The Expanding Education Gap in Women's Earnings: The Role of Fertility Timing

Catherine Doren
Princeton University

Abstract

In recent decades, gaps in earnings and fertility timing between women with and without a college degree have simultaneously grown. College-educated women, compared to less educated women, have been substantially delaying fertility, which is associated with lower motherhood pay penalties, suggesting that this trend may contribute to the growing disparity in women's earnings by education. This paper explores the relationship between rising inequality in women's earnings and shifting patterns of childbearing across education levels. I use data from the Survey of Income and Program Participation Synthetic Beta Data to estimate effects of motherhood on earnings by education and how these effects vary by first birth timing. I also consider how these effects have changed over time. I conclude by posing a counterfactual, asking to what degree the education earnings gap would differ today if all women still had the same first-birth timing as did same-education women in the 1960s.

Extended Abstract

Introduction

In recent decades, the earnings gap between women with and without a college degree has nearly doubled (Autor 2014). Across this same period, college-educated women began substantially delaying childbearing while less educated women did not, leaving recent cohorts with little similarity in the timing of childbearing by education (Cherlin, Talbert, and Yasutake 2014). Delayed fertility is associated with lower motherhood wage penalties, at least for college-educated women, suggesting that this trend may contribute to the growing disparity in women's earnings (Amuedo-Dorantes and Kimmel 2005; Chandler, Kamo, and Werbel 1994; Herr 2015; Miller 2011; Taniguchi 1999). This paper aims to understand this relationship between the rise in women's earnings inequality and the shift in patterns of childbearing across education levels. I ask, to what degree do education differences in fertility schedules account for the education gap in women's earnings? Have the benefits of fertility delay changed over time? And, finally, how would the education earnings gap differ today if all women followed the same fertility schedules followed by same-education women in the 1960s?

Recent studies have considered the role of fertility timing on motherhood wage penalties, asking whether women have smaller penalties if they delay fertility. Generally, these studies have used data from women in the baby boom cohort, who primarily had children during the

1970s-1990s (e.g., National Longitudinal Survey of Young Women 1968, National Longitudinal Survey of Youth 1979). They find that women in these cohorts who delay fertility into their late 20s and early 30s see smaller motherhood penalties than women who have children earlier, some even noting no effect of motherhood or a premium (Amuedo-Dorantes and Kimmel 2005; Doren 2018; Miller 2011; Taniguchi 1999). Yet these benefits of delay are seemingly only present for women with a college degree or more. If this is the case, more educated women will see a double advantage over less educated women given these trends. Not only are less educated women not delaying fertility, but if they do, they do not observe the same benefits. As fertility timing has diverged by education, these disparities may have grown as well. However, as fertility delay has become more common, it may be that the effects of delaying fertility have changed over time.

I use data from the Survey of Income and Program Participation (SIPP) Synthetic Beta Data (SSB) to examine the relationship between fertility timing and women's earnings inequality over time. First, I will divide the sample into cohorts (women who became mothers in the 1960s, 1970s, 1980s, 1990s, and 2000s) and track patterns of age at first birth by education across time. Next, I will estimate the effects of fertility timing on women's earnings by education for each cohort. Finally, I will apply a counterfactual analysis in which I ask how earnings inequality between women with and without a college degree would be different if women having children today had the same fertility timing and effects thereof as women in past cohorts.

Data and Measures

I use SSB data,[1] which links data from the SIPP survey respondents from all nine panels (1984, 1990, 1991, 1992, 1993, 1996, 2001, 2004, and 2008) to Social Security Administration (SSA) and Internal Revenue Service (IRS) records from 1951-2011.[2] The SIPP consists of a continuous

[1] This analysis was first performed using the SIPP Synthetic Beta (SSB) on the Synthetic Data Server housed at Cornell University which is funded by NSF Grant #SES-1042181. These data are public use and may be accessed by researchers outside secure Census facilities. For more information, visit https://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html. Final results for this paper will be obtained from a validation analysis conducted by Census Bureau staff using the SIPP Completed Gold Standard Files and the programs written by this author and originally run on the SSB. The validation analysis does not imply endorsement by the Census Bureau of any methods, results, opinions, or views presented in this paper.

[2] https://www.census.gov/programs-surveys/sipp/about/sipp-introduction-history.html

series of panels of 2 ½ to 4 years for a nationally-representative sample of households. All members of sample households age 15 or older are interviewed. In each panel, the sample size ranged from 14,000 to 52,000 households.

The SSB integrates this survey with administrative tax and benefit data to provide a unique opportunity to examine historical trends in fertility and earnings with high-quality data spanning the life course. Once a respondent participates in the survey, their past and future tax records are linked to their responses. To preserve confidentiality, the Census Bureau synthesizes the data to retain covariate relationships but obscure the actual values of variables. The "Complete Gold Standard Files" are then used to confirm that synthesized data do not produce misleading results.

This unique structure provides some benefits. With earnings data from tax records, SSB has much more accurate data on earnings than most datasets, which typically rely on self-reports. Additionally, these data are paired with information from respondents on a range of topics from the surveys, including their age at first birth and educational attainment. Yet the data are limited in other ways. Most notably, there are a limited number of control variables available, such as work characteristics beyond earnings as reported on W-2 forms. I therefore am constrained to use relatively few variables in my model. For my outcome variable, I use the natural log of annual earnings (adjusted using the CPI inflation index). I take the natural log to account for the skewness of the earnings distribution. Additionally, this transformation allows me to interpret motherhood effects in terms of percentages of earnings as opposed to dollar amounts, which makes effects more comparable across education levels. I estimate effects of motherhood with a series of time-variant age-at-first-birth-, education-, and cohort-specific dummy variables for whether women had had their first birth by a given year. I describe these in further detail below. I define educational attainment as a dummy variable, based on whether women ever received a bachelor's degree or more. I also include controls for age and age squared. Because I am considering the education earnings gap among women, I drop men from my sample.

Methods

First, I divide the sample into cohorts based on the decade in which they became mothers (1960s, 1970s, 1980s, 1990s, 2000s) and I describe women's distributions of age at first birth by education across decades. Next, I estimate the effects of fertility timing on the natural log of

women's annual earnings by education for each cohort, using fixed effects regressions to assess within-person change in log earnings before and after becoming mothers. The model is as follows:

$$\ln(earn)_{eit} = \beta_0 + \Sigma\beta_{ebc}(mom_{itebc}) + \beta(age_{it}) + \beta(age_{it}^2) + a_i + u_{it}$$

where the subscript *e* refers to educational attainment (less than college vs. BA or more), *b* refers to age at first birth, and *c* refers to cohort, *t* refers to time in years, and *i* refers to individuals. Similar to the strategy applied in Doren (2018) mom_{itebc} is a series of age-specific, education-specific, cohort-specific dummy variables (e.g., for women with a BA+ who gave birth at 20 in the 1960s: $mom_{it,BA+,20,1960}$) and estimates the fixed effects of motherhood at specific age of first birth for women with a given level of education in a given cohort. For instance, if a woman with a BA+ had her first birth at 20 in the 1960s, $mom_{it,BA+,20,1960}$ will have values of 0 at ages below 20 and values of 1 at ages 20 and higher, signaling the woman's transition to first birth. She will have values of zero at all ages on, for instance, $mom_{it,<BA,20,1960}$, which represents women who had less than a BA but had a birth at the same age in the same decade, $mom_{it,BA+,20,1970}$, which represents women who had the same education and same age at first birth, but gave birth in the next decade, and all other age-education-cohort combinations. The variables that do not apply to each woman will drop out of her fixed effects model due to the lack of variation, so the model estimates heterogeneous effects of motherhood by age at first birth-education-cohort group. Alpha represents individual fixed effects, u_{it} represents the person-period error term.

Finally, I apply a counterfactual analysis in which I ask how earnings inequality between women with and without a college degree would be different if women having children in the 2000s had the same fertility timing (distribution of age at first birth) and effects thereof (coefficients on motherhood by age at first birth) as women in the 1960s.

Preliminary Results

Figure 1 shows the expected patterns of diverging ages at first birth by educational attainment for women who had their first child in each decade from the 1960s through the 2000s. In the 1960s, average age at first birth for women with less than a BA was just over 22 and it hovered around age 22-23 through the 2000s. Women who attained a BA or more, on the other hand, had considerably more change in age at first birth over time. In the 1960s, their average age at first birth was similar to that of less educated women, at 23. Yet it rose to 25 in the 1970s, 27 in the

1980s, 28 in the 1990s, and 30 in the 2000s. Thus, the initial one-year education gap in age at first birth in the 1960s expanded to a seven-year gap by the 2000s.

Table 1 presents coefficients from a fixed-effects regression looking at effects of motherhood by cohort. These effects suggest that women who had their first birth in later decades had smaller motherhood earnings penalties than did women earlier cohorts. Women who had their first birth in the 1960s had nearly a 15% earnings penalty, which then declined to a 13% penalty for first births in the 1970s and an 11% penalty for first births in the 1980s. For women with first births in the 1990s, the motherhood penalty became a 2% earnings premium, growing to a 7% premium for first births in the 2000s. This may be driven by many factors, but how much of this decline in the penalty is related to rising age at first birth, given that the past literature has identified smaller motherhood penalties (and even motherhood premiums) for later transitions to motherhood? How does this vary by educational attainment, given that only college-educated women have delayed fertility? Next steps are to examine these cohort-specific motherhood effects by education and age at first birth and to apply the decomposition.

References

Amuedo-Dorantes, Catalina and Jean Kimmel. 2005. "The Motherhood Wage Gap for Women in the United States: The Importance of College and Fertility Delay." *Review of Economics of the Household* 3(1):17–48.

Autor, David H. 2014. "Skills, Education, and the Rise of Earnings Inequality among the 'Other 99 Percent.'" *Science* 344(6186):843–51.

Chandler, Timothy D., Yoshinori Kamo, and James D. Werbel. 1994. "Do Delays in Marriage and Childbirth Affect Earnings?" *Social Science Quarterly* 75(4):838–53.

Cherlin, Andrew J., Elizabeth Talbert, and Suzumi Yasutake. 2014. "Changing Fertility Regimes and the Transition to Adulthood: Evidence from a Recent Cohort." in *Annual Meeting of the Population Association of America*.

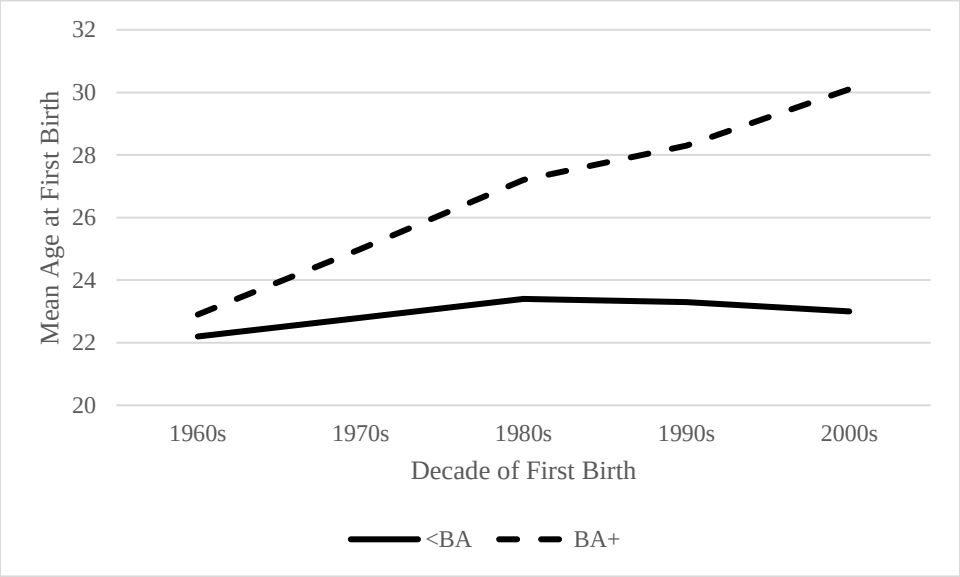
Doren, Catherine. 2018. "Do Some Mothers Pay a Higher Price? Variation in Motherhood Wage Penalties by Education, Parity, and Fertility Timing." *Presented at the Annual Meetings of the Population Association of America*. Denver, CO.

Herr, Jane Leber. 2015. "Measuring the Effect of the Timing of First Birth on Wages." *Journal of Population Economics* 1(May):39–72.

Miller, Amalia R. 2011. "The Effects of Motherhood Timing on Career Path." *Journal of Population Economics* 24(3):1071–1100.

Taniguchi, Hiromi. 1999. "The Timing of Childbearing and Women's Wages." *Journal of Marriage and Family* 61(4):1008–19.

Figure 1. Mean age at first birth by educational attainment in each decade from 1960-2000



Source: SIPP Synthetic Beta Data. U.S. Census Bureau. SIPP Synthetic Beta: Version 6.0 [Computer file]. Washington DC; Cornell University, Synthetic Data Server [distributor], Ithaca, NY, 2015. Not validated with Gold Standard Data.

Table 1. Fixed effects of motherhood on log earnings by decade of first birth

Variable	Coefficient
First birth in 1960s	-0.146 ***
	(0.005)
First birth in 1970s	-0.128 ***
	(0.004)
First birth in 1980s	-0.110 ***
	(0.003)
First birth in 1990s	0.017 ***
	(0.004)
First birth in 2000s	0.068 ***
	(0.008)
Person-Years	3,643,332
Respondents	151,465

Source: SIPP Synthetic Beta Data. U.S. Census Bureau. SIPP Synthetic Beta: Version 6.0 [Computer file]. Washington DC; Cornell University, Synthetic Data Server [distributor], Ithaca, NY, 2015. Not validated with Gold Standard Data.
Note: Includes time-variant controls for age and age squared. Standard errors in parentheses.
*** p<0.001, ** p<0.01, * p<0.05