Impact of Having Children on Women's Careers

Term Paper for ECO475 - Applied Econometrics II

Qien Song, Yexuan Shen

Supervised by Professor Ismael Mourifié

& Teaching Assistant Atom Vayalinkal

Department of Economics

University of Toronto

December 6, 2021

1 Introduction

1.1 Background

The branch of social sciences studying the relationship between family and the labour supply have long been a subject of key interest to social scientists including both economists and demographers. A major topic within this subject is the effect of motherhood on employment outcomes. Much empirical research has consistently shown that an income penalty exists for childbearing mothers who return to work afterwards [Angrist and Evans, 1998][Correll et al., 2007][Keating, 2002]. A common theory proposed by past research is concerned with the time allocation constraint of mother between household and career duties. The natural tension between maternal and work responsibilities is an unavoidable dilemma that working mothers have to face. In many circumstances, mothers consciously shift their focus away from employment duties towards child-caring activities, which stalls their potential income growth. Another plausible explanation for the phenomenon of income penalty is based on societal perception of the role of mothers. Past research has shown that employers may hold innate bias against pregnant women and consider them as less reliable, less rational and less committed to work [Halpert et al., 1993]. The same prejudice may exist for working mothers in general, and give rise to a wage penalty.

While existing research papers have generally focused on maternal effect on employment outcomes in developed nations, specifically the United States and Scandinavian countries where a plethora of high quality data is available, this paper aims to explore the relationship between childbearing and employment outcome in China, which has a lower level of economic development. Understanding how having children affect a mother's income in developing nations holds both theoretical and practical significance. Firstly, it can help economists to understand whether the penalty is universal across countries of different economic levels. Secondly, developing nations typically experience higher levels of economic growth. Modeling the effect of having a child on labour market outcomes in a high growth macroeconomic environment can generate additional insights. Lastly, women's labour force participation rate has experienced a sharp decline in China despite consistently high economic growth. The country's female labour force participation rate has dropped from 73.24% to 60.57% from 1990 to 2020 [WorldBank, 2021]. As most current research is completed on countries with either steady or increasing women's labour force participation, this phenomenon represents a unique socioeconomic situation that is valuable to be investigated.

1.2 Research Question

Having a child could affect women's earnings in numerous ways, through discrimination imposed by society, cognizant personal decisions, or an ambiguous mix of both. Employer discrimination against working mothers may arise due to low expectation of mothers' commitment to work. Mothers may be viewed negatively in terms of their expected work focus [Poduval and Poduval, 2009] due to the perception that they have to take on childcare duties. Employers may make career promotion decisions based on the expected future output of employees and mothers are naturally disadvantaged as they are more likely to be viewed as family-oriented rather than career-oriented.

Employers may perceive working mothers as more negotiable due to their perceived focus on family, and thus offer them a below-market wage. In certain cases, mothers may also be willing to accept a below-market wage as they understand their commitment to work may not be as high as those who are non-mothers. Therefore, the wage penalty may be a reflection of an equilibrium mutually agreed upon between the employer and working mother. From mothers' perspective, they may actively pursue employment opportunities that are closer to home and require less time commitment so they could spend time raising their children.

Another dimension to this particular topic is the marginal effect of having a child. Under reasonable assumptions, the income effect of having an additional child is not constant. This paper limits the scope of investigation to the marginal effect of having a second child for a better controlled setting. Firstly, the employer discrimination effect largely diminishes in a second child case as employers' distinguish between mothers and non-mothers more so than between one-child mothers and two-children mothers. Secondly, analyzing the second child effect allows for leveraging a Chinese birth-control policy to construct an appropriate instrumental variable, which is elaborated further in the methodology section.

In the construct of the generic question concerning childbearing's effect on personal income, there exists a double endogeneity issue. Endogeneity is defined as a situation where both the independent and dependent variable are correlated with an unobserved factor. In this particular case, the decision to have a child and the decision to work are both endogenous to personal income as many factors such as education level, spousal income can lead to causation simultaneously. However, monitoring a double endogenous problem presents additional methodology challenges that may lead to inaccurate analysis of either variable. To simplify the question, this paper examines the income of working mothers only to focus on evaluating the selection to motherhood problem. Finally, this paper constructs the research question as the following:

How does having a second child affect a working mother's income in percentage terms?

Given a Chinese context, this question presents high relevance as the contemporary period represents a transition away from the traditional patriarchal society. For policymakers, the research result could bring a renewed understanding of current legislation surrounding women's maternity decisions in the workplace, and whether any deficiencies or loopholes exist. In addition, this paper looks at a more rural subset of the general population, who are more economically left-behind and receive less attention from lawmakers and the general public.

1.3 Summary of Paper

To briefly summarize this paper on mother's income with a focus on the second-child effect, we used data from China Family Panel Studies in 2010 and conducted Two-Stage least squares analysis to explore the causal effect. To solve the endogeneity problem of fertility, the gender of first child was used in this paper as an appropriate instrumental variable, based on the special one-and-half-child policy in China.

From the IV regression results, we have found that bearing a second child has a negative effect on women's income. Keeping all other variables constant, having a second child approximately decreases income by 42.48% on average, compared to those who do not have a second child. Although the result is not statistically significant at the 5% significance level, it exhibits a marginal trend towards significance with a p-value of 0.07. We interpret the result as having a second child does not have a positive effect on mothers' potential income but the exact impact of fertility decision of having a second child remains unknown. Certain limitations such as income ambiguity, internal migration and marriage culture may have prevented the model from reaching a precise conclusion.

2 Literature Review

The decision to have children has been leading to reduced opportunities and increased discrimination for women in the labor market, which usually is reflected by lower average earnings and fewer women being in the labor force compared to men. These phenomena may be attributed to the fact that lots of companies tend to hire fewer female employees or women who want to spend more time on children tend to work less or choose simple jobs with low wages. There are a large number of previous studies on the relationship between having children and women's careers,

which have estimated the causal effect of motherhood on female labor market outcomes, such as the finding of a large and long-lasting negative effect on earnings [Lundborg et al., 2017].

A new instrumental variable (IV) approach based on in vitro fertilization (IVF) was introduced in the study by Lundborg, Plug, and Rasmussen in 2017. IVF is defined as a medical intervention to help infertile women to get pregnant and have children. This previous study uses data on IVF treated women in Denmark and constructs a two-stage least-squares model to estimate the causal impact of fertility on women's earnings. It states that IVF treatment success makes a plausible instrument for fertility because IVF treatment outcome is independent of labor market characteristics and explains exogenous variation in the likelihood to have children. Thus, success at the first IVF treatment among childless women can be used as the instrumental variable in order to solve the endogeneity problem of fertility indicator (equals 1 if a woman has children and 0 if not). Specifically, the study is interested in the causal effect of fertility on labor earnings t years after these women had their first IVF treatment, where t can be 0-1, 2-5 and 6-10 to account for all short term, medium term and long term effect. The researchers have found that there exists a negative and statistically significant effect of fertility on women's earnings no matter in the short run or in the long run, yet, the short run effect is much larger than the long run effect, which could be attributed to maternity leave taken with low compensation before and after the childbirth. Another interesting finding is that the decrease in earnings estimated for having children is much more than that for having additional children among women who already have children. This finding is closely related to our topic studying the second child effect on women's earnings, and thus we are interested to verify if this statement still holds using different data with a developing Asian country background.[Lundborg et al., 2017]

Another study conducted by Markussen and Strøm in 2015 also points out that the endogeneity problem of fertility decision makes it complicated to study the causal relationship between motherhood and labor market outcomes. This study uses miscarriage as a biological shock to fertility in order to estimate the causal impact of motherhood on female labor market outcomes the first five years after birth. It estimated the marginal effect of the first three children using multiple instrumental variables because there are more than one endogenous causal variables they want to study, such as effect of having first child, effect of having second child, etc. Data used in this study is Norwegian women who gave birth or experienced miscarriage in the period from 2001 to 2004. The main finding is that the first child has a big impact on women's income, with an average reduction of 18 percent. On the other hand, for women at the higher end of the income distribution, the impact

of having a child on their income is small. Furthermore, the study found that labor force participation is not greatly affected by childbirth, which shows that the combination of work and family obligations is obviously realizable in the Norwegian background.[Markussen and Strøm, 2015]

Overall, although previous studies have examined the significant negative causal effect of motherhood on women's earnings using various IV methods, they mostly focus on developed countries and first-child effects. Few research have been conducted on China due to the lack of household-level data prior to 2010.

In order to further supplement and advance on the basis of previous research, we plan to propose a different IV in terms of the special female labor market background and the unique restriction policy on having the second child in China. By focusing on the second-child effect, we will then compare it to the significant negative effect of first-time motherhood which has been proven by previous studies.[Mu and Xie, 2014]

3 Endogeneity Challenge & IV Strategy

3.1 Challenge

Fertility is commonly regarded as a highly endogenous variable [Angrist and Evans, 1998]. Therefore, the major challenge of this study is the endogeneity associated with fertility, where multiple factors may influence motherhood and income concurrently. Mothers who choose to have a second child and those who do not may have drastically different traits, both observed and unobserved, ranging from spousal situation, career opportunities to societal values. Previous research has found that hours worked is a representative example which both correlates with the decision to have a child and income, where the more hours worked leads to higher total income and a lower probability to have another child [Angrist and Evans, 1998]. Not properly addressed endogeneity issues mask the true effect of childbearing due to the omitted variable bias it creates and thus this may pose challenges for policymakers to properly address the issue.

3.2 Strategy

This paper proposes an instrumental variable approach leveraging a Chinese birth control policy that makes the gender of first child a suitable candidate. The one-and-half-child policy is introduced to complement the one-child policy which prohibits households from bearing a second child. This specific one-and-half-child policy gives an exemption to rural households, whose first child is a girl,

to have a second child due to strong preferences for male offspring in many rural regions across China.

The actual implementation of the one-and-half-child policy varies by province and sometimes varies by county level. In certain provinces, namely Jiangsu, Shanghai and Beijing, the policy was never introduced and the one-child policy is enforced strictly [of People's Republic of China, 2013]. In Zhejiang, the policy was not universal across all rural regions and only applied to rural households where neither member works in government or state-owned institutions [ZhejiangProvincialDepartmentofFinance, In Shandong, the policy applied to all rural households [Congress, 2014]. Besides the differentiated implementation, the enforcement also varies considerably across regions. In provinces such as Guangdong and Fujian, the policy was introduced in the late 80s but it was never enforced strictly and having two children was a common phenomenon in rural households [Tribunal, 2011].

For this particular IV to be effective, it has to adhere to the dependence relationship shown in Figure 5. Gender of the first child should be independent from other explanatory variables and the target variable, and only influences the mother's income through the decision to have a second child. From Table 1 Summary Statistics, we can see 48.31% of the first child in our data is a girl, which is very close to 1/2. With the law to prevent sex-selective abortions in China, it indicates the randomness of first child gender in China. Therefore, gender of the first child seems independent of other variables except for the decision to have a second child. We will also examine the validity of the IV using both a qualitative and statistical approach in the Results section.

4 Data

4.1 Data Overview

This paper uses China Family Panel Studies, which is a longitudinal survey collected biannually starting from 2010. This survey encompasses all aspects of Chinese society including economy, demographics, education and health. About 33,000 households participated in the initial survey in 2010 [Institute of Social Science Survey, 2015].

There are multiple approaches used depending on the nature of the data. A significant portion of the past research papers deploy a time fixed-effect model to control for endogeneity as the data typically examined are longitudinal surveys. Upon initial screening of the data, about 40% of households that responded in the 2010 survey did not respond to the survey in 2014, limiting the sample size of the data. In addition, the 2010 survey collected extra information that was not

collected in follow-up surveys so the use of panel data limited the scope of variable selection. The survey data from 2016 and 2018 cannot be incorporated into the study as a universal two-child policy replaced the one-child policy in 2015. A 4-year time-frame may be too short to capture accurate trends in income. Due to these data issues, we use the 2010 survey data as a static snapshot of the general population.

4.2 Preprocessing

This paper performs data preprocessing to select relevant variables and data entries. The detailed preprocessing procedures are available in [Python Notebook]. In order to control for some exogenous variations, the sample first restricts on female individuals with an agricultural hukou (household registration) status. This status should be distinguished from living in a rural place of residence. As a consequence of China's strict household registration law, rural migrants to urban areas in the wave of large scale internal migration, are likely to maintain an agricultural hukou status and they continue to be subjected to the one-and-half-child policy.

As the one-and-half-child policy can be viewed as a partial exemption to the one-child policy, the paper excludes populations who are subject to other exemptions. This includes couples with at least one ethnic minority (non-Han) member, and couples who are both only-child. This exclusion is necessary because both groups are completely exempted from the one-child policy, so their decision to have a second child is independent of the gender of their first child.

To further restrict the sample size, the paper examines households who are in their first marriage only. This restriction is applied because a divorced mother can marry again to have a second child, which may complicate the observations. In addition, the sample filters on mothers aged between 20 and 40, whose first child is less than 18, and who has income. After the child reaches adulthood, they become more economically independent and may not require maternal commitment to education and health anymore. Also, the age between 20 and 40 represents the peak reproductivity age. Also, it has been demonstrated that motherhood penalties tend to diminish in the long run [Lundborg et al., 2017] so this dataset restricts to the peak reproductivity age between 20 and 40. The income requirement is needed as this paper controls for the selection to work endogeneity.

The last layer of filter is on the provincial level where regions with strict enforcement are included. Table 5 illustrates the evidence of wide disparity in enforcement level across four provinces. In Shandong, amongst 30 households with a female first child, 50% decided to have a second child,

while only 1 out of the 41 households with a male first child had a second child. Similar strong enforcement is visible in Zhejiang as well. In contrast, 60% out of 80 households have a second child when the first child is a boy, 71.2% out of 83 households have a second child when the first child is a girl in Guangdong, exemplifying a weak enforcement of the policy. To filter a dataset that is a holistic representation of the Chinese population, geographic factors such as *North VS South*, *Coastal VS Inland* are considered, and 11 provinces are included in the final dataset.

4.3 Summary Statistics

Table 1 provides us a quick summary of our dataset and some major variables. With a high standard deviation, the distribution of personal income is highly right skewed. This is also evident from Figure 1. The income ranges from 160 to 150000 with the average income being 10193.24.

Most of the key variables are binary indicators with the exception of prestige score, which is on a continuous scale, and highest education, which is an ordinal variable. From the mean value of bearing a second child, about 34.11% of the women in our sample data have a second child and about 2.54% of them have a third child, which clearly reflects the birth control policy in China. The discrepancy between agricultural registry and agricultural occupation is visible from the "In Agricultural Work" variable where about 30% responded "Yes". This indicates that the respondents may also be urban residents with agricultural hukou, or rural residents engaged in non-agricultural work. We observe a mean value of 0.70 for the "Ever Took Train" variable, which aligns with our initial research aim, as this indicates about 30% within our sample has never taken a train before, so we are examining the economically left-behind households.

For exploratory data analysis, we examine the box plot of income of mothers with a second child and those without in Figure 3. A slight gap is visible with the income of mothers with a second child trailing those without. Also, we look at the scatter plot of income of mothers broken down by education level, where mothers with higher education receive higher income no matter their fertility decision in Figure 4. For detailed definition of variables see Table 2.

5 Methodology

5.1 Variable Selection

This paper proposes a 2-stage linear regression model with the target variable, log(*personal income last year*). The paper chooses to use logarithmic income upon observing the highly right-

skewed distribution of income, which may lead to a violation of normal distribution assumptions of multivariate regression. Figure 2 shows the distribution of income after log transformation which resembles closer to a normal distribution. In addition, the paper investigates the relative effect of having a second child on mothers' earnings in percentage terms. Logarithmic income allows for interpretation of coefficients as percentage change which suits the objective.

The variables are initially screened using the forward and backward selection technique in Python, and adjusted manually by adding interaction effects and filtering through statistical significance. The education level of both the respondent and her spouse are included in the final model as education may have direct effect on women's decision to have children and direct effect on a person's income. The age of the first child is selected as a proxy variable for the number of years of experience, which is not available in the original dataset. The response to 'Have you ever taken a train before?' is selected as a proxy variable for the economic status of the respondent as those who have never taken a train tend to have fewer contact with the outside world and are less likely to be affluent. Prestige score is another economic variable included as a standardized way to measure the socioeconomic level of the respondent's occupation. Lastly, a geographic variable is used to categorize provinces as northern and southern according to the definition of Qinling-Huaihe Line, with the economic structure and policy enforcement greatly differing across the country.

Interaction variables are included in the model to reduce the bias arising from multicollinearity. The variable of agriculture related work is included due to the distinction between agricultural household registry and rural residence. Although the sample includes households where both members are of agricultural household registry, this does not imply their main occupation is agriculture-related work. Due to large-scale industrialization and urbanization, they could be migrant workers who live in urban areas, or non-agricultural workers who live in rural areas. In Table 6, the correlation between the variable (In Agricultural Work) and its spousal version is significant at 0.39 as this indicates that the couple is likely to be both in agricultural production or both not in. Similarly, daily hours committed to a job is correlated with full time status, so an interaction effect is used as well. However, the interaction effect of education level is not included in the final model as education level is not a binary indicator but rather a categorical variable. Adding an interaction effect would lead to insufficient sample size in each subcategory and may overfit.

5.2 Model

We construct the final model as the following:

First Stage

Equation-1 shows the first stage regression performed by regressing the endogenous independent variable (*Bear_Second_Child*) against all the other independent variables, including the instrument variable (*First_Child_Sex*).

Second Stage

Equation-2 shows the second stage regression performed by regressing the target variable against the predicted endogenous variable obtained from Equation-1 and the other independent variables.

6 Results

6.1 IV Validity Check

Having a valid IV is crucial to the robustness of the overall model. For an instrument to be valid, it has to satisfy the following properties: causation on endogenous variable, exclusion restriction and monotonicity [Mourifié and Wan, 2017].

The causation on the endogenous variable can be verified in the first stage regression. The variable $First\_Child\_Sex$ is statistically significant with a p-value of 0, validating the policy effect on the decision to have a second child. A coefficient of 0.33 implies that holding all other factors constant, a mother with a female first child is 33% more likely to have a second child. This causality is also evident in Table 7 which shows the number of observations for each combination of first child gender and fertility decision. Given the first child is male, 47 out of 244 (19%) households have a second child. In contrast, 114 out of 228 (50%) households have a second child if the first is female. This 31% gap in decision to bear a second child largely aligns with the 33% coefficient obtained from the regression result. Notably, this decision gap is significantly larger than a similar IV based on child gender used in Angrist and Evans's study in 1998, where 56% of family with first two children with same sex had a third child, comparing with 51% whose first two children were of opposite sex [Angrist and Evans, 1998].

To satisfy the exclusion-restrictions requirement, the sex of first child should only influence

income through decision to have a second child, and should not directly influence income. Although there is no consensus on testing methods to validate this requirement, it is reasonable to assume that a mother's income is independent from the gender of her first child. Through a rich history of past research on factors that affect income, gender of the first child is not considered as a factor that holds any statistical significance [Angrist and Evans, 1998][Lundborg et al., 2017].

The last property that needs to be satisfied is the monotonicity assumption, equivalent to the negligibility of "defiers". In this particular context, this refers to households who would bear a second child if the first child is a boy, but not so if the first child is a girl. From a cultural perspective, such preference for girls is uncommon and the existence of birth law would further discourage defying behaviour. However, it is still possible that such defiers exist in a significant proportion of the general population. Therefore, a rigorous testing approach is required to check for monotonicity assumptions.

This paper uses the testing framework from Mourifié and Wan's research work, who have shown that the testable implications - monotonicity assumptions and causation on endogenous variable, can be tested using two conditional inequalities with the Chernozhukov intersection bounds framework [Int, 2013]. The test result shows that the null hypothesis is not rejected at the 1% significance level, which indicates that the sex of first child is statistically significant and is a valid instrument.

6.2 Regression Interpretation

Table 3 shows the first stage regression results, with coefficient estimates of each variable alongside with their standard error and significance level. The R-squared value of 0.3350 shows that 33.50% variation in the decision to have a second child is explained by the independent variables in the first stage. Firstly we found that for each 1 unit of increase in women's education level (see Table 2 for variable unit definition), the woman is 8.3% less likely to have a second child, agreeing with prior studies which have also found that highly educated women tend to have fewer children in developing countries [Cornett, 2020]. However, the spousal education level is deemed insignificant by the model, showing that the husband's influence is trumped by the mother herself. This result is an indicator that the patriarchal construct has declined significantly in China. Secondly, a mother is 3.3% more likely to have a second child when her first child ages by 1 year. Our exclusion of older women, whose fertility have declined, ensures a positive relationship as an older child needs less time to be taken care of, then the mother will have more time to take care of a newborn baby if she is still in prime childbearing age. Thirdly, women living in the north regions

are 10.6% less likely to bear a second child, which could imply either stricter enforcement or weaker cultural preference for more children. With the inclusion of all 3 northeastern provinces in the data, the result largely corresponds with the observed "ultra-low fertility" in the region.[Huang, 2019] Lastly, the IV's significance has already been illustrated.

Table 4 presents the second stage regression results and we can summarize some key findings as follows. The R-squared value of 0.3358 shows that 33.58% variation in log income is explained by the independent variables in the second stage. The main finding which answers our research question is that bearing a second child is likely to cause a 42.4% penalty in income holding all other variables constant. With a p-value of 0.07, we cannot claim statistical significance on the effect of having a second child at 0.05 significance level. Nonetheless, this result approaches the borderline of significance as the p-value is still less than 0.1. However, previous studies have examined a statistically significant negative effect of having a first child on women's earnings but none have suggested a magnitude exceeding 40%. However, the negative effect of an additional child should be decreasing [Lundborg et al., 2017] as mothers get more prepared when taking care of a second child and women's position in the workplace becomes more stable as their age and experience grow. We deem the exact coefficient obtained from the regression model as not necessarily a meaningful value that represents the true penalty due to its large standard error and its magnitude. However, with a provisionally significant p-value, we can interpret the sign of the coefficient with reasonable confidence that having a second child does not have a positive effect on working mothers' income. This is further supported by the fact that the 95% confidence interval of the coefficient is between -0.88 and 0.03.

Combining the first-stage and second-stage result, we can clearly identify the endogeneity of motherhood as a few variables are jointly significant in either stage. For example, for every one level higher in education, the mother's income increases by 25.69% on average. This result reflects the fact that a higher education level leads to better job opportunities and higher wages. Combining this with the first-stage result, we can see that high education simultaneously has a negative effect on fertility and a positive effect on income. In comparison, a more educated spouse has a positive influence on mother's income but the effect on fertility is insignificant. With an education level correlation of 0.4633 (in Table 6), we can conclude that more educated people are likely to marry each other and thus boost their personal income. At the same time, we can interpret the insignificance of spousal education as mothers in a more educated household have more power in her own fertility decision.

Numerous factors including whether the woman has ever taken train, the prestige score, the interaction term of full time work and daily job commitment not larger than 6 hours, and the interaction term of full time work and daily job commitment larger than 6 hours are all statistically significant and exert a positive effect on income. This shows that *ever_took_train* variable is a strong proxy for economic status as more affluent women are more likely to travel, with those responding "Yes" having a 21% income boost on average. On the other hand, significant factors including whether living in the north region, the interaction term of woman in agricultural work and spouse not in agricultural work, the interaction term of both woman and her spouse in agricultural work are negatively correlated to income. The negative coefficient of agricultural work agrees with a wide gap in urban/rural income, with the urban rural income ratio standing at 3.33 as of 2009 [Chen et al., 2020]. Although this measure is based on place of residence rather than household registry, it is still valid as agricultural registry households are more likely to live in rural areas.

7 Limitation Conclusion

7.1 Limitation

A major limitation that may have prevented the variable of bearing a second child from reaching a statistically significant level is the nature of rural income. Compared with urban households, it is much more likely for a rural household to misreport true income, particularly those engaged in agricultural work. A 2005 paper that investigates income inequality in rural China [Benjamin et al., 2005] has pointed out that total reported income often combines both agricultural and non-agricultural income from all sources including that of other family members. The monetary value reported in this dataset is aggregated income as labour income is not separately surveyed. Therefore, this may not fully represent the mother's actual income as she may have aggregated her income together with the family if they all work collectively on a piece of land. Thus, the attribution of family income in the survey is subject to the respondent's own judgement and may be a poor measure of her true income. In addition, agricultural income tends to fluctuate due to crop conditions and commodity prices. This adds another layer of noise to the reported income number where the underlying factor influencing it is unobservable. Evidenced by geographic diversity in the dataset, farmers may cultivate different crops and experience different market impacts from crop price fluctuations.

Another limitation that exists in the current implementation of the model is sample selec-

tion bias. Due to the policy constraint, the model only examines households with agricultural registry as of 2010. However, from 1990, large-scale urbanization has happened where the urban population as percentage of total population has risen from 26.41% to 49.68%. In terms of household registry status, non-agricultural registry has increased from 25.89% to 39.34% over the period [NationalBureauofStatistics, 2010] [NationalBureauofStatistics, 1990]. This indicates that there may exist a significant number of households who were agricultural when deciding to have a second child, but converted to urban registry prior to 2010. There are somewhat stringent requirements for an individual to switch from agricultural to non-agricultural registry, usually related to their income and education level. This is reflected in the census that urbanization rate is approximately 10% higher than non-agricultural registry rate. Therefore, the exclusion of these data points may mask the true income penalty as households eligible to switch to urban registry would on average have earned higher income and attained higher education level.

Lastly, with reasonable intuitions we treat the IV as satisfying the exclusive restriction property as gender of the first child is not typically associated with the mother's income. This assumption could be weakened given China's marriage culture that imposes a financial burden on males. An expensive 'bride price' averaging a few years' of average rural household's disposable income, is payable by the groom's family to the bride's family upon marriage. According to a 2012 paper on the customs of bride price in rural China [Jiang and Sánchez-Barricarte, 2012], some impoverished households are forced to economize and save for decades and in certain scenarios to borrow money to afford their male offspring's marriage. Under these circumstances, the IV may directly cause changes in the mother's income level. If the first child is a boy, the mother may devote more time and effort to employment to plan for a potential marriage in the long term, thus directly increasing her potential income. On the other hand, if the first child is a girl, the mother may devote less time and effort to employment as an anticipation of higher future income in the form of a 'bride price'.

7.2 Conclusions

Fertility is often regarded as a factor that negatively affects women's earnings potential. Using the gender of the first child as an instrumental variable to control for the exogenous variation in selection to motherhood, we have reached a conclusion that supports this opinion but cannot accurately assess the exact impact of fertility. It is important to note that this study examines a small subset of the general population of women and the results are exclusively related to young mothers

having a second child. Therefore, the result may not necessarily generalize to the employment outcome of mothers having a first child or any subsequent children, which is more representative of the general population.

Adding new variables may lead to more accurate results, namely the information about grandparents, who often take the responsibility of "grandchild-care" in Chinese social settings. The survey also records some personal opinions on certain subjects that can be extracted to reveal more traits on individuals that potentially influence their income and decision to bear a second child.

The implications from this study may hold practical significance as China has recently introduced a three-child policy to boost its declining birth rate. Assessing the effect of having a second child can provide policy-makers valuable insights on how maternity decisions can be shaped or encouraged by introducing measures to reduce the penalization against working mothers. The standardization and extension of maternity leaves, and infrastructure roll-out of socialized childcare may all boost women's confidence in having more children.

References

[Int, 2013] (2013). *Econometrica*, 81(2):667–737.

[Angrist and Evans, 1998] Angrist, J. D. and Evans, W. N. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 88(3):450–477.

[Benjamin et al., 2005] Benjamin, D., Brandt, L., and Giles, J. (2005). The evolution of income inequality in rural china. *Economic Development and Cultural Change*, 53(4):769–824.

[Chen et al., 2020] Chen, Y., Luo, P., and Chang, T. (2020). Urbanization and the urban–rural income gap in china: A continuous wavelet coherency analysis. *Sustainability*, 12(19):1–14.

[Congress, 2014] Congress, S. P. P. (2014). Regulations of shandong province on population and family planning.

[Cornett, 2020] Cornett, A. (2020). Analyzing the relationship between female education and fertility rate. *Drake Undergraduate Social Science Journal*.

[Correll et al., 2007] Correll, S., Benard, S., and Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, 112(5):1297–1338.

[Halpert et al., 1993] Halpert, J. A., Wilson, M. L., and Hickman, J. L. (1993). Pregnancy as a source of bias in performance appraisals. *Journal of Organizational Behavior*, 14(7):649–663.

[Huang, 2019] Huang, W. (2019). China's northeast: the world's ultralow fertility capital.

[Institute of Social Science Survey, 2015] Institute of Social Science Survey, P. U. (2015). China Family Panel Studies (CFPS).

[Jiang and Sánchez-Barricarte, 2012] Jiang, Q. and Sánchez-Barricarte, J. J. (2012). Bride price in china: the obstacle to 'bare branches' seeking marriage. *The History of the Family*, 17(1):2–15.

[Keating, 2002] Keating, M. (2002). Ann crittenden, the price of motherhood: Why the most important job in the world is still the least valued. *Public Choice*, 113:485–488.

[Lundborg et al., 2017] Lundborg, P., Plug, E., and Rasmussen, A. W. (2017). Can women have children and a career? iv evidence from ivf treatments. *American Economic Review*, 107(6):1611–37.

[Markussen and Strøm, 2015] Markussen, S. and Strøm, M. (2015). The effects of motherhood. (No. 19/2015).

[Mourifié and Wan, 2017] Mourifié, I. and Wan, Y. (2017). Testing local average treatment effect assumptions. *The Review of Economics and Statistics*, 99(2):305–313.

[Mu and Xie, 2014] Mu, Z. and Xie, Y. (2014). Motherhood penalty and fatherhood premium? gender disparities of fertility effects in china. *Demographic Research*, 35:1373–1410.

[NationalBureauofStatistics, 1990] NationalBureauofStatistics (1990). Fourth national population census 1990.

[NationalBureauofStatistics, 2010] NationalBureauofStatistics (2010). Sixth national population census 2010.

[of People's Republic of China, 2013] of People's Republic of China, S. C. (2013). .

[Poduval and Poduval, 2009] Poduval, J. and Poduval, M. (2009). Working mothers: How much working, how much mothers, and where is the womanhood? *Mens sana monographs*, 7:63–79.

[Tribunal, 2011] Tribunal, A. G. R. R. (2011). China – chn36267 – guangdong – one child policy – children born overseas – fines.

[WorldBank, 2021] WorldBank (2021). Labor force participation rate, female (% of female population ages 15+) - china.

[ZhejiangProvincialDepartmentofFinance, 2016] ZhejiangProvincialDepartmentofFinance (2016). .

A TABLES

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Personal Income Last Year	472	10193.24	12510.95	160	150000
Birth Year	472	1976.57	4.97247	1967	1990
First Child Sex	472	.48305	.50024	0	1
Bear Second Child	472	.34110	.47458	0	1
Bear Third Child	472	.02542	.15758	0	1
Highest Education	472	1.99576	.60429	1	4
Highest Eduction of Spouse	472	2.14831	.5483398	1	4
In Agricultural Work	472	.30932	.46270	0	1
Spouse in Agricultural Work	472	.28390	.45137	0	1
Ever Took Train	472	.70339	.45725	0	1
Prestige Score	472	38.3276	6.74588	13	78
North Region	472	.63983	.48056	0	1
Full Time Work	472	.69492	.46093	0	1
Daily Job Commitment 6 hours	472	.36017	.48056	0	1

Table 2: Variable Definition

Variable	Definition	Range Definition
Personal Income Last Year	Reported Personal Income in 2009	N.A.
Birth Year	Reported Birth Year	N.A.
First Child Sex	Binary Indicator of Sex of First Child	1=Female, 0=Male
Bear Second Child	Binary Indicator of whether individual has second child	1=Yes, 0=No
Bear Third Child	Binary Indicator of whether individual has second child	1=Yes, 0=No
Highest Education	Level of Education Attained	1=Illiterate/Semi-Illiterate 2=Primary/Middle School 3=High School 4=College or Above
Highest Education of Spouse	Level of Education Attained	1=Illiterate/Semi-Illiterate 2=Primary/Middle School 3=High School 4=College or Above
In Agricultural Work	Binary indicator of full-time occupation in agriculture, fishing, grazing	1=Yes 0=No
Spouse in Agricultural Work	Binary indicator of full-time occupation in agriculture, fishing, grazing	1=Yes 0=No
Ever Took Train	Binary indicator of respondent ever used train as transportation	1=Yes 0=No
Prestige Score	Rating of main occupation	13=Least Prestigious, 78=Most Prestigious
North Region	Province located above Qinling-Huaihe Line	1=Shandong, Heilongjiang, Jilin, Liaoning, Hebei, Shaanxi, 0=Anhui, Zhejiang, Hubei, Hunan, Jiangxi
Full Time Work	Binary Indicator of Full Time Employment	1=Yes 0=No
Daily Job Commitment ¡ 6 hours	Binary Indicator of working less than 6 hours a day	1=Yes, 0=No

Table 3: First Stage Regression Results

VARIABLES	Bear Second Child
Highest Education	-0.08344**
	(0.03626)
Highest Education of Spouse	0.055416
	(0.03823)
First Child Age	0.03323***
	(0.00389)
Ever Took Train	-0.07049
	(0.0404)
Prestige Score	0.00220
	(0.00284)
Region North	-0.1083***
	(0.03909)
Not in Agricultural Work * Spouse in Agricultural Work	0.06597
	(0.06156)
In Agricultural Work * Spouse Not in Agricultural Work	0.10869
	(0.0693)
In Agricultural Work * Spouse in Agricultural Work	0.14654**
	(0.05900)
Not Full Time Work * Daily Job Commitment >6 hours	0.261766***
	(0.07496)
Full Time Work * Daily Job Commitment <= 6 hours	0.16507**
	(0.07196)
Full Time Work * Daily Job Commitment >6 hours	0.2873**
	(0.09382)
First Child Sex	0.3318199***
	(0.03675)
Constant	-0.30549*
	(0.16210)
Observations	472
R-squared	0.3350
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Table 4: IV Regression Results

VARIABLES	log(Income)
Bear Second Child	-0.42476*
	(0.23424)
Highest Education	0.25430***
	(0.079801)
Highest Education of Spouse	0.21648***
	(0.0808)
First Child Age	0.02655**
	(0.01112)
Ever Took Train	0.21561**
	(0.0859)
Prestige Score	0.0197818***
	(0.00603)
Region North	-0.33627***
	(0.08665)
Not in Agricultural Work * Spouse in Agricultural Work	0.0046326
	(0.13094)
In Agricultural Work * Spouse Not in Agricultural Work	-1.0099***
	(0.14722)
In Agricultural Work * Spouse in Agricultural Work	-0.76462***
	(0.12790)
Not Full Time Work * Daily Job Commitment >6 hours	-0.01246
	(0.17178)
Full Time Work * Daily Job Commitment <= 6 hours	0.6238***
	(0.15819)
Full Time Work * Daily Job Commitment >6 hours	0.52694**
	(0.20848)
Constant	6.818***
	(0.338758)
Observations	472
R-squared	0.3358
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Table 5: Provincial Disparity in Policy Enforcement

Province	Enforcement	Z = 0	Z = 1	P(D = 1 $\|Z = 0$)	P(D = 1 $\|Z = 1$)
Shandong	Strong	41	30	2.4%	50%
Zhejiang	Strong	23	13	0.0%	23.1%
Guangdong	Weak	80	83	60%	71.2%
Fujian	Weak	15	13	46%	54%

*** D = Bear Second Child (1=Yes)

*** Z = First Child Sex (1=Female)

Table 6: Correlation Matrix of Select Variables

	In Agricultural Work	Spouse In Agricultural Work	Full Time Work	Daily Job Commitment <= 6 hours	Highest Education	Highest Education of Spouse
In Agricultural Work	1.0000					
Spouse In Agricultural Work	0.3985	1.0000				
Full Time Work	0.4439	0.0384	1.0000			
Daily Job Commitment <= 6 hours	0.0149	0.0186	-0.5179	1.0000		
Highest Education	-0.2390	-0.1908	0.0182	-0.0093	1.0000	
Highest Education of Spouse	-0.1645	-0.2049	0.0794	-0.0348	0.4633	1.0000

Table 7: IV Validity Check

	Z = 0	Z = 1	Total
D =1	47	114	161
D = 0	197	114	311
Total	244	228	472

*** D = Bear Second Child (1=Yes)

*** Z = First Child Sex (1=Female)

B Equations

Equation-1: First Stage Regression (Condensed Form)

Bear_Second_Child = constant + $\alpha_{1LS} First\_Child\_Sex +$

$\beta_1 Highest\_Education + \alpha_2 Highest\_Education\_of\_Spouse +$

$\alpha_3 First\_Child\_Age + \alpha_4 Ever\_Took\_Train + \alpha_5 Prestige\_Score + \alpha_6 Region\_North +$

$\alpha_7 Not\_in\_Agricultural\_Work * Spouse\_in\_Agricultural\_Work +$

$\alpha_8 Not\_Full\_Time * Daily\_Job\_Commitment < 6hours + \epsilon$

*** ϵ is the error term in the first stage regression

Equation-1: First Stage Regression (Expanded Form)

$$
\begin{aligned}
Bear\_Second\_Child = constant + \alpha_{1LS} First\_Child\_Sex + \\
\alpha_1 Highest\_Education + \alpha_2 Highest\_Education\_of\_Spouse \\
+ \alpha_3 First\_Child\_Age + \alpha_4 Ever\_Took\_Train + \alpha_5 Prestige\_Score + \alpha_6 Region\_North + \\
\alpha_7 * (Not\_in\_Agricultural\_Work, Spouse\_in\_Agricultural\_Work) + \\
\alpha_8 * (In\_Agricultural\_Work, Spouse\_Not\_in\_Agricultural\_Work) + \\
\alpha_9 * (In\_Agricultural\_Work, Spouse\_in\_Agricultural\_Work) + \\
\alpha_{10} (Not\_Full\_Time, Daily\_Job\_Commitment > 6hours) + \\
\alpha_{11} (Full\_Time, Daily\_Job\_Commitment <= 6hours) + \\
\alpha_{12} (Full\_Time, Daily\_Job\_Commitment > 6hours) + \epsilon
\end{aligned}
$$

Equation-2: Second Stage Regression (Condensed Form)

log_personal_income_last_year = constant + $\beta_{2LS} \hat{Bear\_Second\_Child} +$

$\beta_1 Highest\_Education + \beta_2 Highest\_Education\_of\_Spouse +$

$\beta_3 First\_Child\_Age + \beta_4 Ever\_Took\_Train + \beta_5 Prestige\_Score + \beta_6 Region\_North +$

$\beta_7 Not\_in\_Agricultural\_Work * Spouse\_in\_Agricultural\_Work +$

$\beta_8 Not\_Full\_Time * Daily\_Job\_Commitment < 6hours + \mu$

*** μ is the error term in the first stage regression

*** $\hat{Bear\_Second\_Child}$ is the predicted value of $Bear\_Second\_Child$ obtained from first stage regression

Equation-2: Second Stage Regression (Expanded Form)

$$log\_personal\_income\_last\_year = constant + \beta_{2SLS} \hat{Bear\_Second\_Child} +$$
$$\beta_1 Highest\_Education + \beta_2 Highest\_Education\_of\_Spouse$$
$$+ \beta_3 First\_Child\_Age + \beta_4 Ever\_Took\_Train + \beta_5 Prestige\_Score + \beta_6 Region\_North +$$
$$\beta_7 * (Not\_in\_Agricultural\_Work, Spouse\_in\_Agricultural\_Work) +$$
$$\beta_8 * (In\_Agricultural\_Work, Spouse\_Not\_in\_Agricultural\_Work) +$$
$$\beta_9 * (In\_Agricultural\_Work, Spouse\_in\_Agricultural\_Work) +$$
$$\beta_{10}(Not\_Full\_Time, Daily\_Job\_Commitment > 6hours) +$$
$$\beta 11(Full\_Time, Daily\_Job\_Commitment <= 6hours) +$$
$$\beta 12(Full\_Time, Daily\_Job\_Commitment > 6hours) + \mu$$

C Figures

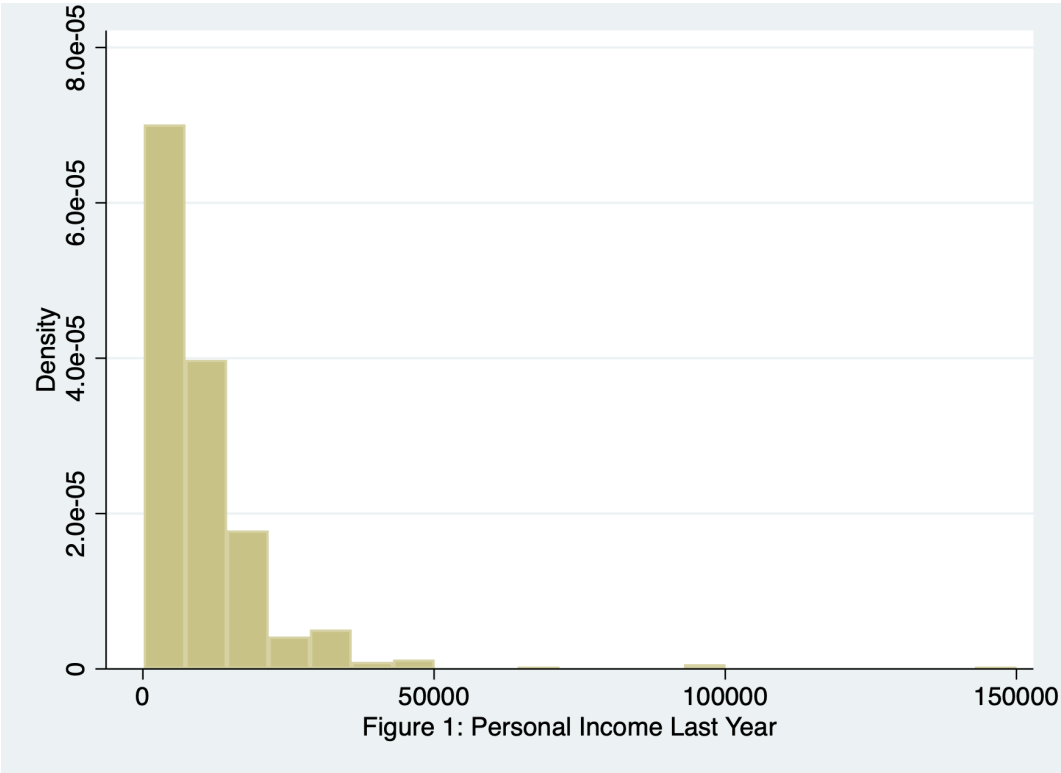

Figure 1

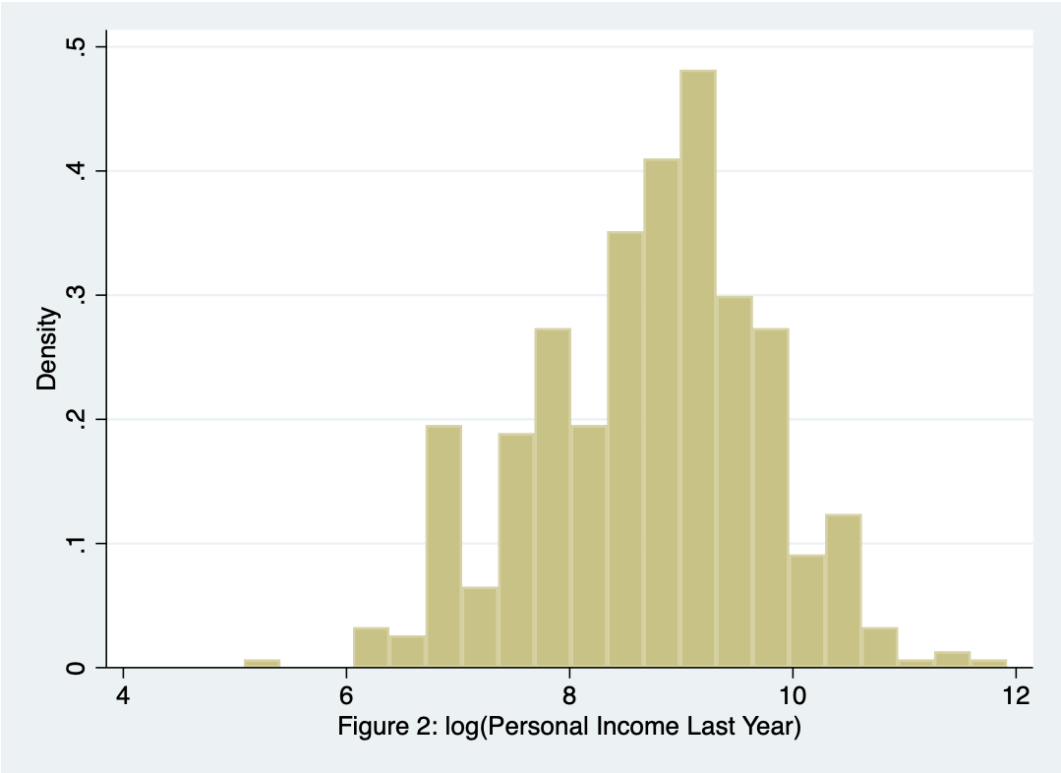

Figure 2

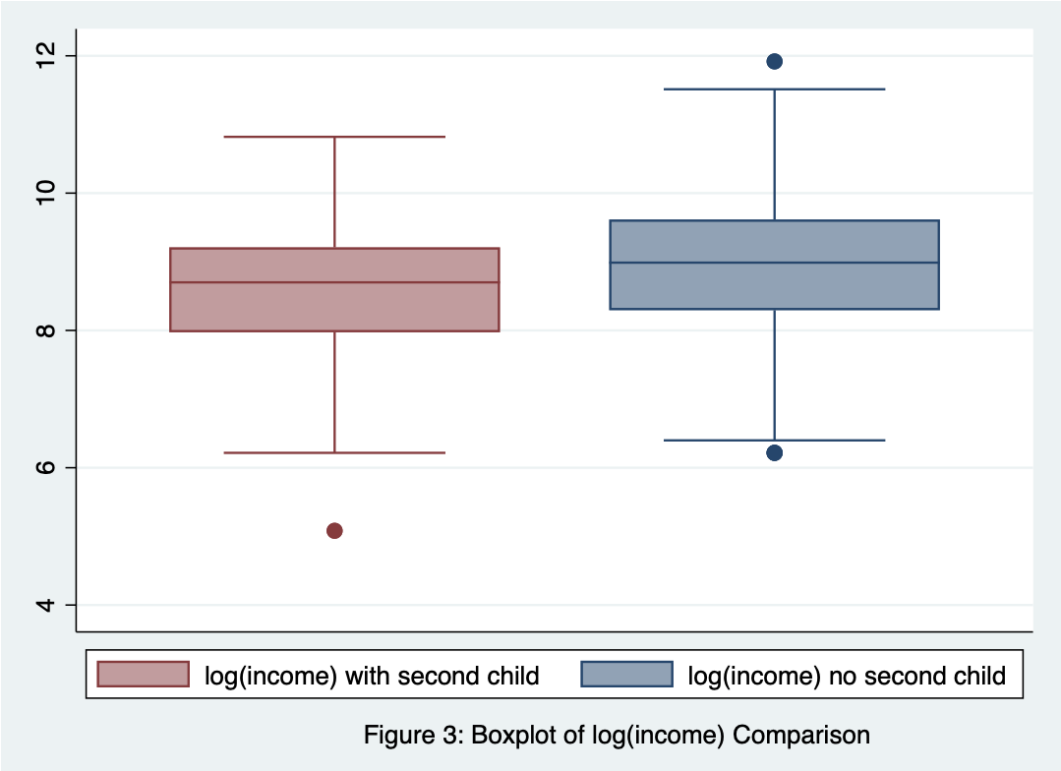

Figure 3

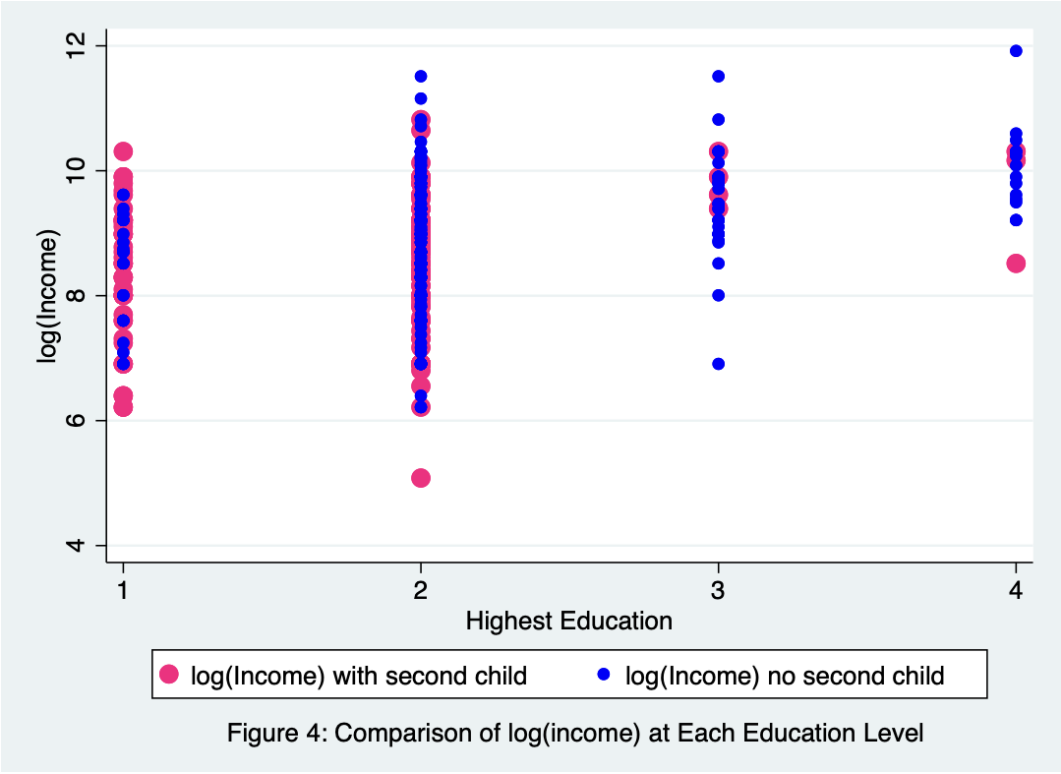

Figure 4

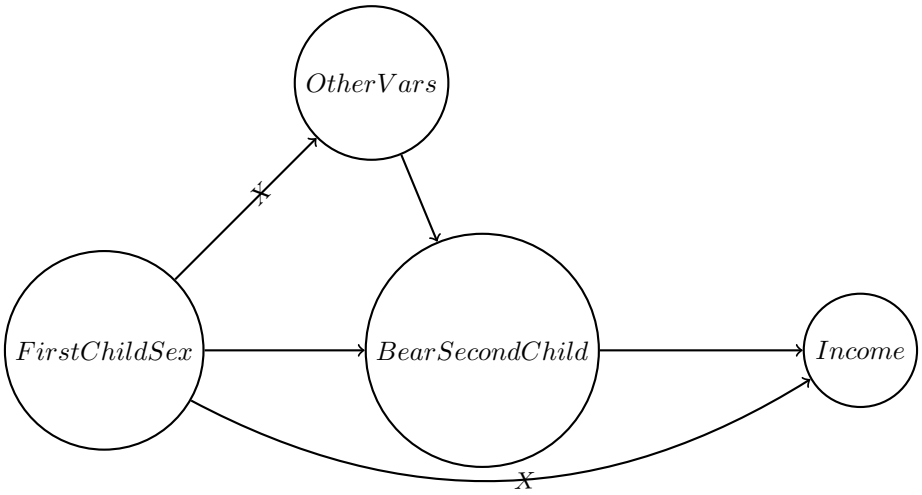

Figure 5: IV Causation conditions