How Socioeconomic Factors Affect Average Age of First Birth in Canada*

Sevnur Kulak   Dongqi Bi   Ivan Li

17 March 2022

Abstract

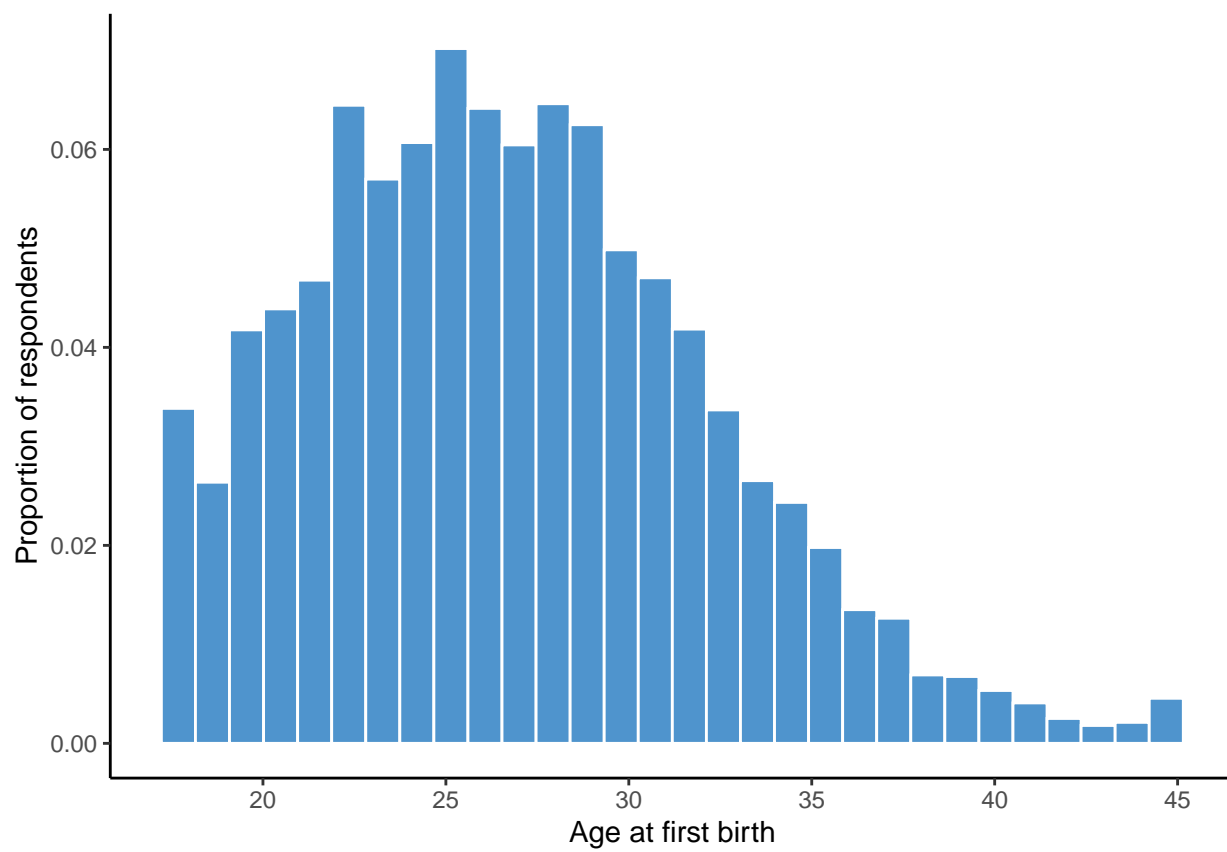write abstract here

Contents

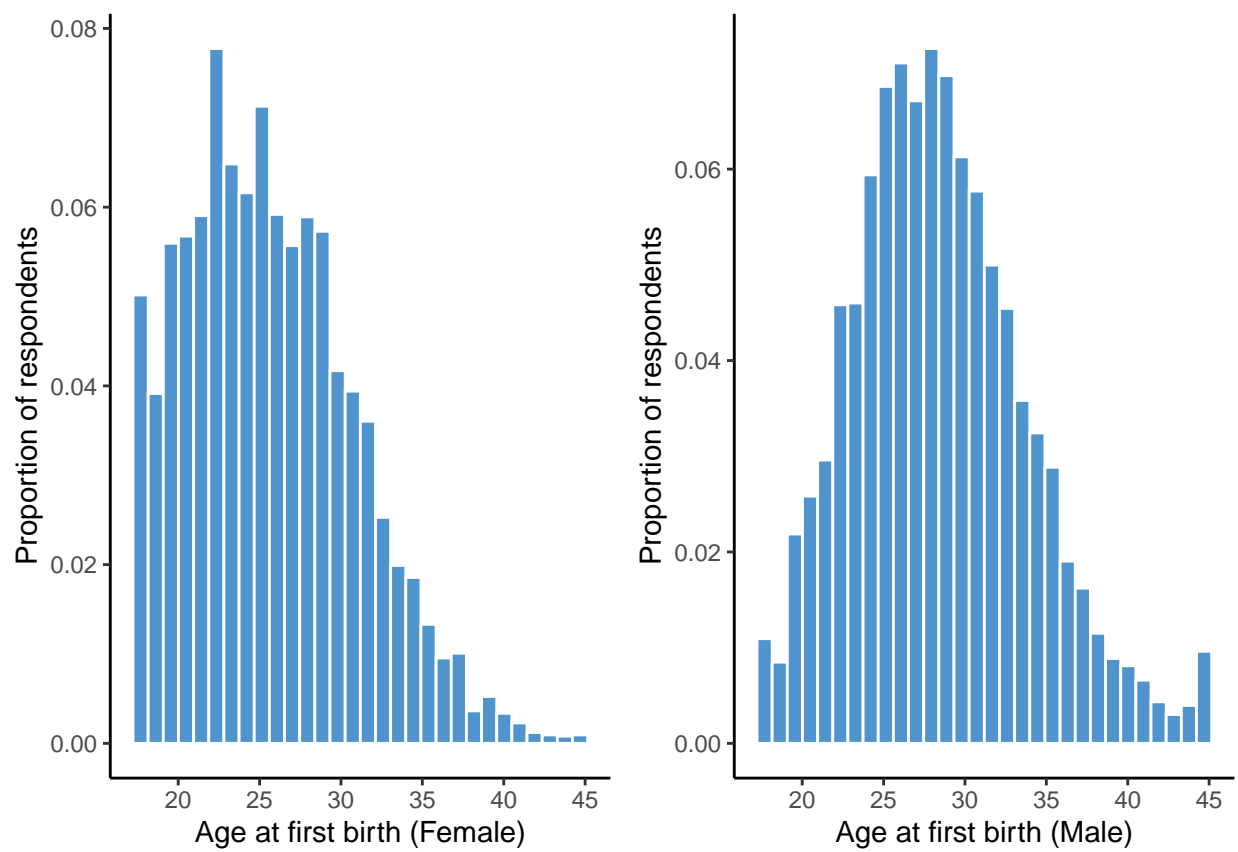Figure 1: Distribution of age at which respondents have their first child

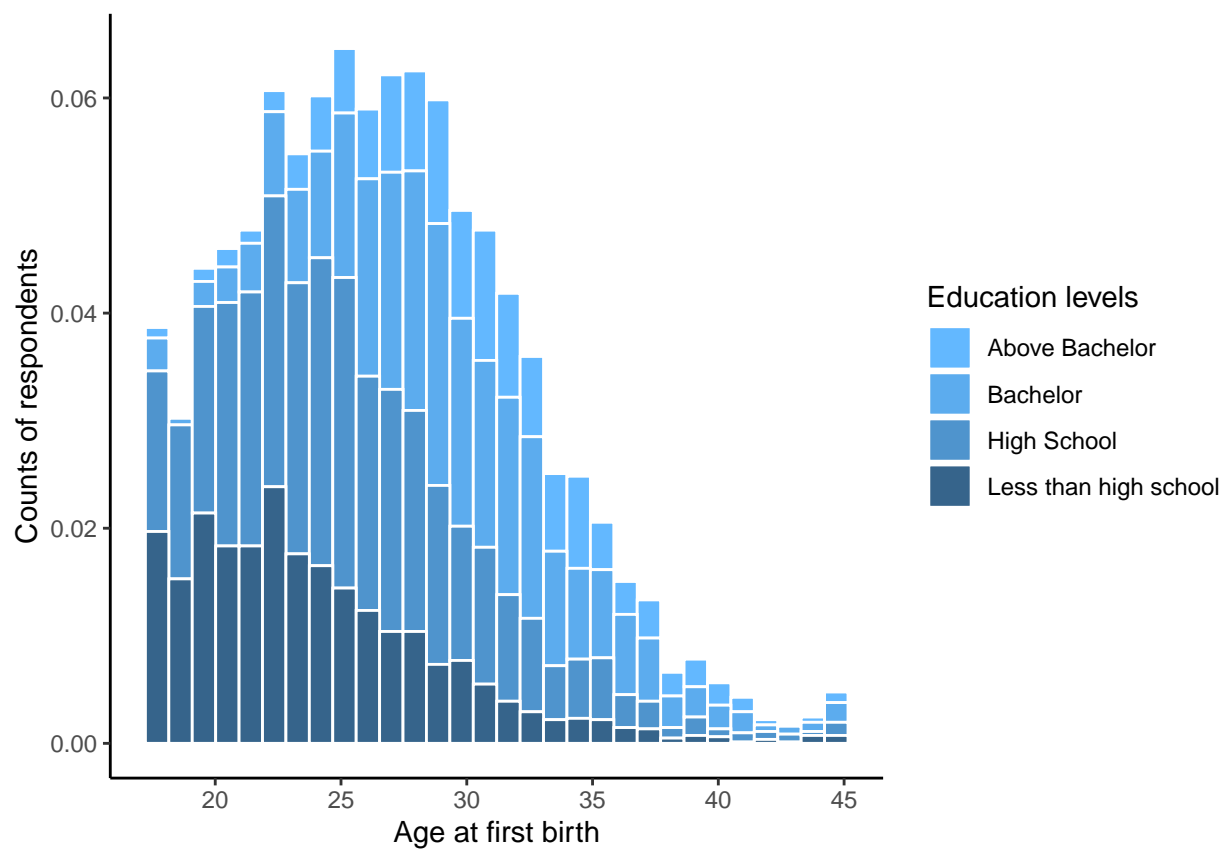Figure 2: Distribution of age at which respondents have their first child by gender

Figure 3: Distribution of age at which respondents have their first child by education levels
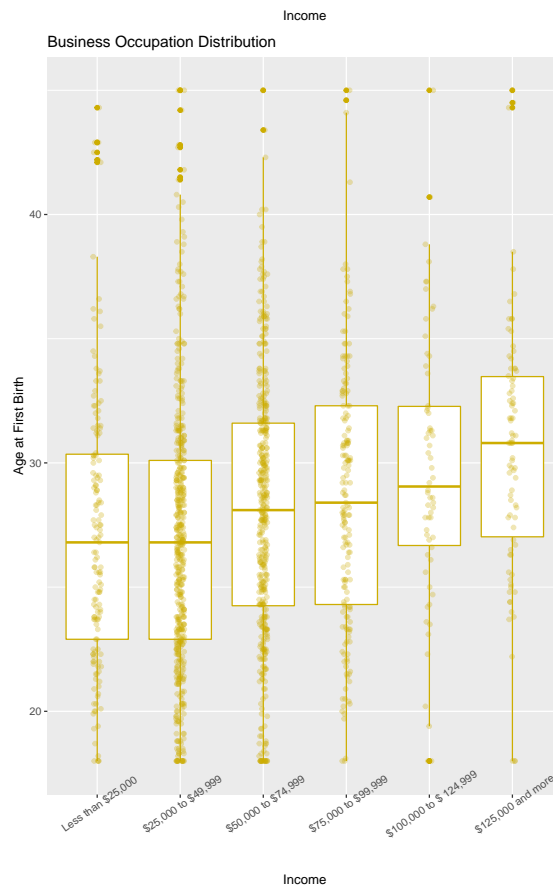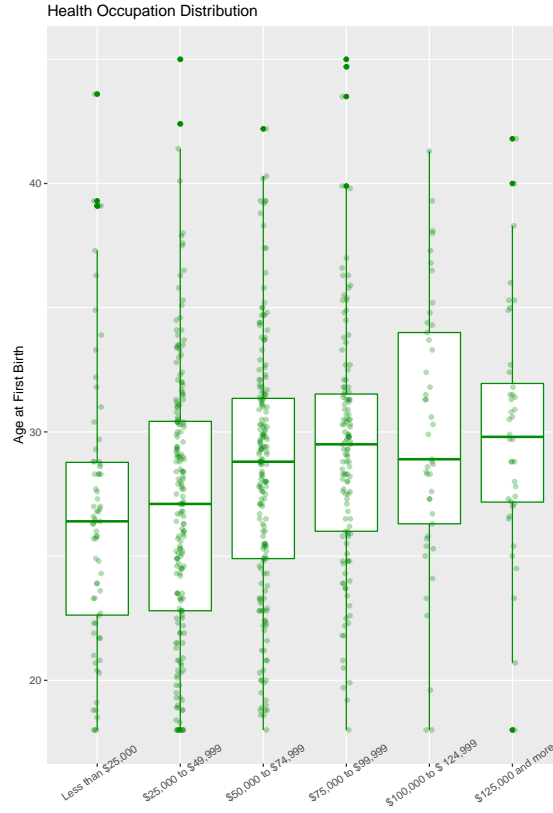
Figure 4: Part1 : Distribution of age at which respondents have their first child by occupation and income
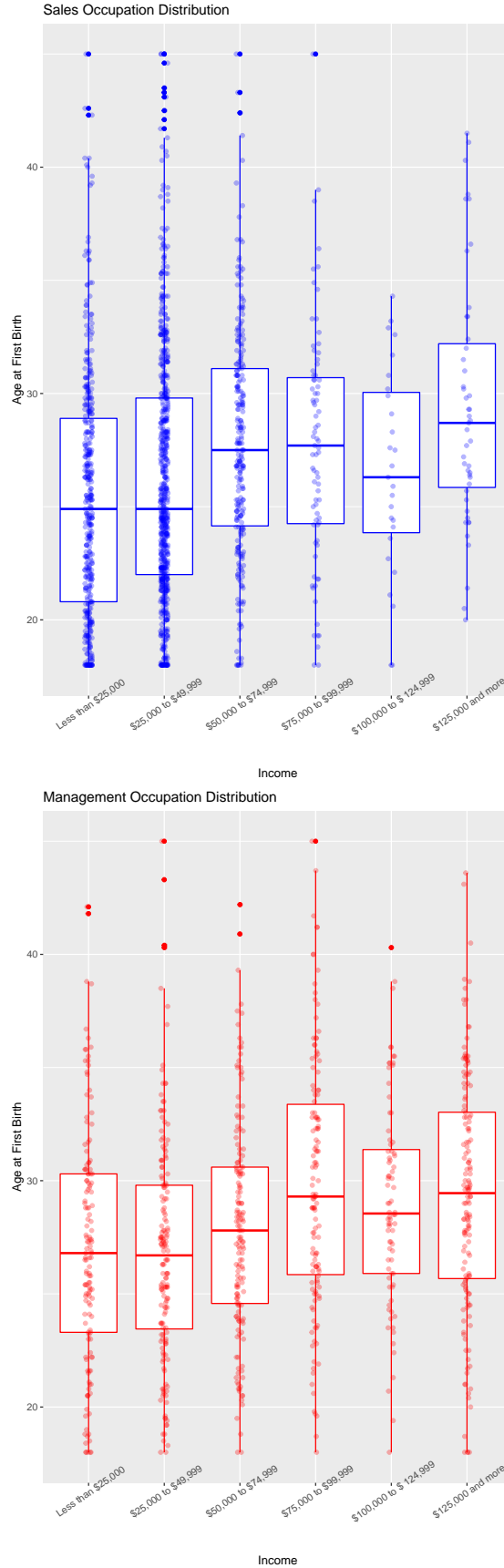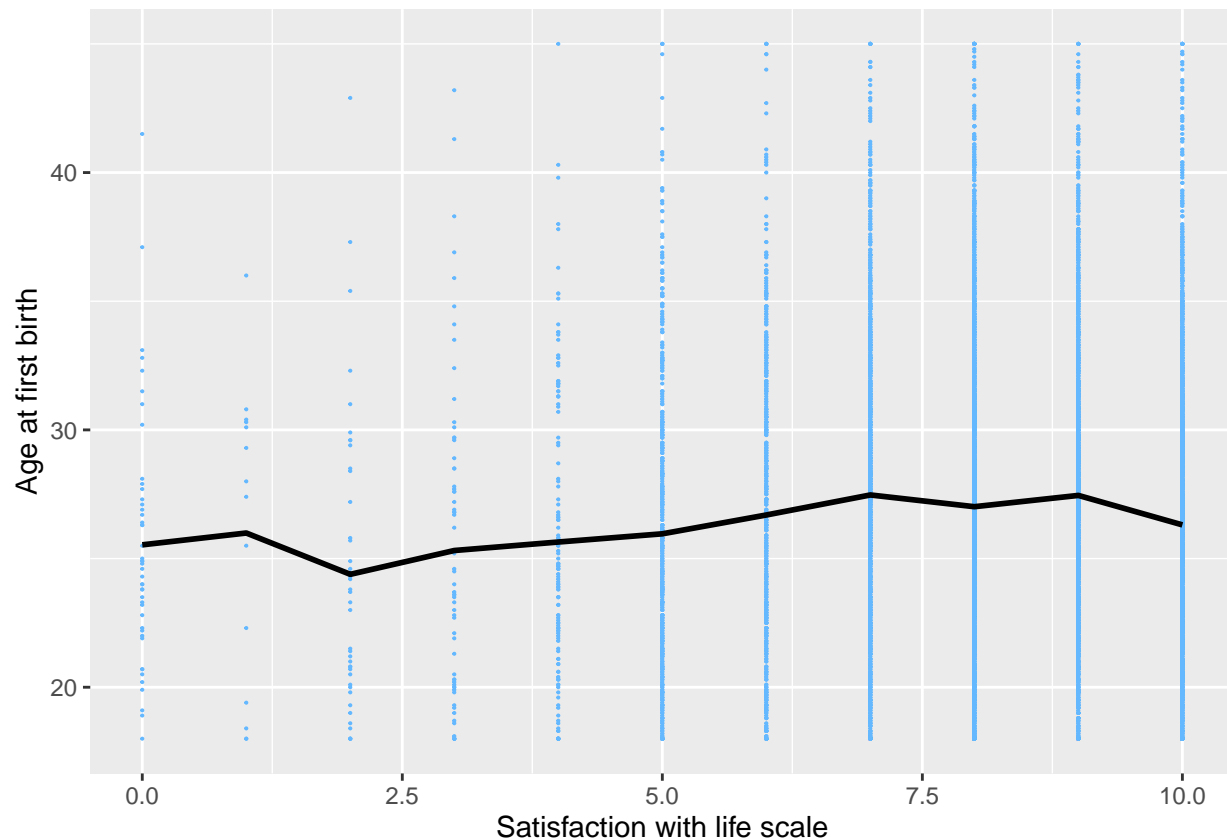
Figure 5: Part2 : Distribution of age at which respondents have their first child by occupation and income

1 Introduction

2 Data

The dataset we chose is collected from the Canadian General Social Survey (GSS) Cycle 31 Main Survey - Family. The GSS program was developed in 1985 in Canada with six survey themes, each focusing on a different topic in depth. (reference 1) The six themes are caregiving, families, time use, social identity, volunteering, and victimization. (reference 1) Each of them is used to aim at a different aspect of improving Canadian's well-being, informing social research, and acting as an effective training tool for quantitative analysis. (reference 1) Our paper is interested in the socioeconomic factors that could potentially affect the average age of first birth in Canada. Since our subject is closely related to the parental history, household information, education and health information of Canadian families, and these are the survey's subjects; we choose the "Families" theme over other themes. It contains all the variable information that we needed.

In particular, we choose Cycle 31 of the GSS Families survey; The data is collected from 2017-02-01 to 2017-11-30 with a target demographic of all non-institutionalized people aged 15 and up who live in Canada's ten provinces. (reference2). The data is collected using a cross-sectional design, combing landline and cellular telephone numbers. (reference2) The selection process of the data collection used a stratified sampling method, a method of dividing a population into smaller groups called strata and selecting random samples from each strat (reference3). In this survey, the strata are at the province level; proxy replies are not authorized for this survey.

Our response variable is "age at first birth," which is the respondent's age when they have their first child. We used other variables in this paper as factors to analyze their relationships with the response variable is gender , including females and males. Income levels which contains 6 levels: Less than $25,000, $25,000 to

---

*Code and data are available at: https://github.com/hellowoshibb/Age_of_First_Birth_Analysis

$49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 to $124,999 and $125,000 and more. Occupation types include heath occupation, business occupation, sales occupation and management occupation. And the respondent's rating about the level of satisfaction with life scaling from 1 to 10.

Figure1, we have the distribution of age at which respondents have their first child. The x-axis is the age at which respondents have their first child, and the y-axis shows the percentage of the respondents in this age group of the entire sample. The age group is divided evenly into small groups with individual bars representing them from age 18 to age 45. Each bar is approximately representing one year of age. At two endpoints of the graphs, the first bar represents age 18 and under age 18, and the last bar represents age 45 and above age 45. The heights of the bars show the popularity of this age group. Figure 2 is similar to figure 1; the difference is that we divide the distribution graph by gender, and we use it to analyze how gender as a factor influences the age of people who have their first child. Figure 3 is another distribution graph; we split the distribution into different education levels, we can see the portion of people in each education level in a diverse age group. For instance, at age 25, most people have their first kid in a high school education level, and a tiny portion of people have above bachelor education. In figure 4 and figure 5, we have four boxplots. Each boxplot analyzes data based on one type of occupation; we have heath, business, sales and Management. Each boxplot also contains six boxes representing an income level from less than $25,000 to more than $125,000. Within each box, it draws from the first to the third quartile of the data of age at which respondents have their first child. At the median, a vertical run through the box. We also add raw data points to our boxplots to better visualize the distribution of data points. In our last graph, figure 6, we have x-axis as the respondents rating regarding their satisfaction with life from 1 to 10, and y-axis as the age of first birth. The trendline shows the average age at first birth with each satisfaction rating with life, and we also add raw data points to this plot to visualize the spread of sample points.

3   Results

4   Discussion

5   Survey

6   References

7   Appendix