Education, Child Bearing, and the Mother's Personal Income

Langwen Guan, Yuhang Ju, Zike Peng

October 19, 2020

#Education, Child Bearing, and the Mother's Personal Income

By: Langwen Guan, Yuhang Ju, Zike Peng

19/10/2020

##Abstract

This study documents an analysis of data from the Canadian General Social Survey of 2017. The data provides ample data with significant variables which affect a family. Using this data, this study explores the relationship between the age at which Canadian women give birth to their first children, their levels of education, and their level of personal income. The study used a linear model to explain these relationships between these factors. The study is a contribution to the understanding of gender relations in the society.

##Introduction

Personal income is a very important factor in people?s lives. It enables one to be independent and to an extent, free. Typically, this matter is more complex for women due to the issues of childbirth and education which impact personal income differently. These two factors also have a peculiar relationship with each other. Childbirth is one of the most important events in a woman?s life. In addition to bringing a young one to life, birth also changes various factors of the mother?s life. Some of these changes are temporary and so they subside with time. For instance, weight gain which occurs after birth reverses shortly afterwards. However, other changes have a lasting impact. They include the academic performance and professional competitiveness of the new mothers. As such, they usually limit the set of opportunities available to them. Education usually has an opposite effect on a person?s professional and economic life. Education provides one with many opportunities to fulfill professional and financial needs. As a result, people who are more educated tend to generate a higher income and have a higher quality of life. This report documents a study which explains the relationship between the age of a mother when she gives birth to her first born, her education, and her income. It is significant to evaluate the opportunities presented to women and specifically, mothers. The study seeks to answer these research questions: 1. How does a woman?s income vary with the age at which she got her first born? 2. How does her education vary with the age at which she gave birth to her first born? 3. How does her education impact her income?

##Data

The data used to study this relationship was obtained from the Canadian General Social Survey of 2017. The data was made from an interview or a questionnaire which were filled by individuals about themselves. This dataset contained a large amount of data about various aspects of family. In essence, the dataset had 20,602 rows of data divided into 80 categories. These categories included the

respondents? age, occupation, educational level, incomes, and more information about themselves, their partners, and their children. The respondents were residents in Canada aged between 15 and 80 years of age. This means that each one of them could provide reliable information about himself or herself. The dataset contained both qualitative and quantitative data since it was created to be as exhaustive as possible. It was then coded into .csv format in order to summarize it in concisely. The information was coded appropriately in order to facilitate easy and effective analysis of the data. However, the data was limited to the Canadian demographic. It also lacked the aspect of change with time and was only useful for a cross section study of the variables of age at first birth, personal income, and education. The change in the variables with time is crucial to identify the progress the country is making towards supporting women as an effort towards achieving gender equality in all matters. Further, the respondents were provided with choices in giving their responses, while this increased the efficiency for the research process, it made the analysis less effective. This is because the data appeared clustered although it was not actually clustered in real life.

##Model

In order to analyze this data, this study isolated these variables and formatted them in such a way to facilitate the creation of a linear model. After importing the dataset into R as a .csv file, the age at first birth dataset, named {r} raw_data$achb1c in the original file, was changed from a factor into a numeric to facilitate the creation of a model. The NAs were stripped from the data before defining the linear model. A linear model was chosen due to the direct relationships between the three variables. Furthermore, the hypotheses made by the study before fitting the model. The data was fitted to the linear model summarised in the table below. In order to understand the relationship between the independent variables, the study also included separate linear models which represented the relationship between education and personal income, and between age at first birth and personal income.

```
##
## Attaching package: 'janitor'

## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test

## -- Attaching packages -------------------------------------------------------------- tidyverse

## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.1
## v tidyr   1.1.1     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0

## -- Conflicts -------------------------------------------------------------- tidyverse_confl:
## x dplyr::between()   masks data.table::between()
## x dplyr::filter()    masks stats::filter()
## x dplyr::first()     masks data.table::first()
## x dplyr::lag()       masks stats::lag()
## x dplyr::last()      masks data.table::last()
## x purrr::transpose() masks data.table::transpose()

## Warning: package 'tinytex' was built under R version 4.0.3
```

###Linear Model Representing the Relationship Between Personal Income, Age at First Birth,

and Education

```
##
## Call:
## lm(formula = income ~ fbirth + ed)
##
## Residuals:
##     Min     1Q  Median     3Q     Max
## -1.8780 -1.2702 -0.3721  0.6526  3.7349
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.4061854  0.0177267 135.738  < 2e-16 ***
## fbirth      -0.0014660  0.0002654  -5.523 3.37e-08 ***
## ed           0.0050321  0.0007719   6.519 7.24e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.356 on 20599 degrees of freedom
## Multiple R-squared:  0.003482,   Adjusted R-squared:  0.003386
## F-statistic: 35.99 on 2 and 20599 DF,  p-value: 2.489e-16
```
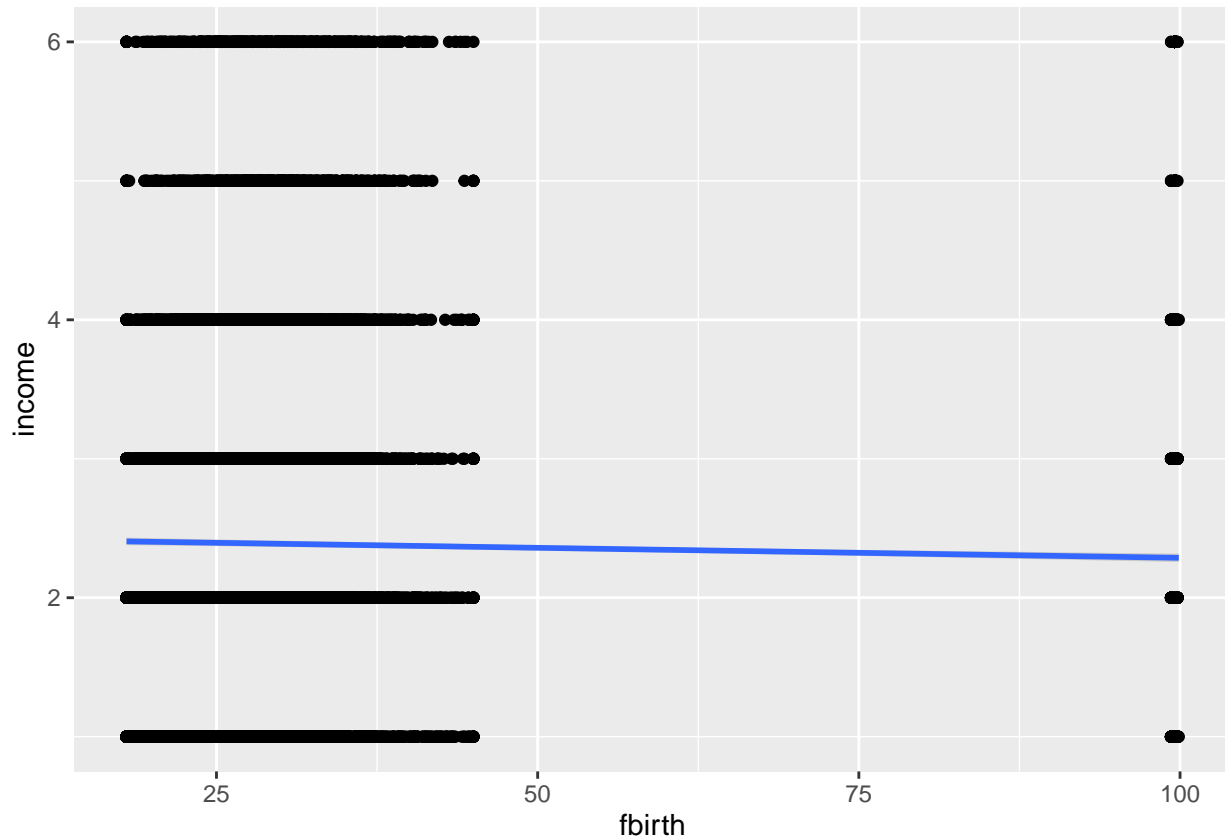
###Linear Models Representing The Relationship Between Education and Age at First Birth

```
##
## Call:
## lm(formula = ed ~ fbirth)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.469 -3.143 -1.468  0.532 93.935
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.976032   0.156205   31.86   <2e-16 ***
## fbirth      0.004935   0.002396    2.06   0.0394 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.24 on 20600 degrees of freedom
## Multiple R-squared:  0.000206,   Adjusted R-squared:  0.0001574
## F-statistic: 4.244 on 1 and 20600 DF,  p-value: 0.0394
```

In making this model, the study did not split the income or education vector into males and females. The idea behind the unified vectors was that in Canada, gender equality is maintained in education. In personal income, the distribution is random between men and women. In essence, the study assumed that sex does not determine a person?s income or education. After creating the model, the study cleaned the data and came up with a new dataset in .csv format. This dataset represented qualitative data as such. It also had easier to understand column titles in order to facilitate the interpretation of the data. While it is usually advisable to clean the data before creating the

regression model, the cleaned dataset was in characters, making it difficult to analyze. Results The study discovered a direct relationship between both education and age at first birth; and personal income. As the graph of income against age at first birth depicts, these two variables have a direct negative relationship. This means that women who have children at a younger age generally make more money than those who have children later on in life. Conversely, education showed a direct positive relationship with personal income. This means that women who have achieved higher levels of education generally generate more income as compared to women who have not achieved education at such levels.

`geom_smooth()` using formula 'y ~ x'

##Discussion

The study revealed a relationship between personal income, education, and age at first birth which has not been explored before. First, the gradient in the linear model of the relationship between education and personal income represented the unitary improvement in personal income with a unit increment in the educational level. The gradient was positive, implying that higher education leads to higher personal income. However, the gradient was very low. This means that higher education is not sufficiently rewarded. However, the study appreciates that this disparity may have arose from non-occupational income. For instance, investment, inheritance, and other supplementary income. Second, the study proved the hypothesis that women who have children at an early age get a lower personal income. However, the study could not prove a causal relationship between the data. As such, instead of having a child at an early age leading to a smaller personal income later on in life, the data may have shown that women who have high personal incomes prefer to have children at a

young age. Third, the study revealed that there was no significant relationship between education and women?s age at first birth. This means that generally, a woman?s level of education does not influence or inform her decision to have a child at a young age or later on in life. Conversely, having a child at a young age or at a relatively advanced age does not influence a person?s academic endeavors. This study was centered on women?s educational progress, the age at which they gave birth to their first children, and their personal income, and the relationship between these variables. It established that the relationships between personal income and both the level of education and the age at first birth are direct and linear. Although this study was successful in creating a reliable body of knowledge, it was limited in how much data it could make use of. In future, it recommends that a longitudinal study be conducted to investigate the trend of these variables with time. The longitudinal study will be instrumental in determining whether Canadian women can maintain the balance they have between these three issues going into the future. Further, the study should look more closely to represent how these variables vary in different provinces in Canada. Indeed, many more studies should be conducted to establish whether women in different countries are able to navigate the relationship between their age at first birth, their educational level, and their personal income. This way, the issue of gender equality can be understood and tackled as the global problem that it is. This would pave way for the academic discourse towards creation of solutions which foster more holistic gender equality all around the world. The academic discourse would enable various authorities to implement policies which protect women and look out for them.

This code can be found on github in this public repository: https://github.com/Juyuhang/Problem-Set-2.git