

Lyah Muktavaram

Professor Anthony Keats

Quantitative Methods in Economics

5/19/23

Educational Attainment on Women's Fertility Rate

What, if any, is the causal effect of educational attainment on the number of children a woman has? The independent variable in this study is educational attainment (educ), denoting the highest year of school or degree a respondent completed. The dependent variable (chborn) reports the number of children ever born to each woman; this number reflects total live births by all fathers, regardless of whether or not the children are still living. According to the World Economic Forum, fertility rates have decreased by 50% worldwide for the past 70 years (WeForum). As a prospective environmental studies minor, I've learned about the consequences of the worldwide decline in total fertility rates on the environment. Beyond environmental dynamics, however, a worldwide decline in fertility has significant effects on social and economic patterns and development. Firstly, a mass decrease in fertility rates correlates with a reduction in the encompassing effects of a growing population. The impacts of global warming, deforestation, and air pollution, for instance, will be lessened due to the limited population. Many environmental issues stem from the power of the extensive global population, and thus, lower fertility rates may help facilitate a more sustainable planet and population. On the other hand, critics of low fertility rates cite concerns regarding the aging population coupled with longer life expectancies. An older population risks economic destabilization as healthcare costs may rise and the global workforce will be limited. Ultimately, fertility rates and the average number of children born to a woman supply paramount insights into worldwide affairs. As a

result, it is essential to understand the variables that affect fertility rates. Educational attainment, for example, may empower women to pursue a long-term career or employment as opposed to having children. Furthermore, higher educational attainment may provide women with a better understanding of sex education, and therefore, reduce the risk of unplanned pregnancy. Country governments may institute educational programs to encourage more women to participate in the workplace, perhaps in an effort to combat the effects of a growing population or stimulate the economy.

Literature Review

McCray and Royer (2011) evaluated age-at-school entry policies to consider the effect of female education on fertility and infant health. Essentially, McCray and Royer sought to answer the following question: Do school entry policies have a causal effect on female education and subsequently fertility and infant health? By using 1989-2001 Texas and 1989-2002 California natality data, McCray and Royer focused on the contrasts in education, fertility, and infant health between women born before and after the school entry threshold. The natality files, acquired from each state's Department of Health, covered approximately 800,000 births per year. Each mother and her healthcare provider completed a survey regarding maternal and paternal behaviors, demographic characteristics, and the health of the infant. McCray and Royer restricted their data to mothers born in the state in which they gave birth, first-time mothers (to ensure the independence assumption holds), mothers who are 23 years old or younger, and potential mothers born between 1969 and 1987. First-time mothers do not have previous experience with childbirth, thus their general knowledge of pregnancy is likely derived from their education. McCray and Royer found no difference in fertility behaviors for those born just before and after

the school entry cut-off dates. Additionally, McCray and Royer found no statistically significant difference in fertility behaviors associated with school entry policies. The key findings from this study suggest women born just after the school entry date 1) have lower schooling, 2) are equally likely to become mothers, 3) give birth at similar ages, and 4) give birth to similarly healthy infants.

It is also important to recognize the study's lack of external validity. McCray and Royer recognize their estimates are localized to women whose education at motherhood is affected by school entry policies. The women may be negatively selected due to their parents' willingness to comply with school entry policies (common among lower socioeconomic groups). As a result, McCray and Royer's results can not be generalized and applied to other groups or subpopulations. McCray and Royer's study is nevertheless noteworthy when considering the causal relationship I am examining. Since my paper does not consider the influence of school entry policies on a woman's educational attainment, I was concerned with the possible variance and discontinuity that may arise. However, McCray and Royer's findings demonstrate that even though women born just after the cut-off date received substantially lower schooling, there is no distinguishable difference between their infants' health and motherhood practices. While McCray and Royer's findings don't note the number of children ever born to each woman, their results are closely related to my paper's focus on the fertility rate. McCray and Royer's study, for example, looked into the age at which a mother first gives birth. A woman who conceives her first child at a younger age may be more likely to have multiple births. These results, consequently, force me to reconsider the influence of a marginal change in educational attainment on the fertility rate. If a woman born just after the cut-off date drops out of school at the same time a woman born just before the cut-off date drops out of school, will their fertility

rates differ? McCray and Royer's study findings suggest there would be no difference in their likelihood of becoming mothers, their childbearing age, and their infants' health. What does this say about the marginal increase of educational attainment on the fertility rate?

Götmark and Andersson (2022) consider the relationship between education, economy, religion, and contraceptives and family planning programs on fertility levels in six global regions. The study used data from Eastern Europe, Western Europe, Latin America and the Caribbean, the Middle East, Sub-Saharan Africa, and Asia, amassing 141 countries. Götmark and Andersson weren't able to include Russia, China, and other countries in their findings due to a lack of available data. In considering shared political and economic history, Götmark and Andersson also differed from the United Nations grouping by placing Eastern Europe as a separate region. Table One lists countries and the total fertility rate for a five-year period (2010-2015). Götmark and Andersson's data concerning education reflects the number of years of education received by women aged 25 and older. The findings for GDP per capita come from data from the World Bank from 2011-2015, and the data for religion comes from standardized surveys by Gallup from 2005 to 2012. Götmark and Andersson examined the contraceptive prevalence rate (CPR) by dividing "the number of women of reproductive age who are married or in a union and who are currently using a method of contraception by the number of women of reproductive age who are married or in a union" for the period 2006-2010. A total score for family planning programs was calculated (based on policy, service, records and evaluation, and contraceptive access) and used in Götmark and Andersson's analysis.

Götmark and Andersson found that the total fertility rate negatively correlates with education, gross domestic product (GDP) per capita, and contraceptives and family planning programs. On the other hand, the total fertility rate correlates positively with religion. In Western

Europe, interestingly, the total fertility rate increases with education, a finding that deviates from the rest of the study's regions. Götmark and Andersson acknowledge that "more work is needed to establish causality and relative importance of the factors." While the data I am using for my paper is limited to the United States, Götmark and Andersson's study provides interesting and applicable information. Firstly, their results confirm my hypothesis that a woman's total fertility rate negatively correlates with her educational attainment. The study, however, did not compile data from the United States, which differs economically, socially, and politically from the rest of the countries surveyed. Western European countries likely share fundamental similarities with the United States, yet interestingly, their results regarding fertility and education differ from the other global regions. As a result, it is difficult to verify if Götmark and Andersson's study has external validity that can be applied to my paper. It is nevertheless helpful to consider Götmark and Andersson's results as I write my paper because the frameworks for their study are largely similar. Increased educational attainment, for instance, positively correlates with other factors that may decrease fertility rates (contraceptives and family planning programs). It is consequently important for me to understand the ways in which educational attainment influences total fertility rates for my research.

Bailey (2013) assesses the long-term effects of United States family planning policies, specifically those which increase legal and financial access to contraceptives. The study leverages data after policy changes that occurred during the late 20th century: the birth control pill's reintroduction after *Griswold versus Connecticut* was repealed and federal funding for local family planning programs was expanded. Bailey considers the influence of these policies on fertility rates, the child's college completion, labor force participation, wages, and family incomes decades later. It is important to note that family planning programs in the United States

were not randomly assigned to representative participants. As a result, there may be selection bias regarding the women who have access to or participate in these family planning programs. Furthermore, while fertility rates fell in the late 20th century, it is difficult to identify if this indicates a causal relationship between family planning policies or if it is simply a consequence of a longer-term, overarching national trend. To best address this concern, older data in Bailey's study used results from cross-sectional surveys to limit omitted variable bias and newer data was the product of localized, randomized interventions that aimed to reduce teen pregnancies. By using data from a myriad of findings, Bailey's study maintains internal validity.

Bailey found that children whose mothers had greater access to family planning programs went on to live in high-earning households. Increasing legal and financial access to family planning programs is associated with a 2% to 3% increase in family income. Furthermore, children whose mothers had greater financial access to contraceptives were more likely to receive higher educational attainment. In fact, these children were 2% to 7% more likely to attain 16 or more years of education. Bailey also concluded that after assessing the costs and benefits of investing in family planning programs, the programs provide abundant opportunities and longer-term benefits. While Bailey's findings are incredibly expansive, the takeaways found are applicable to my paper. Firstly, Bailey's concern with family planning programs is important to consider as I examine the influence of educational attainment. As Götmark and Andersson explain, educational attainment positively correlates with family planning programs, and thus, it is interesting to examine a study that exclusively focuses on access to the programs. Bailey's findings are also fascinating as they consider the long-term influence of family planning programs on the child's educational attainment.

Data and Descriptive Statistics

The variables I included in Table One are as follows: year, sex, age, marital status, age at first marriage, children ever born, educational attainment, labor force status, hours worked last week, and wage and salary income. The variables I included were sourced by IPUMS (United States). Firstly, the variable “year” denotes the year in which the household was surveyed. For the purposes of this paper, my “year” variable has a minimum value of 1930 and a maximum value of 1940, as reported in Table One. The variable “sex” reports whether the person surveyed identified as a male or a female; this dummy variable is especially important because my paper concerns female educational attainment and fertility rate. Furthermore, the numeric variable “age” documents the respondent’s age (in years). According to Table One, the mean age of the respondents in the data is 29.77 years old. The variables “sex” and “age” help describe the basic profile of the observations.

I also included the variable “mar” to note the person’s marital status at the time of the census. I concluded that early 20th-century traditional family values likely included childbearing, and thus, one’s marital status could influence the age at which their first child is born. The dummy variable “mar” takes on the value 0 if the spouse is not present and the value 1 if the spouse is or was present (divorcees or widows). The variable “agemarr” denotes the respondent’s age at their first marriage. I hypothesized that a married respondent would be likely to have children at a younger age. According to Table One, the mean age for “agemarr” is 23.15 years old (I dropped the missing observations for respondents who aren’t married). The dummy variable “labforce” signifies whether the respondent participated in the labor force. “labforce” acts as a control for my regression as a woman in the labor force may not have time for children. The numeric variable “hrswork1” (hours worked last week) also acts as a control for my

regression. Additionally, the variable “incwage” reports the respondent’s total pre-tax wage and salary income. “incwage” not only acts as a control for my regression, but it also provides information about the respondent’s profile. Finally, the independent variable for my study is educational attainment (educ), and the dependent variable is the number of children born to each woman (chborn).

	mean	sd	min	max	count
census year	1940.00	0.00	1940.00	1940.00	121446
sex	2.00	0.00	2.00	2.00	121446
age	44.91	16.55	14.00	100.00	121446
mar	0.92	0.27	0.00	1.00	121446
agemarr	21.65	5.16	12.00	98.00	121446
children ever born	3.10	2.42	1.00	61.00	121446
educational attainment [general version]	3.44	2.37	0.00	11.00	121446
labforce	1.21	0.41	1.00	2.00	121446
hours worked last week	7.54	17.28	0.00	98.00	121446
wage and salary income	17215.43	129606.66	0.00	999999.00	121446

The timeframe for my study is limited to data from 1930 to 1940. While I originally wanted to expand my timeframe to include data from the 21st century, some of my control variables weren’t available for the later periods. After further research, I realized that census results from this time could make for compelling data. The Great Depression dramatically expanded educational participation with increased access to public school systems. In Detroit, for example, the number of children in schools more than doubled from 1920 to 1930. With women’s mean age for first childbirth at around 20 (years), the timeframe 1930 to 1940 would

begin to showcase the effects of increased access to public education. I am also limiting my data to women in the United States. Firstly, the IPUMS (United States) database only lists data from the United States. More importantly, however, my control variables would have varied significantly if I expanded the scope of my data collection. The control variable “agemarr,” for instance, differs in other countries. According to economist Kumudini Dandekar, the mean age of marriage for women in India was 14.9 years from 1931 to 1941, almost 10 years younger than women in the United States. As a result, it would have been difficult to separate cultural, political, and economic differences between countries from the influence of the control variables in my regression.

Empirical Analysis

Running a simple regression of my dependent variable on my independent variable—regchborn educ, robust—warrants the following results. The estimated coefficient on “educ” is -0.25, meaning that for every one-unit increase in “educ,” a -0.25 increase in “chborn” is expected. In other words, an increase of one year of education decreases predicted children born to a woman by 0.25 children. The estimated coefficient of -0.25 is statistically significant with a p-value of 0.05 or less—the null hypothesis is rejected as the coefficient is statistically significant from 0. The significance level is 95%, and the preselected alpha value is 0.05.

The simple regression I ran, however, did not give an unbiased estimate of the relationship between educational attainment and the number of children born to a woman. Income, for instance, was omitted from the simple regression equation which may cause the estimate to be biased. As a woman’s income increases, she may be more likely to hold off having children. Perhaps the woman is investing in her employment in order to supplement her income,

and thus, doesn't have the time or opportunity to raise children. Furthermore, increasing income provides women with the resources they need to go to college or complete higher education. As a result, income positively correlates with educational attainment (women with higher incomes may complete more years of schooling). In the simple regression, the variable "incwage" is biased upwards because women with more income tend to have higher educational attainment, and a higher income may incentivize women to have fewer children. In order to address this Omitted Variable Bias (OVB), I will re-run the regression with the variable "incwage" accounted for.

In addition to controlling for "incwage" in my regression analysis, I will also include variables for sex, age, marital status, age at first marriage, labor force status, and wage and salary income. The variables "hrswork1" and "incwage" serve as an interaction term because the effect of the independent variable "incwage" on the outcome will be aggregated with an increase in "hrswork1." Running a regression of "chborn" on "educ" and "incwage" results in the estimated coefficient -0.000798 on "incwage." The estimated coefficient is statistically significant with a p-value of 0.05 or less. The significance level is 95% because the preselected alpha value is 0.05. My findings for the second regression support my main hypothesis. A one-unit increase in wage and salary income predicts a -0.0007983 increase in the number of children born to a woman. When I ran a regression of "chborn" on "educ" and "incwage," the estimated coefficient on the primary independent variable, "educ," changed to -0.23. The coefficient of interest changed from -0.25 to -0.23 when the control variable was included in the regression. The change in the value of the main coefficient suggests that the bivariate regression suffered from OVB and inaccurately attributed a one-unit increase in educational attainment to a -0.25 increase in the number of children born to a woman.

Next, I ran a regression of “chborn” on “educ,” “incwage,” “sex,” “age,” “mar,” “agemarr,” and “labforce,” which resulted in the estimated coefficient of 0.0523 on “age,” holding all other variables constant. This regression also presented an estimated coefficient value of -0.12 on “agemarr.” The estimated coefficients are statistically significant with a p-value of 0.05 or less; the significance level is 95%. In addition, the estimated coefficient value for wage and salary income is -8.65×10^{-10} (in thousands). The regression also produced the estimated coefficient of 0.0774817 on marital status. Regarding the effect of labor force status, the coefficient on “labforce” is -0.5853033. The dummy variable “sex” was omitted because there is expected collinearity between women and the fertility rate. The estimated coefficients are statistically significant with a p-value of 0.05 or less; the significance level is 95%. When I ran a regression of “chborn” on “educ,” “incwage,” “sex,” “age,” “mar,” “agemarr,” and “labforce,” my coefficient on my primary independent variable, “educ” changed to -0.12. As a result, my second regression could have been impacted by OVB and assigned a one-unit increase in educational attainment to a -0.23 increase in the number of children born to a woman. Instead, the estimated coefficient -0.12 on the variable “agemarr” predicts that for every one-year increase in the age a respondent is married, a 0.12 decrease in “chborn” is expected. While the estimated coefficient on “agemarr” supports my hypothesis, the coefficient on “age” deviates from my initial predictions. The estimated coefficient on “age” suggests that a one-year increase in age projects a 0.0523 increase in the number of children ever born to a woman. In addition, the estimated coefficient of 0.0774817 on “mar” corroborates my initial predictions—the number of children born to a woman increases by 0.0774817 when the dummy variable switches from no spouse present to spouse present. The coefficient of -0.5853033 on “labforce” also supports my hypothesis as the coefficient on the variable is negative. A negative estimated coefficient

signifies that the number of children born to a woman decreases when the woman is part of the labor force.

Next, I ran a regression of “chborn” on “educ,” “incwage,” “sex,” “age,” “mar,” “agemarr,” “labforce,” and “hrswork1.” The regression resulted in the estimated coefficient -0.0034945 on hours worked last week. The estimated coefficient is statistically significant with a p-value of 0.05 or less; the significance level is 95%. The coefficient on my primary independent variable, “educ” changed to -0.1242274. While the change in the value of the main coefficient is minimal, it could suggest that the previous regression failed to account for OVB and inaccurately attributed a one-unit increase in educational attainment to a 0.12 decrease in the number of children ever born to a woman. The estimated coefficient -0.0034945 on the variable “hrswork1” suggests that for every one-hour increase in the hours worked last week, a 0.0034945 decrease in “chborn” is predicted. The findings from this regression also support my main hypothesis as an increase in hours worked last week negatively influences educational attainment.

I also ran a regression of “chborn” on “educ,” “sex,” “age,” “mar,” “agemarr,” “labforce,” and the interaction term “hrswork1_incwage.” The estimated coefficient on the interaction term is -4.19×10^{-9} (in thousands). The estimated coefficient is statistically significant with a p-value of 0.05 or less; the significance level is 95%. The estimated coefficient on my primary independent variable, “educ,” marginally changed to -0.1207917, possibly suggesting the previous regression suffered from omitted variable bias and accredited a one-unit increase in educational attainment to a 0.124 decrease in the number of children ever born to a woman. The estimated coefficient on the variable “hrswork1_incwage” suggests that for every one-unit increase in the hours worked last week and income, a -4.19×10^{-9} (in thousands) increase in “chborn” is predicted. The findings

from this regression support my main hypothesis. The results from my fourth and fifth regression, nevertheless, seem less impactful—the change in my estimated coefficient value on educational attainment was minimal. While the interaction term was statistically significant, its influence on regression findings wasn't notable.

As I researched possible sources of bias regarding the causal relationship between educational attainment and fertility rate, I realized there were variables I omitted in my regression. Firstly, religiosity or an attachment to religion could influence both educational attainment and fertility rate. A religious woman, for instance, could be more inclined to have children if the religion promotes family and traditional values. Subsequently, the woman may also be less willing to spend time and energy in school—schooling could interfere with her responsibilities as a mother. The variable “religion” would have been negatively correlated with the explanatory variable and positively correlated with the outcome variable, and thus, attribute bias for the estimated coefficient on educational attainment. Furthermore, access to contraceptives may have been helpful to account for in the regression analysis. A woman with practical access to contraception is less likely to have more children. Additionally, a woman with more access to contraception could have achieved higher education, and consequently, have a better understanding of where to find contraceptives. Access to contraception would have been negatively correlated with both the explanatory and outcome variables and thus, attribute a positive bias for the estimated coefficient on educational attainment.

	(1) children ever born	(2) children ever born	(3) children ever born	(4) children ever born	(5) children ever born
educational attainment [general version]	-0.252*** (0.003)	-0.253*** (0.003)	-0.124*** (0.002)	-0.124*** (0.002)	-0.121*** (0.002)
wage and salary income		-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	
agemarr			-0.120*** (0.002)	-0.120*** (0.002)	-0.120*** (0.002)
sex			0.000 (.)	0.000 (.)	0.000 (.)
age			0.052*** (0.000)	0.052*** (0.000)	0.052*** (0.000)
mar			0.077*** (0.023)	0.076*** (0.023)	0.193*** (0.022)
labforce			-0.585*** (0.013)	-0.460*** (0.027)	-0.476*** (0.018)
hours worked last week				-0.003*** (0.001)	
hours worked in one week x income from work					-0.000*** (0.000)
Constant	3.960*** (0.013)	3.969*** (0.013)	4.430*** (0.043)	4.304*** (0.050)	4.196*** (0.046)
R-squared	0.060	0.061	0.221	0.221	0.220
Adj R-squared	0.060	0.061	0.221	0.221	0.220
Observations	121446	121446	121446	121446	121446

Conclusion

I initially hypothesized that educational attainment plays a noteworthy role in the number of children born to a woman. The regressions I ran situate fertility rate as an outcome variable while I consider the influence of educational attainment, wage and salary income, marital status, age at first marriage, children ever born, labor force status, and hours worked last week on the number of children ever born to a woman. The findings from my regressions imply that higher educational attainment is associated with a decreased fertility rate. The control variables, nevertheless, limited the direct influence of educational attainment on a woman's fertility rate. The wage and salary income variable, for instance, changed the coefficient on educational

attainment from -0.25 to -0.23, an essential indication that other factors may also impact both the outcome and the independent variables.

The connection between a woman's educational attainment and the number of children she has is not surprising. Higher educational attainment offers women economic opportunities to advance their careers outside of the home and learn about sex education. As we evaluate the impacts of the fertility rate on the world's economic, environmental, and social dynamics, it is imperative to recognize the influence of factors like women's education. Empowering women through education is not only linked to lower fertility rates but also has well-received implications for the economy and society.

Works Cited

Alvarez, P. (2022). *What does the global decline of the fertility rate look like?*. World Economic Forum.

<https://www.weforum.org/agenda/2022/06/global-decline-of-fertility-rates-visualised/>

Bailey M. J. (2013). Fifty Years of Family Planning: New Evidence on the Long-Run Effects of Increasing Access to Contraception. *Brookings papers on economic activity*, 341–409.

<https://doi.org/10.1353/eca.2013.0001>

Dandekar, K. (1974). *Age at marriage of women - jstor*. Jstor.

<https://www.jstor.org/stable/pdf/4363695.pdf>

Götmark, F., Andersson, M. (2020). Human fertility in relation to education, economy, religion, contraception, and family planning programs. *BMC Public Health*, 20, 265.

<https://doi.org/10.1186/s12889-020-8331-7>

Longman, P. (2010). *Environmental consequences of low fertility rates*. Environmental Consequences of Low Fertility Rates | Newgeography.com.

<https://www.newgeography.com/content/001824-environmental-consequences-low-fertility-rates>

McCrary, J., & Royer, H. (2011). The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth. *The American economic review*, 101(1), 158–195. <https://doi.org/10.1257/aer.101.1.158>