Project Report

Rajvi Jasani

Data:

This study uses data from the National Survey of Family Growth (NSFG) 2017–2019, a nationally representative dataset of individuals in the United States aged 15–50. This study makes use of the Female Respondent file (6141 obs.) and the Male Respondent file (5206 obs.). The files were provided as ASCII data along with a SAS Program Statement which has been used to import data into R. These datasets include many variables of which this study focuses on those related to demographic, socioeconomic and future birth intentions information. Table 1 in the Appendix summarizes the variables of focus for this study. All the categorical independent variables were considered as factors with meaningful reference categories. Additionally, sampling weights (WGT2017_2019) and design variables for stratum (SEST) and cluster (SECU) have also been used to ensure that the estimates reflect the U.S. population accurately. The functions of the survey package in R have been used to create survey designs and perform the rest of the analysis.

Research Question:

How do age, education level, relationship status and employment status influence men's and women's individual intentions regarding future children?

Approach:

R code, data, and other supplementary information can be found on the GitHub repository.

We start by looking at some basic descriptive statistics for gender wise distribution of age, marital status, education level, employment status and intentions for future births to compare how they distribute within female and male datasets. Refer to Graphs 1-5 in <Appendix>. All the summaries and graph show that there are no data inconsistencies that might cause potential logical errors in our analysis. These also help us set the most common categories as reference for the regression models.

To answer the research question, we compare individual effects of the independent variables on the dependent variable in both the datasets individually. We look at the summary and visual statistics to get inferences on how these groups may be individually affecting a person's intentions for having (more) children.

However, in real world, there is not just one factor that affects such decisions. So, we design a regression model to check the influence of each independent variable in a combined setting. We design core models for each gender, which include the universal variables (AGER, RMARITAL, HIEDUC, LABORFOR). We also design extended models for each gender to analyze the effect of the employment status of a spouse/partner (where applicable).

After observing the data, we notice that the observations for "Uncertain" intent as very small in comparison to the total data size so for our regression analysis, we have removed such observations. Our response is a binary variable thus we use a logistics regression model. We design a quasibinomial logistic regression model here. A quasibinomial is like a binomial logistic regression model, except that it accounts for overdispersion by adjusting standard errors. Overdispersion occurs when the variance of the outcome variable is larger than expected

under the binomial model. This could happen in our case because we have clustered data (survey data with complex sampling designs).

Results:

First, we shall discuss the results from individual comparisons to birth intentions:

Age Distribution and Birth Intentions: Refer to Graph 6 in Appendix. Younger men and women are more likely to intend to have children, but this intent declines significantly with age, especially after their 30s. The likelihood of not intending to have children grows with age for both genders, with a sharper and faster decline observed among women.

Marital Status and Birth Intentions: Refer to Graph 7 in Appendix. Married men strongly intend to have children, with few "No" responses, while cohabiting men also lean towards having children, though to a lesser extent. Single, separated, and divorced men are less likely to intend to have children. Married women also show a strong intention to have children, with cohabiting women showing similar trends, though slightly lower. Single and separated/divorced women are less likely to intend to have children, possibly due to economic or personal concerns.

Educational Level and Birth Intentions: Refer to Graph 8 in Appendix. Both men and women with higher education are more likely to not want to have children. Men with lower education levels are slightly more decisive compared to women in the same category.

Labor Force Status and Birth Intentions: Refer to Graph 9 in Appendix. Men in the labour force (working or seeking work) are more likely to intend to have children than those not in the labour force. Among men not in the labour force, there is a noticeable increase in "No" responses, suggesting employment may influence their childbearing intentions. Women in the labour force also show a higher likelihood of intending to have children, though the trend is less pronounced than for men. Both genders show a clear link between being in the labour force and intending to have children.

Spouse/Partner Work Status and Birth Intentions: Refer to Graph 10 in Appendix. Partner employment status appears to have a slight positive influence on the intent to have children for both men and women, though the effect is relatively small.

Next, we shall discuss the results from the regression models:

Each coefficient for the regression models represents the log-odds change in the probability of "Yes" relative to a "No" for one-unit change in the corresponding predictor, holding other predictors constant.

Female Regression Models: A snapshot of the model summary can be seen in Figure 1 in Appendix. We notice that age, higher education levels and stable employment status significantly affect the intent for future births in women. Age and a stable employment status positively affect the decision to have children. Higher education levels, such as Master's, Doctorate, Professional degrees are associated with lower childbearing intentions. However, lower education level graduates such as 9^{th} grade or less, 10^{th} grade graduates have more chances of wanting children. Interestingly, marital status does not significantly affect childbearing intentions. All in all, these results align with the results we saw in the individual

comparisons but the regression model helps us to show each's significance. The odds ratio Plot 1 in Appendix visually summarises these and shows the confidence intervals.

As can be seen in Figure 2 in Appendix, inclusion of spouse/partner's work status variable in the extended model shows minimal additional explanatory power. Women with employed partners slightly increase their childbearing intentions, but the effect is not highly significant and it also tends to affect the significance levels of other, previously significant variables.

Male Regression Models: A snapshot of the model summary can be seen in Figure 3 in Appendix. Similar to the female model, we notice that age and higher education levels significantly affect the intent for future children. Additionally, being married positively and significantly affects intentions for having a child in men. However, interestingly, labour force status does not show any significant effects which logically does not feel sound. It also does not align with the individual comparison findings. This is something that can be explored further in detailed studies. The odds ratio Plot 2 in Appendix visually summarises these and shows the confidence intervals.

As can be seen in Figure 4 in Appendix, similarly to females, partner employment has limited impact in men too.

Conclusion:

In conclusion, this study helps us uncover many trends underlying in the data. Initial visualizations help us notice these individual trends and the regression models help us uncover how much significant do these trends remain in a combined setting. We saw that women's intentions increase more sharply with age compared to men, though both show positive associations. High education levels suppress childbearing intentions in both men and women. Marriage boosts intentions more for men, while being single or cohabiting has minimal effects for both. Part-time or studying women show increased intentions, while men's employment type has limited influence. However, this study in no way can be called complete. There are many statistical tests and analysis that need to be done to be more confident in these results. This can be the subject of a next version.

Attribution of Sources:

This study was created using help from external sources listed below:

- Referred to survey documentation and ChatGPT for help with syntax and useful functions.
- Used ChatGPT to find an appropriate regression model. It suggested using the quasibinomial family given the nature of the data.
- Referred to ggplot2 documentation an ChatGPT for help with syntax of plots.