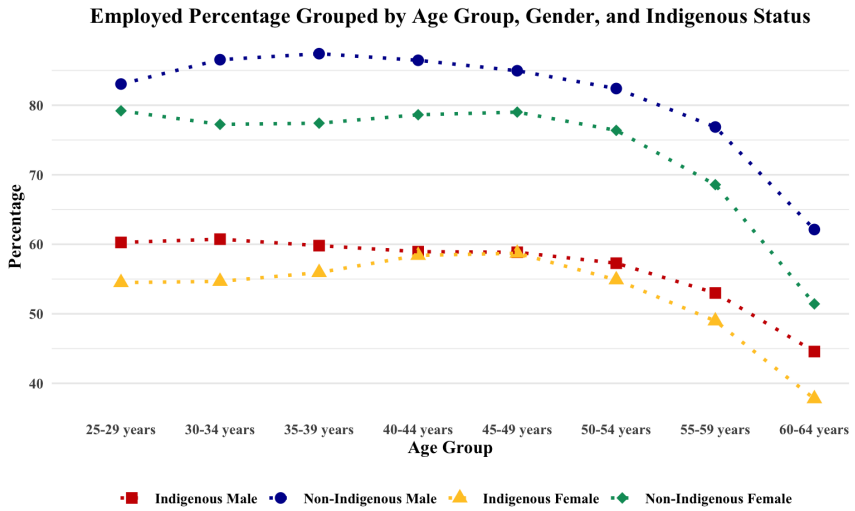

Figure 11: The employment percentage grouped by age group, gender and Indigenous Status

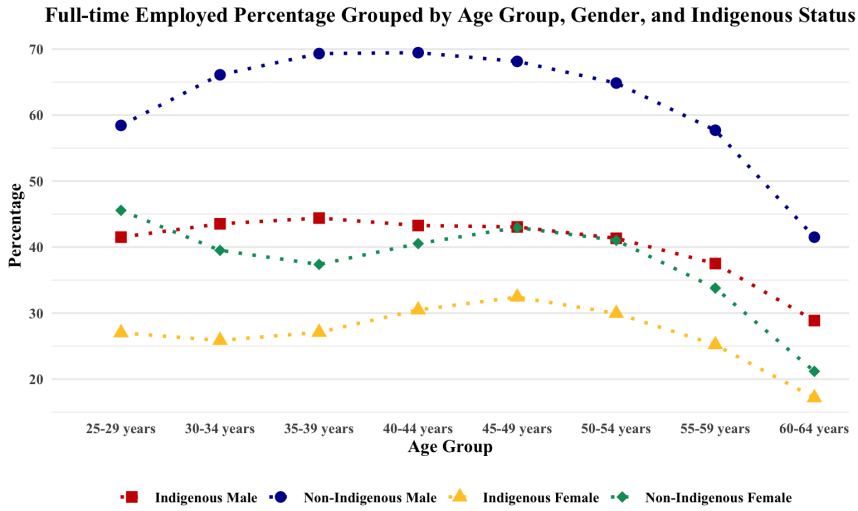

Figure 12: The full-time employment percentage grouped by age group, gender and Indigenous Status

It will surprise few that Indigenous people are less employed than non-Indigenous people. However, the differences in the proportion of Indigenous people and non-Indigenous people who are full-time employed are smaller than those between males and females. Indigenous males are indeed less likely to work full-time than non-Indigenous males, but they are more likely to be employed full-time in most age groups than non-Indigenous females.

6.3 Employment Gap Conditioned on Education

6.3.1 Employed Percentage

By arranging the employed percentage of the population grouped by education level, age, gender and Indigenous status, I obtained the following graphs.

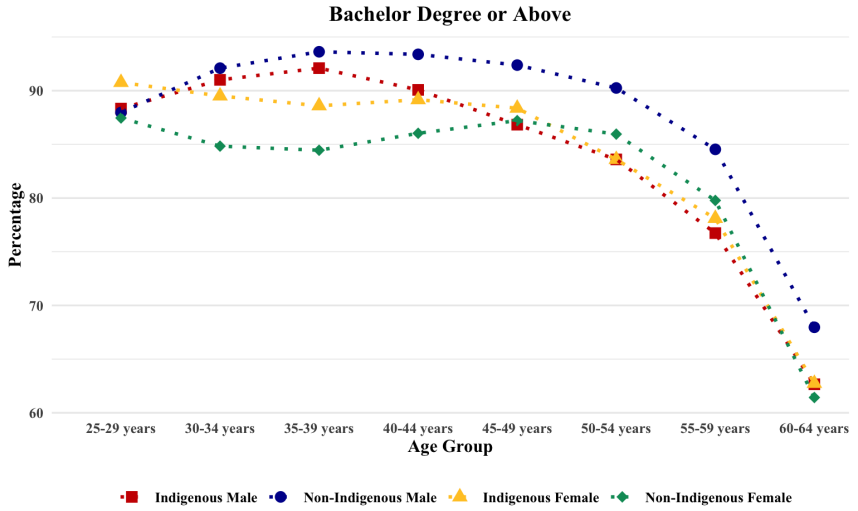

Figure 13: The employed percentage for those who have a Bachelor's or above

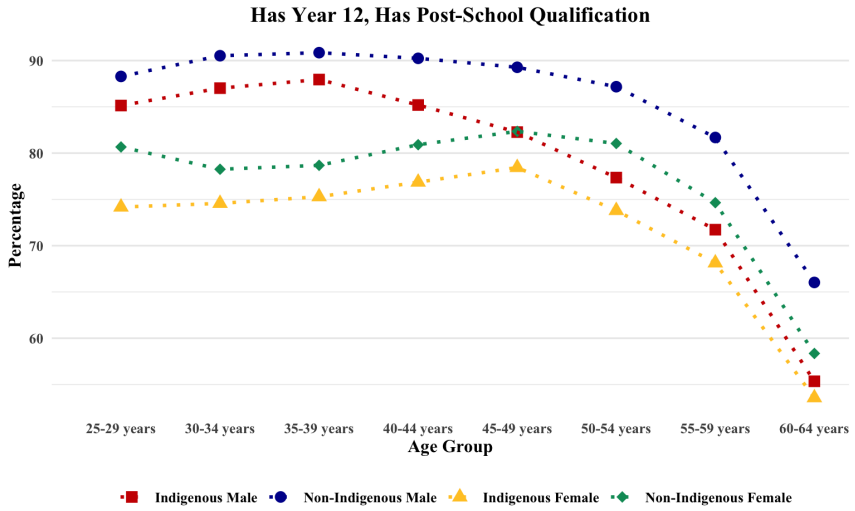

Figure 14: The employed percentage for those who finished year 12, and have a post-school qualification

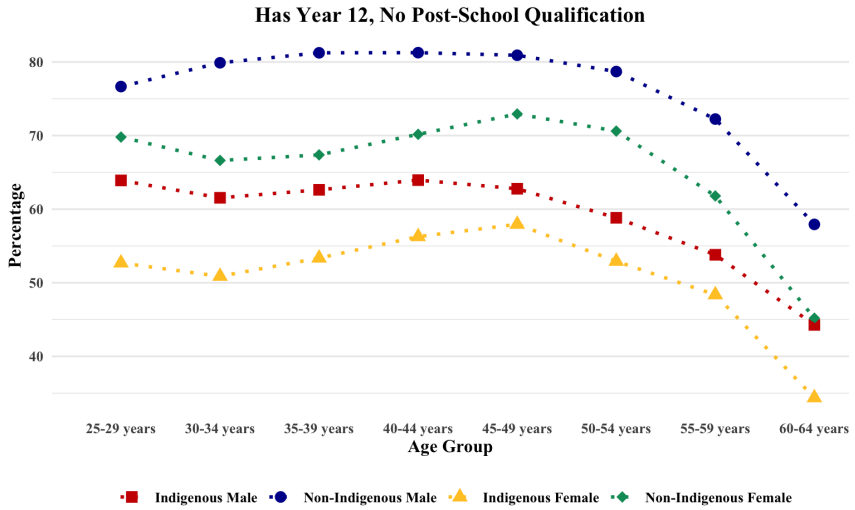

Figure 15: The employed percentage for those who finished year 12, have no post-school qualification

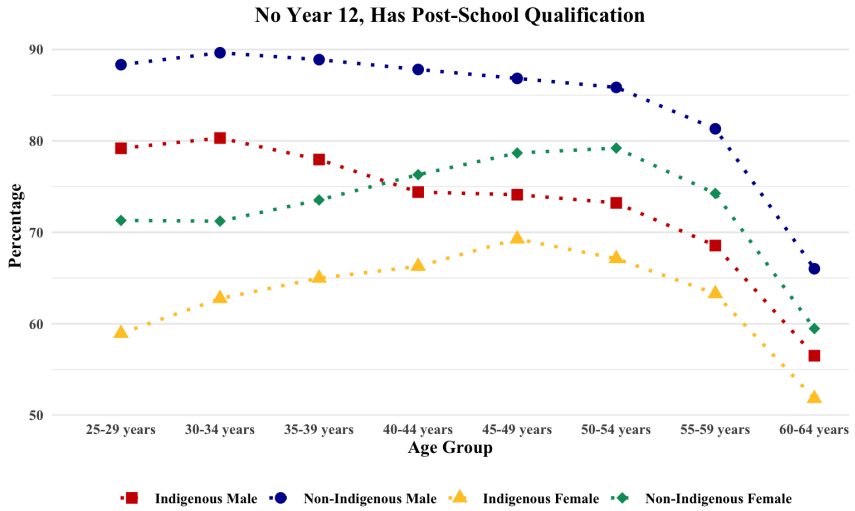

Figure 16: The employed percentage for those who do not have year 12, have post-school qualification

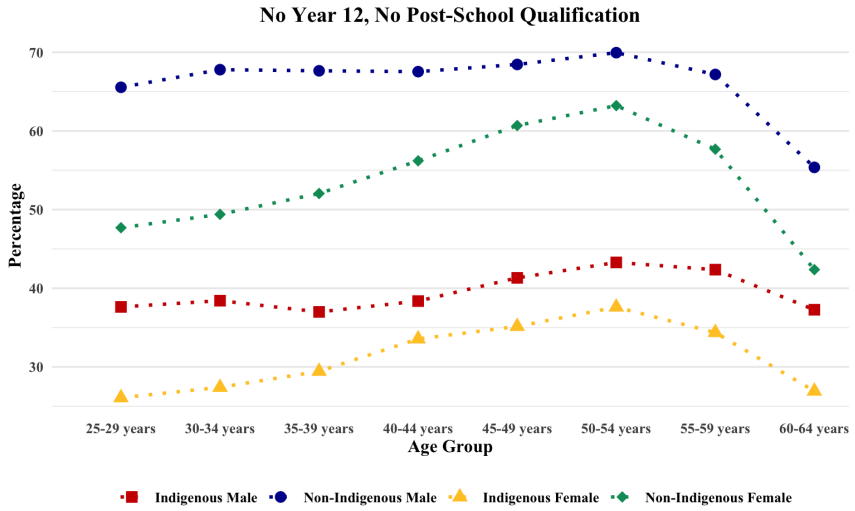

Figure 17: The employed percentage that for those who do not have year 12, have no post-school qualification

For employment percentage, the following are observed:

- Differences in employment rate between Indigenous and non-Indigenous people increase as the level of education decreases.
- For those with higher education levels, the employment gap is smaller for the younger people; while for those with lower education levels, the employment gap is larger for the younger people.
- For the younger cohort, Indigenous status is not always the main factor that drives inequality. For those with higher education levels, the employment gap between males and females is larger than the employment gap between Indigenous and non-Indigenous people. This is shown as young Indigenous males have a higher employment percentage than non-Indigenous females, though lower than non-Indigenous males, for those who have a bachelor's degree or above, or a post-school qualification regardless of whether they have completed year 12 or not. For those with lower levels of education, the Indigenous factor is a larger drive than gender. This is revealed as young Indigenous males are making less than non-Indigenous females for those who do not have a post-school qualification.

Compare these results and observations with the employment gap without grouping by education level, which Indigenous employment gap is larger than the gender employment gap. It emphasises the low share of Indigenous people who do not have a post-school qualification. Closing the gap in Post-school qualification acquisition is not a part of *Closing the gap*'s target, even though young Indigenous people will benefit more in terms of employment level by pursuing a post-school qualification than finishing year 12, which takes roughly the same amount of time.

	Gender	education level	employed percentage
1	Male	Has Year 12, No Post-School Qualification	61.66
2	Male	No Year 12, Has Post-School Qualification	73.82
3	Female	Has Year 12, No Post-School Qualification	52.53
4	Female	No Year 12, Has Post-School Qualification	63.48

Table 10: Comparing the impact on the employed percentage from finishing year 12 or a post-school qualification for the Indigenous people

Focusing on Indigenous employment inequality, the observation put the young Indigenous people with lower education levels (no degree or post-school qualification) as the 'main victims' of the Indigenous employment inequality.

6.3.2 Full-time Employed Percentage

The full-time employed percentage across different education levels has a lot of similarities with the employed percentage: the lower the education level, the larger the Indigenous-non-Indigenous employment gap. For the cohort with a post-school qualification, the full-time employment gap between the Indigenous and non-Indigenous people is smaller for young people, and obtaining a post-school qualification seems to improve the full-time employed percentage more than the impact on such from finishing year 12.

	Gender	education level	employed percentage
1	Male	Has Year 12, No Post-School Qualification	42.28
2	Male	No Year 12, Has Post-School Qualification	54.47
3	Female	Has Year 12, No Post-School Qualification	25.85
4	Female	No Year 12, Has Post-School Qualification	30.54

Table 11: Comparing the impact on the full-time employed percentage from finishing year 12 or a post-school qualification for the Indigenous people

The main difference between the employed percentage and the full-time employed percentage is the role of gender. Gender full-time employment inequality dominates the Indigenous status full-time employment inequality in all education groups as Indigenous males are doing better than non-Indigenous females, though doing worse than non-Indigenous males. The unique role of Indigenous women in full-time employment will be discussed in subsection 8.3.

By arranging the percentage of full-time employed people in the population grouped by education level, age, gender, and Indigenous status, I obtained the following graphs.

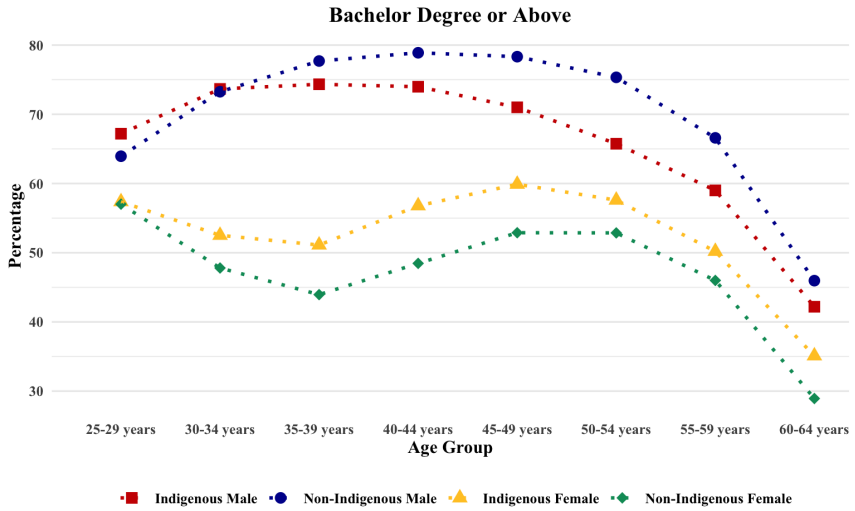

Figure 18: The full-time employed percentage for those who have a Bachelor's or above

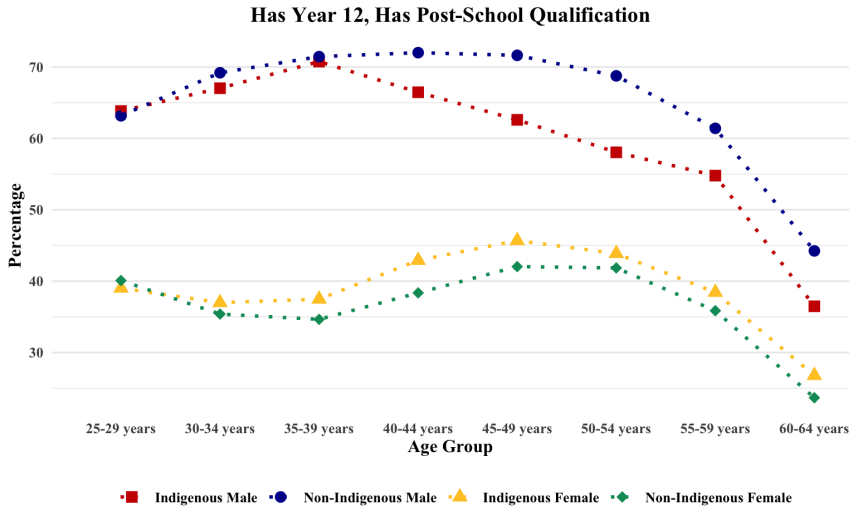

Figure 19: The full-time employed percentage for those who finished year 12, and have a post-school qualification

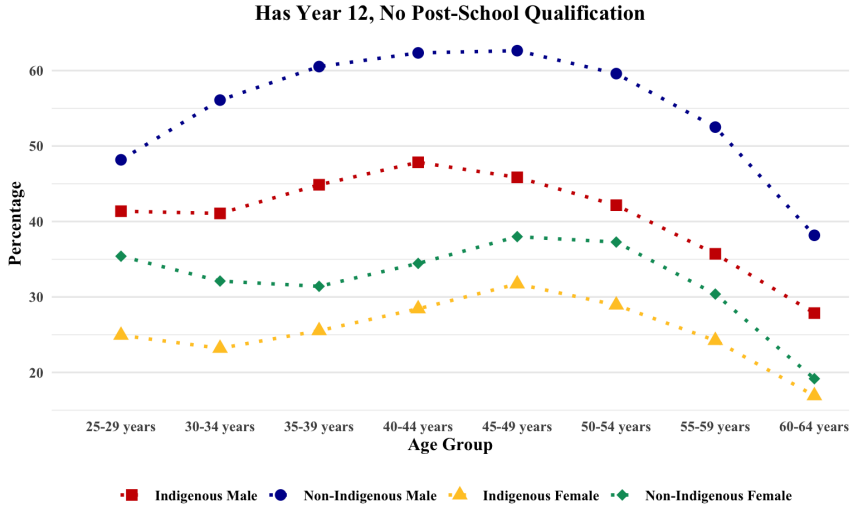

Figure 20: The full-time employed percentage for those who finished year 12, have no post-school qualification

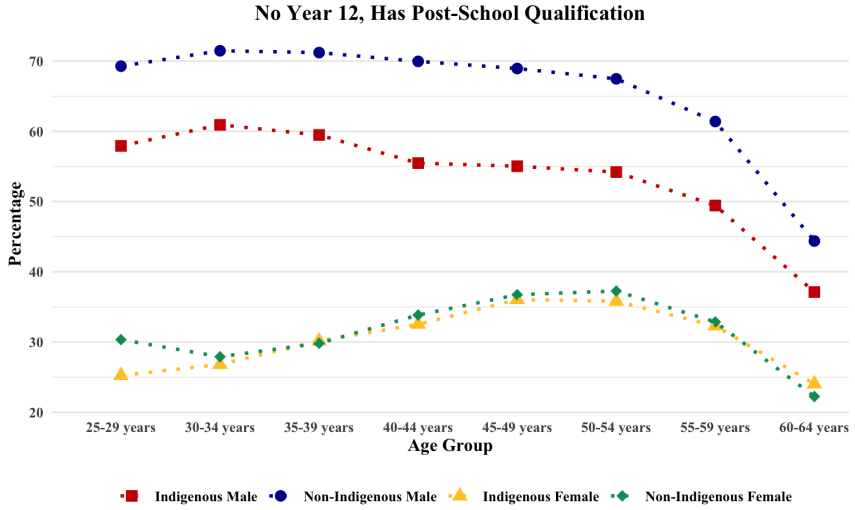

Figure 21: The full-time employed percentage for those who do not have year 12, have post-school qualification

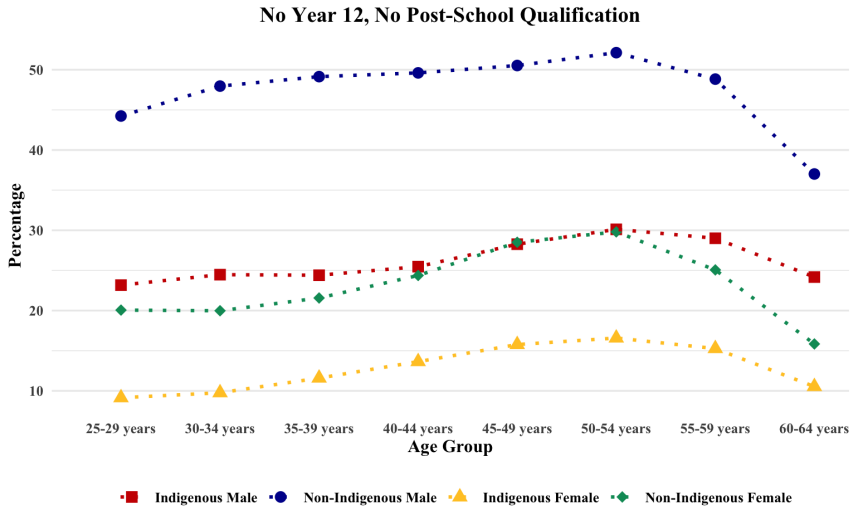

Figure 22: The full-time employed percentage that for those who do not have year 12, have no post-school qualification

7 Shift Share Analysis Result

As demonstrated before, the income gap can be decomposed into 4 components: income function gap, education gap, employment generating gap and gap of employment generated by education difference. The following section decomposes the income gap without shift-share as demonstrated in section 4 using the shift-share analysis method mentioned in subsection 5.2. Due to the random perturbation, the non-Indigenous people's income can be different in different shift-share models.

I present the shift-share analysis result by arranging a table on income ratio δ after each shift share. Since the baseline δ is the same (without shift-share), the shift-share with the largest δ has the largest change in the magnitude of the inequality. I then present the income grouped by age graph to observe the differences in shifting the variable across different age groups.

Overall, giving the education level of non-Indigenous people to the Indigenous people reduces the income gap the largest, almost eliminating the income gap. Particularly for the younger cohort. The shift-share on average weekly income yields little change, implying Indigenous people are receiving a similar amount of income to non-Indigenous people if they have identical education and employment levels. The shift-share on employment conditioned on education also yields a little reduction in income inequality, implying the *employment generating function* is similar between Indigenous and non-Indigenous people, though the employment conditioned on education seems different as previously described, it did not play a big component in income inequality compare to the largely different education level and the employment outcome generated by such. This is also consistent with the result of high differences in education outcomes between Indigenous and non-Indigenous people.

7.1 Shift-Share on Average Weekly Income

This shift share measures the gap induced by the income function gap, that is an Indigenous person receiving less income than a non-Indigenous person who has the same education and employment level.

	Identity	Income
1	Indigenous Male	1280.25
2	Non-Indigenous Male	1693.42
3	Indigenous Female	948.98
4	Non-Indigenous Female	1235.61
5	Indigenous	1104.14
6	Non-Indigenous	1459.74

Identity	Income Ratio δ
Male	0.76
Female	0.77
Total	0.76

Table 12: Income and income ratio δ after shift-share on average weekly income

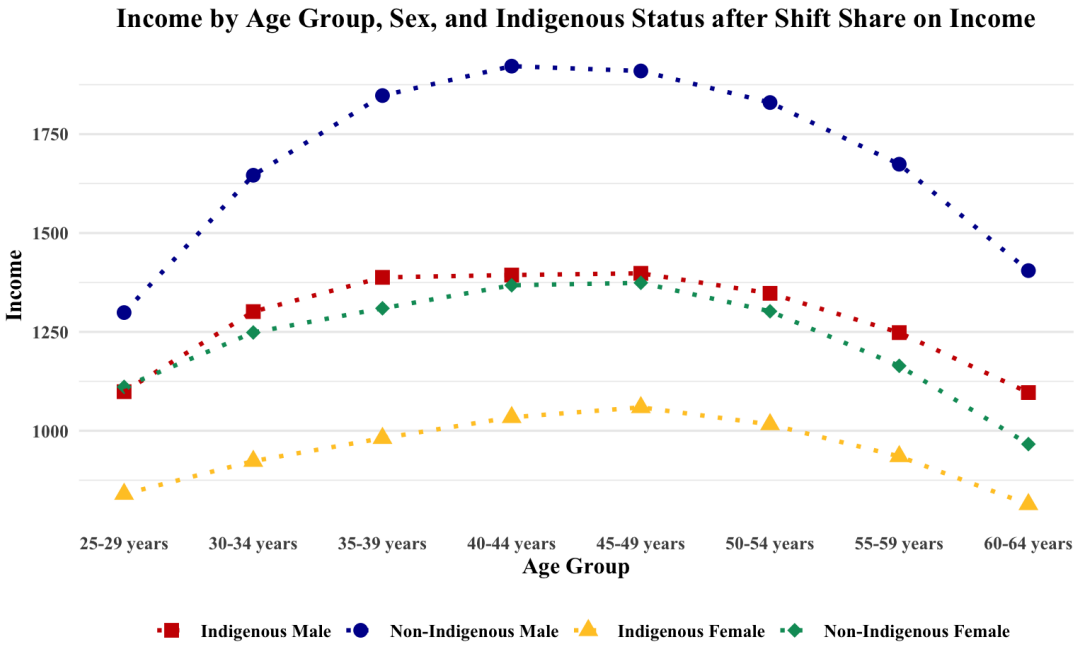

Figure 23: The average weekly income grouped by age, gender and Indigenous status after shift share on income

7.2 Shift-Share on Education

This shift-share measures the income gap generated by the higher education level of the non-indigenous people and *higher education level resulted employment*.

	Identity	Income
1	Indigenous Male	1502.23
2	Non-Indigenous Male	1716.37
3	Indigenous Female	1237.69
4	Non-Indigenous Female	1276.91
5	Indigenous	1361.60
6	Non-Indigenous	1492.08

Identity	Income Ratio δ
Male	0.87
Female	0.96
Total	0.91

Table 13: Income and income ratio δ after shift-share on Education

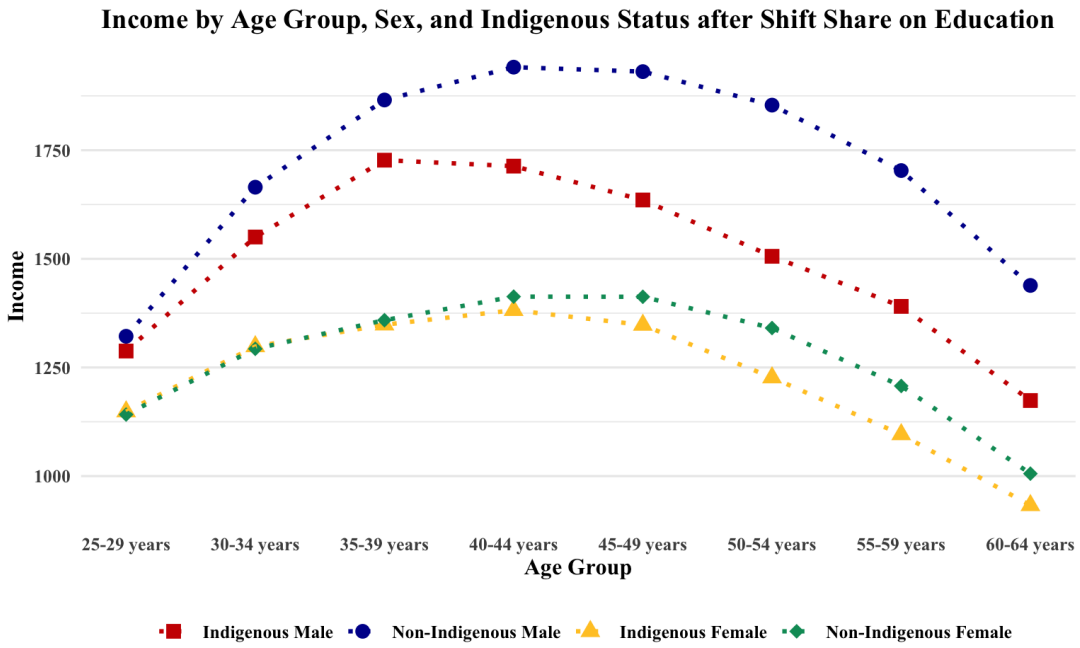

Figure 24: The average weekly income grouped by age, gender and Indigenous status after shift share on education level

7.3 Shift-Share on Employment

This shift share captures the income gap generated by different *employment generating function* and *higher education level resulted employment*

	Identity	Income
1	Indigenous Male	1395.96
2	Non-Indigenous Male	1677.60
3	Indigenous Female	1069.83
4	Non-Indigenous Female	1217.95
5	Indigenous	1222.35
6	Non-Indigenous	1441.80

Identity	Income Ratio δ
Male	0.83
Female	0.87
Total	0.85

Table 14: Income and income ratio δ after shift-share on Employment

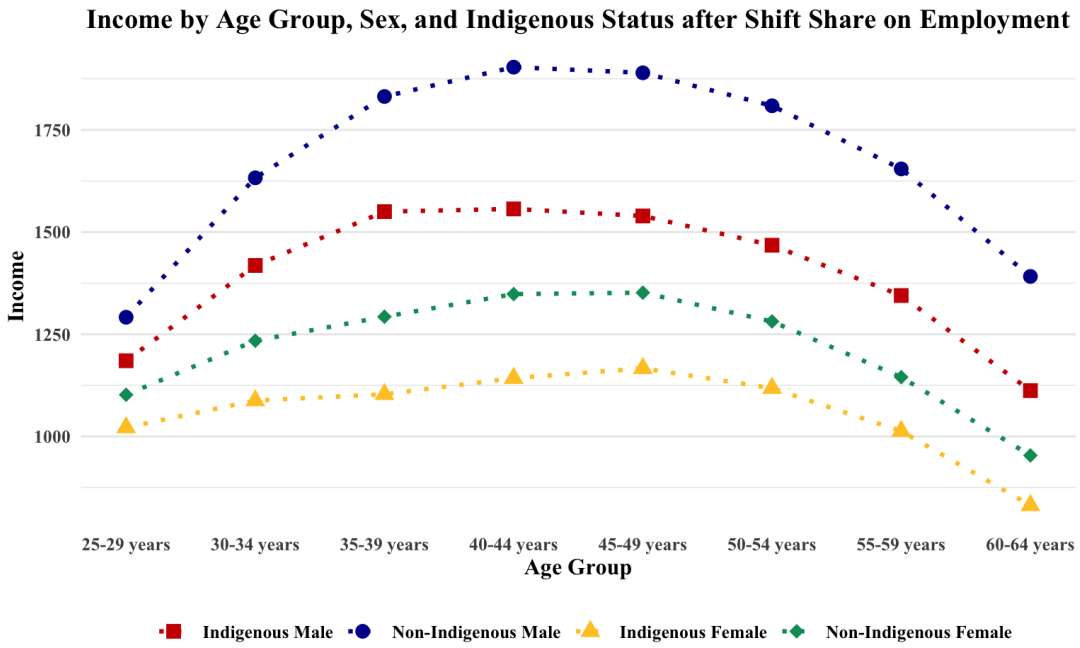

Figure 25: The average weekly income grouped by age, gender and Indigenous status after shift share on employment status

7.4 Shift-Share on Employment Conditioned on Education

This shift-share only captures the income gap generated by different *employment generating function*.

	Identity	Income
1	Indigenous Male	1320.21
2	Non-Indigenous Male	1693.42
3	Indigenous Female	977.89
4	Non-Indigenous Female	1235.61
5	Indigenous	1138.23
6	Non-Indigenous	1459.74

Identity	Income Ratio δ
Male	0.78
Female	0.79
Total	0.78

Table 15: Income and income ratio δ after shift-share on employment conditioned on education

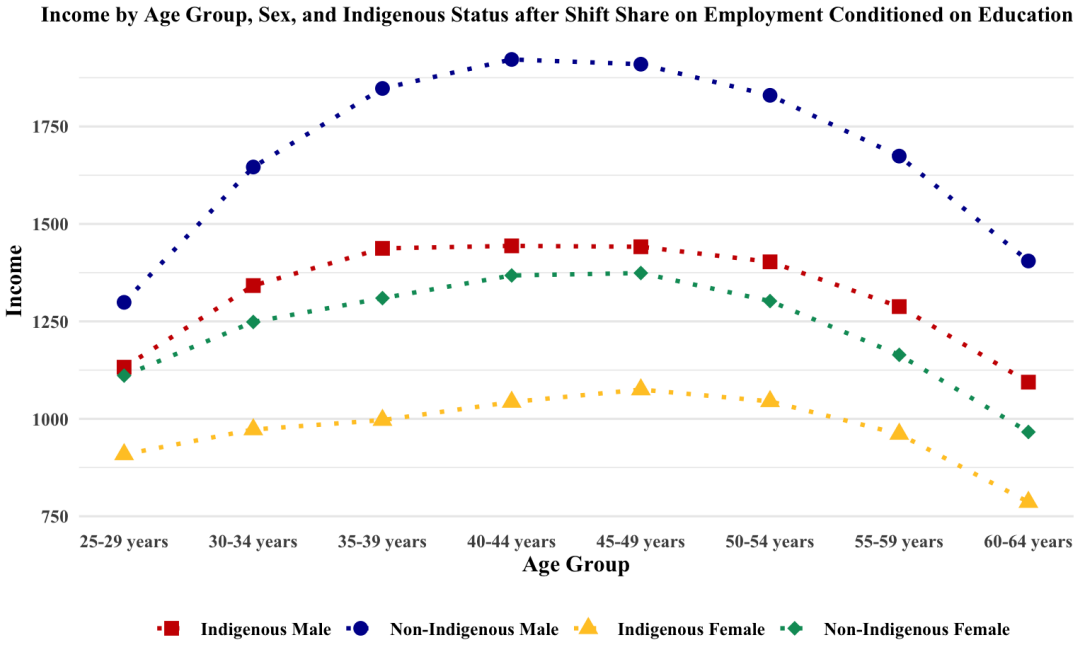

Figure 26: The average weekly income grouped by age, gender and Indigenous status after shift share on employment status conditioned of education level

8 Cross Time Comparison: the Income Gap in 2016

In this section, I performed the shift-share analysis from before on the 2016 census data. Overall, the income inequality is larger in 2016 compared to 2021. I then find the new δ for each shift-shared result.

By giving Indigenous people the average weekly income of non-Indigenous people with identical education and employment levels, the income ratio increased from 0.68 to 0.74, compared to 2021, the same ratio changed from 0.7 to 0.76. This means the income differences between Indigenous and non-Indigenous people who have the same level of education and employment remained relatively stable in the past 5 years. The same goes for shift sharing on education: income ratio changes from 0.68 to 0.88 in 2016, 0.70 to 0.91 in 2021; and employment: income ratio changes from 0.68 to 0.84 in 2016, 0.70 to 0.85 in 2021; and employment conditioned on education: income ratio changes from 0.68 to 0.77, 0.70 to 0.78 in 2021. Therefore, even though the income differences decreased from 2016 to 2021, the decomposition of such does not have a major change. The result is presented as follows:

8.1 Income Gap in 2016

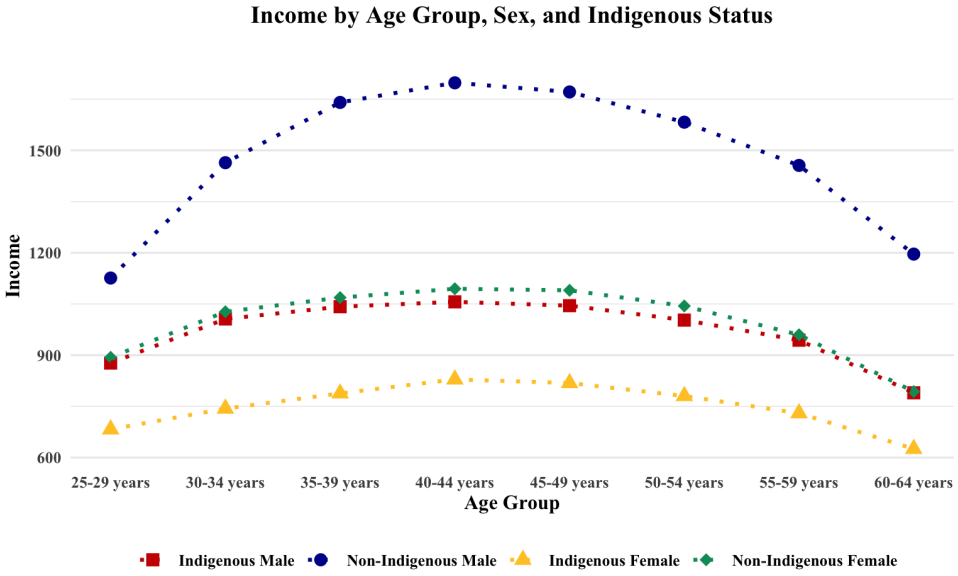

Figure 27: The average weekly income grouped by age, gender and Indigenous status (2016)

	Identity	Income
1	Indigenous Male	967.55
2	Non-Indigenous Male	1491.41
3	Indigenous Female	755.65
4	Non-Indigenous Female	1032.18
5	Indigenous	854.45
6	Non-Indigenous	1256.40

Identity	Income Ratio δ
Male	0.64
Female	0.73
Total	0.68

Table 16: Income and income ratio δ (2016)

8.2 Shift-Share Analysis Result Income Gap in 2016

This graph shows the income grouped by age, gender and Indigenous status under 4 types of shift-share analysis from the 2016 census.

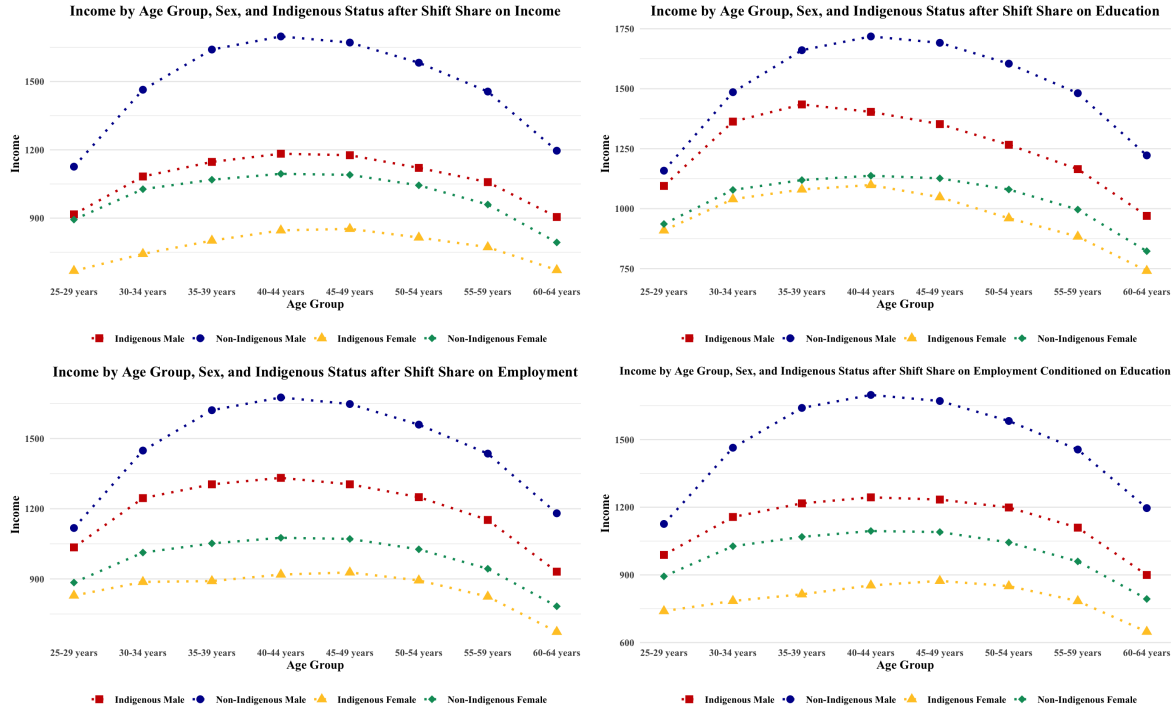

Figure 28: The average weekly income grouped by age, gender and Indigenous status after shift-share (2016)

This table presents the income ratio δ under 4 types of shift-share models: shifting income, education, employment and employment conditioned on education, all results are from the 2016 census.

Identity	Income δ
Male	0.72
Female	0.77
Total	0.74

Identity	Education δ
Male	0.84
Female	0. 94
Total	0.88

Identity	Employment δ
Male	0.82
Female	0.88
Total	0.84

Identity	Empl conditioned on Edu δ
Male	0.76
Female	0.79
Total	0.77

Table 17: Shift-share analysis result for income after age standardised

8.3 The Story of Indigenous Women

Recall the 5 graphs on full-time employment percentages grouped by education level present in subsubsection 6.3.2. Here are a few observations:

- Indigenous males are always doing better than non-Indigenous females, meaning here the gender inequality is larger than Indigenous inequality.

- For the only statistic presented in this report: Indigenous women are doing better than non-Indigenous women if they have a bachelor's degree or above, or have both year 12 and post-school qualifications.
- For the women has a bachelor's degree or above, or have a post-school qualification. The full-time employed percentage goes down then goes up at around the age of 30 to 40. This trend became unclear for those who have a lower level of education.

Few would disagree that this "down and up" trend is due to maternity leave or "return to family". The question is how this trend connects to the high-achieving Indigenous women outperforming non-Indigenous women. The higher the education level is, the more likely an individual is to be full-time employed, but it is less likely for an Indigenous woman to have both year 12 and a post-school qualification, or have a bachelor's degree or above.

These 2 facts capture the selection of the selection performed by education institutions. If the Indigenous women have obtained a higher education level, they have been selected as the few of the few. Unlike non-Indigenous women, who can choose to go to university without being at the top of their cohort, Indigenous women have to put in more effort or be more willing to go to university or complete a post-school qualification. Once they are in the labour force, it is less likely for them to leave the workforce due to the high sunk cost of receiving an education, or a high opportunity cost to give up the education they value highly due to family and childcare.

As a result of a lower education level for Indigenous women, the hardship of obtaining a higher level of education selected those who were truly willing to receive more education, making those who were successful in the system receive a better employment outcome. This captures the willingness of education acquisitions for some people and reveals the potential to improve the education level of the Indigenous community.

9 Conclusive statement

One great aspect of the analysis methods used in this report is one does not have to quantify every input in the model. We can obtain the characteristics of the model from the change in outcome. The result is clear and straightforward: the main factor of the income gap is coming from the gap in education level. If Indigenous people have the same level of education as non-Indigenous people, the Indigenous people will have 91% of non-Indigenous people's income, instead of 70%. A higher level of education also induces a higher employment level, which is correlated with higher income. This report did not find strong evidence to suggest forms of discrimination, regardless it is the income yields by identical education and employment level, or employment level yields by identical education level between Indigenous and non-Indigenous people.

References

Australian Bureau of Statistics (2016), 'Sexp sex, age5p - age in five year groups, ingp indigenous status, heap - 1 digit level, hscp highest year of school completed and lfhrp labour force status and hours worked not stated by incp total personal income (weekly)'. [2016 Census TableBuilder, Employment, Income and Education], (accessed October 2024).
URL: *https://www.abs.gov.au/statistics/microdata-tablebuilder/tablebuilder/topics*

Australian Bureau of Statistics (2021), 'Sexp sex, age5p age in five year groups, ingp indigenous status, lfsp labour force status, hscp highest year of school completed and 1-digit level heap level of highest educational attainment by incp total personal income (weekly)'. [2021 Census TableBuilder, employment, income and education], (accessed October 2024).
URL: *https://www.abs.gov.au/statistics/microdata-tablebuilder/tablebuilder/topics*

Biddle, N. (2024), Towards an estimate of aboriginal and torres strait islander human capital and how it is changing over time, Working Paper 8/2024, Tax and Transfer Policy Institute (TTPI).
URL: *https://ssrn.com/abstract=4894537*

Biddle, N., Howlett, M., Hunter, B. and Paradies, Y. (2013), 'Labour market and other discrimination facing indigenous australians', *Australian Journal of Labour Economics (AJLE)* **16**(1), 91–113.
URL: *https://EconPapers.repec.org/RePEc:ozl:journl:v:16:y:2013:i:1:p:91-113*

Birch, E. and Marshall, D. (2017), 'Revisiting the earned income gap for indigenous and non-indigenous australian workers: Evidence from a selection bias corrected model', *Journal of Industrial Relations* **60**.

Howlett, M., Gray, M., Hunter, B. et al. (2015), 'Unpacking the income of indigenous and non-indigenous australians: wages, government payments and other income'.
URL: *http://hdl.handle.net/1885/147848*

Kalb, G., Le, T., Hunter, B. H. and Leung, F. (2012), Decomposing differences in labour force status between indigenous and non-indigenous australians, IZA Discussion Papers 6808, Institute for the Study of Labor, Bonn.
URL: *https://hdl.handle.net/10419/62473*

National Indigenous Australians Agency (2009), 'Closing the gap report 2009'.
URL: *https://www.dss.gov.au/sites/default/files/documents/05*$_2$*012/closing*$_t$*he*$_g$*ap.pdf*

Oaxaca, R. (1973), 'Male-female wage differentials in urban labor markets', *International Economic Review* **14**(3), 693–709.
URL: *http://www.jstor.org/stable/2525981*

Appendix

A Survey Information

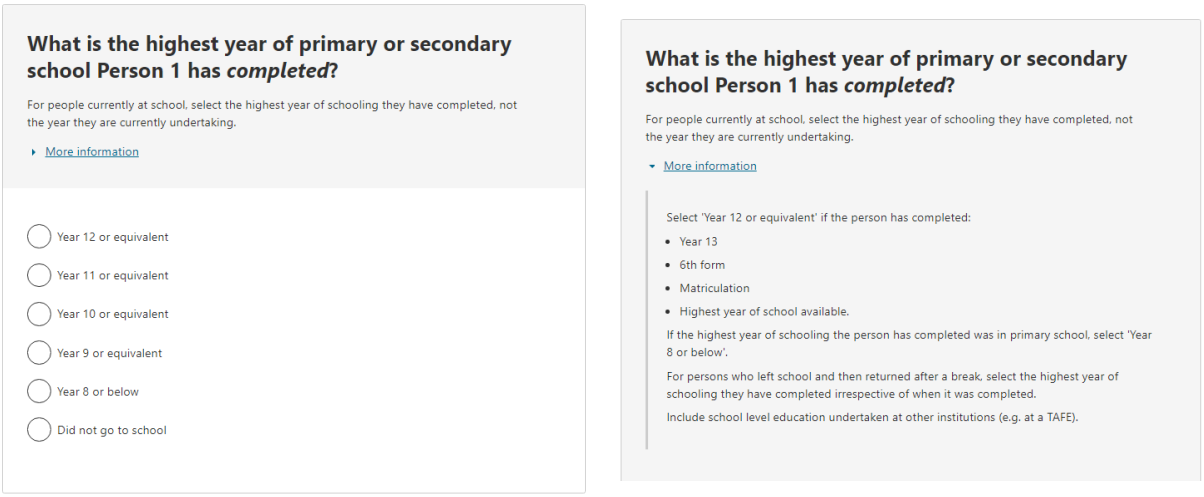

Figure 29: HSCP questions and information provided in the census (2021)

Highest year of school completed	Certificate nfd*(500)	Certificate III & IV nfd*(510)	Certificate IV(511)	Certificate III(514)	Certificate I & II nfd*(520)	Certificate II(521)	Certificate I (524)	Level of education inadequately described(011)	Level of educationnot stated(&&&)
Year 12 (1)	Inadequately described	Certificate III & IV, nfd	Certificate IV	Certificate III	Year 12	Year 12	Year 12	Inadequately described	Not stated
Year 11 (2)	Inadequately described	Certificate III & IV, nfd	Certificate IV	Certificate III	Year 11	Year 11	Year 11	Inadequately described	Not stated
Year 10 (3)	Inadequately described	Certificate III & IV, nfd	Certificate IV	Certificate III	Year 10	Year 10	Year 10	Inadequately described	Not stated
Year 9 (4)	Inadequately described	Certificate III & IV, nfd	Certificate IV	Certificate III	Certificate I & II nfd	Certificate II	Certificate I	Inadequately described	Not stated
Year 8 or below (5)	Inadequately described	Certificate III & IV, nfd	Certificate IV	Certificate III	Certificate I & II nfd	Certificate II	Certificate I	Inadequately described	Not stated
Did not go to school (6)	Inadequately described	Certificate III & IV, nfd	Certificate IV	Certificate III	Certificate I & II nfd	Certificate II	Certificate I	Inadequately described	Not stated
Not stated (&)	Not stated	Certificate III & IV, nfd	Certificate IV	Certificate III	Not stated	Not stated	Not stated	Not stated	Not stated

Note: the number in brackets refer to the classification code used for Highest year of school completed (HSCP) and Non-school qualification: level of education (QALLP).
*nfd: not further defined
Source: Australian Bureau of Statistics, Level of highest educational attainment (HEAP) 2021

Figure 30: Derivation of the HEAP decision table.

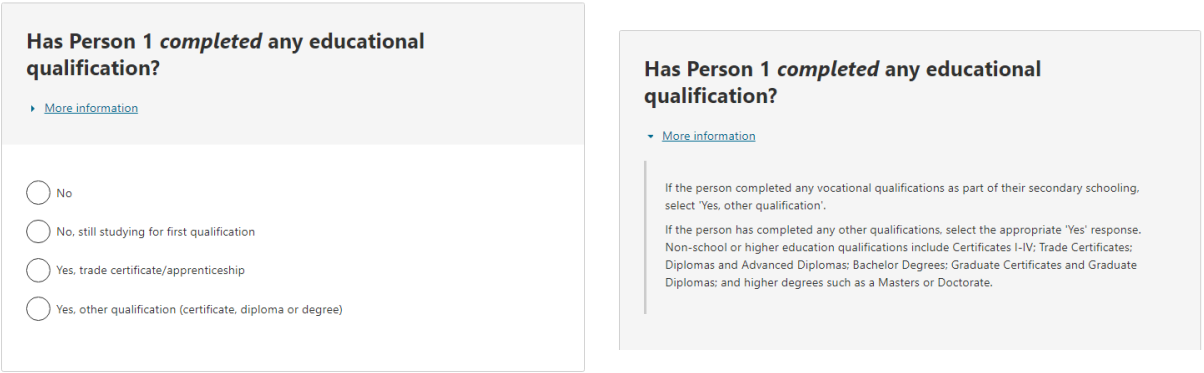

Figure 31: HEAP questions and information regarding completion provided in the census (2021)

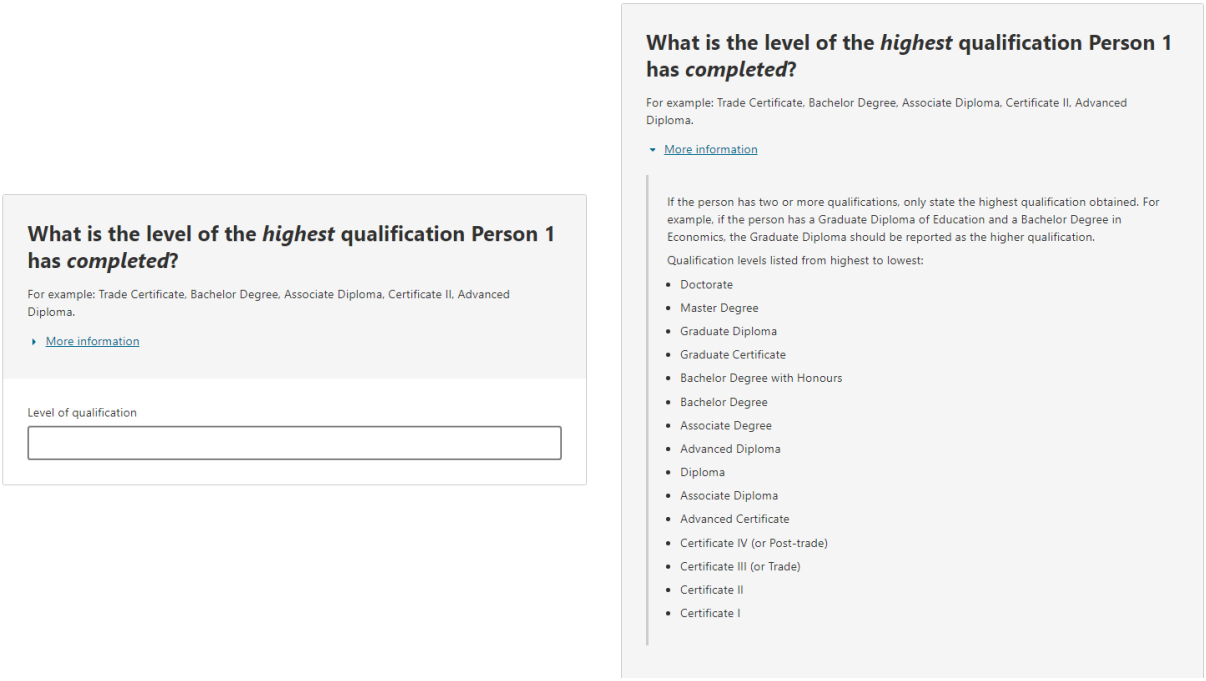

Figure 32: HEAP questions and information regarding highest level education provided in the census (2021)

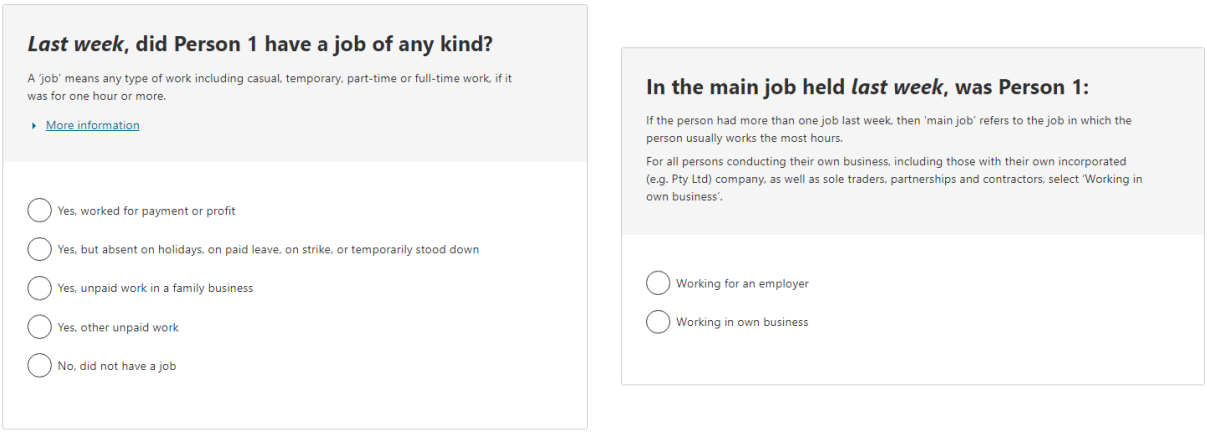

Figure 33: LFSP questions regarding whether or not employed in the census (2021)

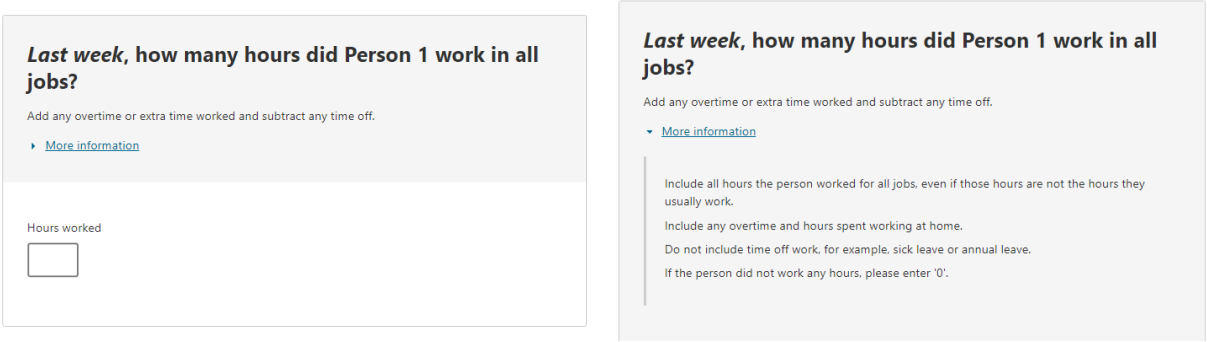

Figure 34: LFSP questions and information regarding hours of work in the census (2021)

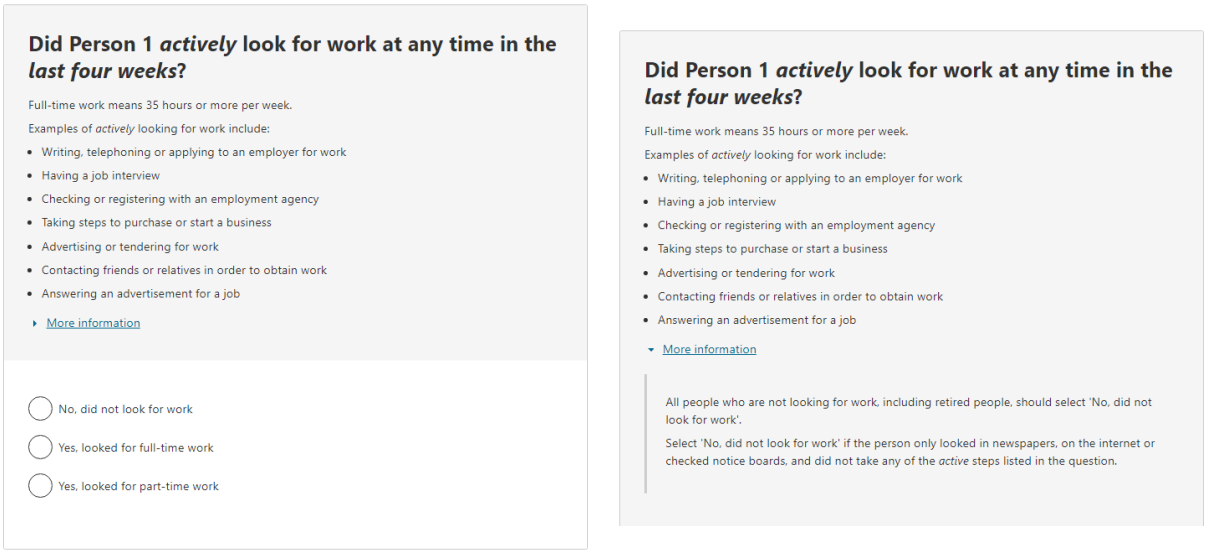

Figure 35: LFSP questions and information regarding job-seeking actions in the census (2021)

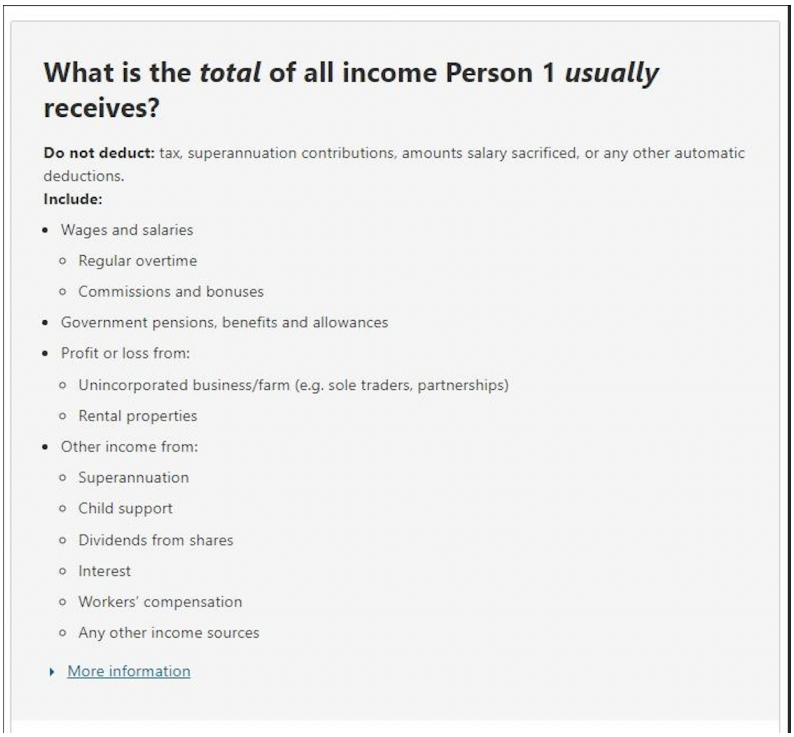

Figure 36: INCP Questions in the census (2021)

B Additional Data

	Identity	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	6.85	7.21	7.41	7.71	8.50	8.30	7.67	8.20
2	Non-Indigenous Male	37.90	39.03	40.46	38.53	33.10	28.14	24.67	23.90
3	Indigenous Female	13.19	13.47	14.48	15.26	15.43	13.92	13.00	12.51
4	Non-Indigenous Female	51.36	53.29	52.44	48.58	41.35	33.75	28.22	26.18

Table 18: The share of the population that has a Bachelor or above

	Identity	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	19.34	18.60	18.36	16.91	15.50	10.18	6.97	6.86
2	Non-Indigenous Male	22.38	24.33	24.33	23.06	22.08	17.36	14.46	14.10
3	Indigenous Female	22.44	21.06	19.61	18.56	16.24	11.65	7.40	5.94
4	Non-Indigenous Female	19.94	20.24	21.22	21.35	21.23	17.28	14.63	13.03

Table 19: The share of the population that finished year 12, have a post-school qualification

	Identity	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	22.27	16.06	13.78	12.68	12.09	9.22	6.37	6.32
2	Non-Indigenous Male	19.66	14.23	12.77	13.09	13.82	12.41	10.97	12.06
3	Indigenous Female	20.24	15.62	14.44	13.35	12.83	9.33	6.83	7.19
4	Non-Indigenous Female	15.82	12.52	11.79	13.01	14.68	14.04	13.57	13.57

Table 20: The share of the population that finished year 12, does not have post-school qualification

	Identity	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	15.42	19.83	20.81	21.08	21.76	25.56	27.18	26.02
2	Non-Indigenous Male	10.14	12.38	12.10	13.07	16.23	21.74	25.04	24.33
3	Indigenous Female	13.93	17.60	18.05	19.05	19.14	22.42	23.55	20.74
4	Non-Indigenous Female	6.31	7.03	7.11	7.61	9.51	13.81	15.14	13.76

Table 21: The share of the population that do not have year 12, have post-school qualification

	Identity	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	36.11	38.30	39.63	41.62	42.14	46.75	51.82	52.60
2	Non-Indigenous Male	9.92	10.04	10.34	12.25	14.77	20.35	24.86	25.60
3	Indigenous Female	30.19	32.26	33.42	33.78	36.35	42.68	49.22	53.62
4	Non-Indigenous Female	6.57	6.92	7.43	9.45	13.22	21.12	28.44	33.46

Table 22: The share of the population that does not have year 12, does not have post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	60.27	60.74	59.80	58.94	58.85	57.28	52.99	44.57
2	Non-Indigenous Male	83.05	86.55	87.41	86.46	84.96	82.40	76.87	62.12
3	Indigenous Female	54.49	54.67	55.93	58.40	58.71	54.89	48.98	37.78
4	Non-Indigenous Female	79.20	77.25	77.42	78.63	79.01	76.38	68.57	51.43

Table 23: Employed percentage grouped by age, gender and Indigenous status

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	41.51	43.53	44.40	43.27	43.06	41.34	37.51	28.87
2	Non-Indigenous Male	58.44	66.11	69.32	69.46	68.14	64.86	57.71	41.51
3	Indigenous Female	27.01	25.87	27.09	30.49	32.43	29.96	25.21	17.18
4	Non-Indigenous Female	45.57	39.51	37.38	40.53	42.96	41.04	33.79	21.18

Table 24: Full-time Employed percentage grouped by age, gender and Indigenous status

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	85.14	87.02	87.95	85.20	82.27	77.35	71.73	55.35
2	Non-Indigenous Male	88.29	90.52	90.85	90.24	89.27	87.17	81.69	66.03
3	Indigenous Female	74.17	74.57	75.29	76.86	78.44	73.80	68.13	53.57
4	Non-Indigenous Female	80.66	78.25	78.68	80.91	82.35	81.04	74.64	58.35

Table 25: The employed percentage for those who have a Bachelor's or above

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	85.14	87.02	87.95	85.20	82.27	77.35	71.73	55.35
2	Non-Indigenous Male	88.29	90.52	90.85	90.24	89.27	87.17	81.69	66.03
3	Indigenous Female	74.17	74.57	75.29	76.86	78.44	73.80	68.13	53.57
4	Non-Indigenous Female	80.66	78.25	78.68	80.91	82.35	81.04	74.64	58.35

Table 26: The employed percentage for those who finished year 12, have a post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	63.90	61.54	62.63	63.94	62.77	58.83	53.80	44.28
2	Non-Indigenous Male	76.66	79.88	81.24	81.26	80.91	78.69	72.24	57.93
3	Indigenous Female	52.69	50.88	53.35	56.26	57.94	52.92	48.37	34.38
4	Non-Indigenous Female	69.80	66.62	67.38	70.17	72.93	70.61	61.79	45.19

Table 27: The employed percentage for those who finished year 12, does not have post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	79.19	80.31	77.96	74.40	74.12	73.22	68.55	56.50
2	Non-Indigenous Male	88.34	89.64	88.89	87.82	86.84	85.85	81.32	66.02
3	Indigenous Female	58.94	62.76	64.98	66.28	69.27	67.12	63.29	51.82
4	Non-Indigenous Female	71.30	71.22	73.54	76.31	78.68	79.22	74.26	59.47

Table 28: The employed percentage for those who do not have year 12, have post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	37.63	38.42	36.99	38.37	41.32	43.28	42.36	37.27
2	Non-Indigenous Male	65.55	67.79	67.64	67.54	68.45	69.95	67.18	55.37
3	Indigenous Female	26.08	27.40	29.44	33.56	35.14	37.60	34.37	26.92
4	Non-Indigenous Female	47.69	49.40	52.05	56.20	60.68	63.22	57.68	42.37

Table 29: The employed percentage that for those who do not have year 12, does not have post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	67.20	73.67	74.34	73.98	71.01	65.75	59.01	42.19
2	Non-Indigenous Male	63.94	73.26	77.70	78.89	78.33	75.34	66.60	45.96
3	Indigenous Female	57.37	52.51	51.11	56.78	59.89	57.59	50.20	35.10
4	Non-Indigenous Female	57.02	47.79	43.95	48.46	52.89	52.86	46.00	28.94

Table 30: The full-time employed percentage for those who have a Bachelor's or above

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	63.83	67.03	70.76	66.46	62.60	58.05	54.78	36.48
2	Non-Indigenous Male	63.16	69.19	71.45	72.01	71.63	68.76	61.41	44.24
3	Indigenous Female	39.02	36.99	37.48	42.91	45.68	43.87	38.43	26.78
4	Non-Indigenous Female	40.08	35.39	34.64	38.35	42.04	41.85	35.86	23.66

Table 31: The full-time employed percentage for those who finished year 12, have a post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	41.37	41.07	44.87	47.84	45.85	42.16	35.71	27.86
2	Non-Indigenous Male	48.17	56.09	60.53	62.34	62.64	59.60	52.51	38.17
3	Indigenous Female	24.92	23.19	25.53	28.43	31.72	28.93	24.23	16.91
4	Non-Indigenous Female	35.40	32.11	31.41	34.45	37.99	37.27	30.38	19.17

Table 32: The full-time employed percentage for those who finished year 12, does not have post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	57.93	60.92	59.48	55.48	55.04	54.20	49.44	37.13
2	Non-Indigenous Male	69.27	71.47	71.21	69.96	68.94	67.47	61.40	44.39
3	Indigenous Female	25.25	26.82	30.17	32.53	36.05	35.81	32.31	24.05
4	Non-Indigenous Female	30.35	27.92	29.83	33.85	36.74	37.27	32.87	22.26

Table 33: The full-time employed percentage for those who do not have year 12, have post-school qualification

	Status	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years
1	Indigenous Male	23.17	24.47	24.40	25.47	28.27	30.12	29.01	24.18
2	Non-Indigenous Male	44.24	47.96	49.14	49.61	50.53	52.13	48.83	37.01
3	Indigenous Female	9.14	9.77	11.59	13.64	15.77	16.58	15.27	10.54
4	Non-Indigenous Female	20.06	19.98	21.57	24.37	28.51	29.80	25.07	15.84

Table 34: The full-time employed percentage that for those who do not have year 12, does not have post-school qualification