IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17419

Income and Fertility of Female College Graduates in the United States

Zhengyu Cai
John V. Winters

OCTOBER 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17419

Income and Fertility of Female College Graduates in the United States

Zhengyu Cai
Southwestern University of Finance and Economics

John V. Winters
Iowa State University and IZA

OCTOBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Income and Fertility of Female College Graduates in the United States

Fertility rates have fallen below replacement levels in many economies. We examine the relationship between female incomes and fertility for college graduates in the United States. Female income is likely endogenous to fertility, and candidate instrumental variables are likely imperfect. We use the Nevo and Rosen (2012) imperfect instrumental variable procedure to estimate two-sided bounds for the effect of female income on fertility. The effect of female income on fertility is unambiguously negative and non-trivial, but the magnitude is relatively small. Our results suggest that the recent fertility slowdown in the U.S. is not primarily due to higher female incomes.

JEL Classification: J13, J16

Keywords: fertility, children, motherhood, female income

Corresponding author:
John V. Winters
Iowa State University
Department of Economics
460B Heady Hall
518 Farm House Lane
Ames, Iowa 50011-1054
USA
E-mail: winters1@iastate.edu

1. Introduction

Low and declining birth rates are a major problem for many high- and middle-income countries around the world including the United States (Kuehn 2019; Buckles, Guldi, and Schmidt 2019). Figure 1(a) shows a steady decline in the overall fertility rate in the United States from 2000 to 2019, which has consistently been below the replacement fertility rate of 2.1. Similarly, Figure 1(b) demonstrates a sharp decline in the number of children for American women aged 39 to 41 from 1970 to 1990 and then a relatively flat trend 1990 to 2019 at historically low levels. The recent decline in 1(a) is especially driven by younger women and will likely result in renewed further declines in the number of children for women ages 39 to 41 in coming years.[1] Declining birth rates cast doubt on whether the future workforce will be able to support an aging population (Stone 2018; Miyazaki 2018).[2] Economic development is expected to cause demographic transition, but fertility below replacement rate is especially worrisome and not an obvious or unavoidable result. Furthermore, women have made substantial labor market gains in recent decades, but significant gender gaps remain and childcare is an important factor (Goldin 2006; Juhn and McCue 2017). Women are still the primary caregivers in most households and face especially salient tradeoffs between time investments in career and family (Kim 2020; Craig and Churchill 2021; Kuziemko et al. 2018; Kleven, Landais, and Søgaard 2021). These tradeoffs produce concerns that

[1] The outcome in Figure 1(a) is a flow and that in 1(b) is a stock, so it takes time for changes in the former to show up in changes in the latter.

[2] In the U.S., the Medicare and Social Security trust funds are forecast to become depleted in the near future hindering the ability of the programs to provide long promised benefits unless tax rates are raised (Franck 2021).

future labor market gains for women may further lower fertility rates.

The theoretical effect of female income on fertility is ambiguous, as it depends on the relative strength of income and substitution effects. Higher female income may increase fertility by raising the affordability of children, or it may decrease fertility by raising the opportunity cost of childbearing. Empirical studies have used various methods and data sources to estimate this effect, but they often face challenges such as endogeneity, and instrumental variable strategies are standard (Schultz 1997).

Schultz (1997) discusses a common approach for estimating an exogenous wage profile throughout a female's lifecycle by employing instrumental variables to impute wages that are independent of prior time allocation choices, career paths, and fertility decisions. This approach is grounded in the human capital earnings model proposed by Mincer (1974). However, one would argue that if a woman's human capital simultaneously influences her fertility choices and her potential wage, then human capital does not satisfy exogeneity and cannot be a valid excluded instrumental variable. Cygan-Rehm and Maeder (2013) and DeCicca and Krashinsky (2023) both employed instrumental variables based on compulsory schooling laws to address the endogeneity issue in the impact of educational levels on fertility outcomes. Both studies find that the level of education has a negative influence on fertility outcomes. Moreover, both studies mentioned the impact of education on women's earnings as a potential mechanism. However, exogenous variation in education due to compulsory schooling is only one source of lifetime income variation, and perhaps not the most significant one.

Other researchers have explored alternative instrumental variables to address the endogeneity between female income and fertility. For example, Fleisher and Rhodes (1979) used husband's wage rate at age 40 and other demographic characteristic variables as instrumental variables to account for the endogeneity of female average lifetime wages and found a negative effect on fertility. Specifically, a $1,000 increase in female lifetime wage reduces the number of children by 0.4, with a partial elasticity of fertility with respect to wages being -0.43. However, the instrumental variable they employed may still be problematic, as there is generally positive assortative mating by income, and the effects on fertility of male and female income may be in opposite directions.

More recent studies have used instrumental variables at a more macro level to examine the impact of female labor market conditions or income on fertility outcomes. For instance, Schaller (2016) used a Bartik style industry shift-share instrument and found that improvements in women's labor market conditions are negatively correlated with birth rates at the state level, but the effect is small and not always significantly different from zero. Incorporating panel data across countries and employing international oil price shocks as an instrumental variable, Hailemariam (2024) estimated the impact of national per capita income on fertility rates. The author discovered that national per capita income generally exerts a negative and significant influence on the total fertility rate.

Overall, as concluded in the review by Jones, Schoonbroodt, and Tertilt (2010), the relationship between female incomes and fertility is overwhelmingly negative. However, there are still numerous concerns about identification and whether estimated results in previous literature represent unbiased causal effects. Female income is likely endogenously related to fertility, so ordinary least squares estimates are likely biased. Nevertheless, available instrument variables in the literature are likely imperfect. Nevo and Rosen (2012) show that under certain conditions, we can use imperfect instrumental variables to provide informative two-sided bounds on a causal effect of interest. Thus, we use the Nevo and Rosen (2012) imperfect instrumental variable procedure to provide two-sided bounds for the effect of earned income on fertility for female college graduates in the United States.

Specifically, we examine the effect of annual earned income on the total number of children at ages 39–41. We focus on female college graduates because college graduates are generally more career oriented than their non-college educated counterparts and are especially likely to face tradeoffs between career and family (Goldin 2006). Figures 1(a) and (b) show that for women with a bachelor's degree, both their total fertility rate and the number of children for women aged 39 to 41 have been almost consistently the lowest among all educational groups. Furthermore, education levels have risen substantially in recent decades, and 40.74 percent of women ages 39–41 in 2019 have a bachelor’s degree or higher level of education, so this is a large and important population to study. We examine women at ages 39–41 because this a point at which most women have completed their family size and before their older children generally move out, but we also examine robustness to older age ranges.

We find that the effect of female earnings on fertility is unambiguously negative for college graduate women. However, the effect magnitude is non-trivial but also relatively small. Our estimated bounds indicate that a 10 percent increase in female income decreases fertility by less than one percent, i.e. a decrease in the mean number of children from around 1.93 to 1.91. From a long-term perspective, our results explain 21.9% to 32.6% of the change in fertility outcomes among college-educated women aged 39 to 41 from 1970 to 2019. Our results suggest that while rising incomes for females have contributed to lower fertility among college graduates, this contribution is not dominant. Other factors such as social norms, childcare availability, and family policies may be more important in explaining the decreased fertility in recent years.

Our study relates to previous literature on fertility and the role of labor market outcomes. While a number of previous studies have examined the relationship between female incomes and fertility, relatively few have used instrumental variables and some earlier studies that do use instrumental variables make strong assumptions that may not be satisfied (Fleisher and Rhodes 1979; Schultz 1986). In particular, instruments may be imperfect and typical methods could lead to biased point estimates.

We make a novel contribution relative to previous literature by using imperfect instrumental variables to provide two-sided bounds of the causal effect of female income on fertility. To our knowledge, ours is the first study in the fertility literature to apply the Nevo and Rosen (2012) imperfect instrumental variable procedure. Our estimated bounds are informative. We provide clear and convincing evidence that female income has a negative effect on fertility that is meaningful but not very large.

2. Empirical Framework

We use the 2009–2019 American Community Survey (ACS) pooled cross-sectional microdata obtained from IPUMS (Ruggles et al. 2019) to investigate the effect of female earned income on fertility.[3] We measure a female's fertility using the number of own children she has living in her household when she is ages 39–41. Specifically, we restrict the analytical sample to native-born married college graduate females ages 39–41 with annual earned income whose spouses are native, male, and college graduates. Our study is intentionally focused on highly educated women with highly educated spouses (both with a bachelor's degree or higher) because of the hypothesized tradeoffs between career and family that are frequently discussed for these women (Kuziemko et al. 2018). We limit the sample to the native-born population to increase cultural homogeneity; cultural norms among different immigrant groups may result in different relationships depending on country of origin and length of time in the United States.

The ACS does not report total lifetime fertility, so we approximate it based on the

[3] The sample starts in 2009 because the construction of one of our instrumental variables requires information on individuals' college majors, which the ACS began to provide in 2009. The sample ends in 2019 because the COVID-19 pandemic that erupted in 2020 might have a systematic impact on the analysis.

number of own children in the woman's household at ages 39–41. We focus on ages 39–41 as a middle range of ages where most college graduate women have completed their family size yet their oldest children have yet to move out. Admittedly, some women have children very early who move out before the women are ages 39–41 and some women have children after ages 39–41, with the latter being especially notable among college graduates. Therefore, we also report additional analysis below for ages 42–44 and 45–47.

Another reason we restrict the sample to ages 39–41 is due to the independent variable of interest, namely, female income. As we are interested in the effect on lifetime fertility, lifetime income is a conceptually more appropriate independent variable for use in the empirical model. However, the ACS only provides information on current annual earnings. To calculate individual lifetime incomes, ideally, we would have a balanced panel dataset at the individual level covering all working ages. Yet, the ACS is pooled cross-sectional. It only allows us to track the same cohort over time, and the dataset is not a full work-life balanced panel for each cohort as we cannot observe the same age range for each cohort within our sample range. In a similar case, Abeysinghe and Gu (2011) show that the incomes of different cohorts in a pooled cross-sectional dataset can be decomposed into a life-cycle component and a cohort effect. For example, we can calculate the average incomes for age groups of 18–65 in the 2009 ACS sample. Yet, these average incomes not only contain the information of life-cycle component, but also the information of the cohorts born in 1944–1991 and observed in

2009. As we have the ACS samples from 2009 to 2019, we can repeat the calculation for each year and use a cohort fixed effect to isolate the life-cycle component. As the age range is set to 18–65, we can impute the life-cycle income profiles for 1944–2001 birth-year cohorts.[4] Specifically, we estimate the model below to impute an income profile from age 18 to 65 for every cohort of college educated women in our sample:

$$\log Income_{ijt} = \theta_0 + \sum \rho_i Age_i + \sum \delta_j Cohort_j + \varepsilon_{ijt} \quad (1)$$

where $Income_{it}$ is annual earning of female i at year t. $j = 1, \dots, J$ is the jth cohort. Age stands for age dummies which capture the life-cycle component, and $Cohort$ is cohort dummies which capture the cohort component. Following Abeysinghe and Gu (2011) and for simplicity, we assume that the life-cycle component does not vary with cohorts. The differences between cohorts only come from the cohort fixed effects.

The estimated lifetime income profiles for the youngest (2000), median (1972), and the oldest (1944) birth-year cohorts are presented in Figure 2. Consistent with the results in the literature (Abeysinghe and Gu 2011; Orazio P. Attanasio and Browning 1995; Attanasio and Weber 1995), the income profiles are hump-shaped and reach their maximums during the late 50s. The horizontal lines represent lifetime average incomes for the selected cohorts calculated using the income profiles. We can observe that the income profiles and lifetime average incomes intersect at ages 39–41 for the selected cohorts. As discussed above, the life-cycle component is the same for all the cohorts.

[4] We actually impute the income profiles for each cohort of 1944–2000 as there is no college graduate among the cohort of 2001 in our sample.

Thus, the annual incomes at 39–41 are at the lifetime average for all the cohorts. As our sample is restricted to ages 39–41, we can use annual earnings to proxy lifetime incomes.[5] Of course, any proxy has limitations. Income at ages 39-41 may be a good proxy for lifetime income on average but not for all individuals. Furthermore, its use implies that individuals form rational expectations at early ages about career earnings and behave accordingly in their fertility decisions. In reality, some individuals will over- or underestimate career earnings and the impact of their fertility on their future earnings. Finally, some individuals may not be sufficiently rational and forward-looking.

In our main analysis, we estimate variants of the model below:

$$Nchild_{it} = \alpha \, log \, Income_{it} + \boldsymbol{X}_{it}\beta + \boldsymbol{S}_{it}\gamma + \tau_t + \varepsilon_{it} \tag{2}$$

where $Nchild_{it}$ is number of children of female i at year t. $log \, Income_{it}$ is natural log annual total income earned by female i in year t. $\boldsymbol{X}_{it}$ is a set of detailed female characteristic control variables including dummy variables for 3 age groups (39 – 41), a dummy indicating whether living in one's home state, 6 categories for race/ethnicity (white, black, American Indian or Alaska native, Asian, other races, and Hispanic), 51 categories for birth state, 4 categories for education level (bachelor's degree, master's degree, professional degree beyond a bachelor's, doctoral degree), 3 categories for times married (once, twice, thrice or more), 27 categories for years

[5] Additionally, Appendix Table A1 provides evidence of correlation in mean hourly incomes across the life cycle by industry and college major. The correlations are generally large for both but somewhat higher by industry than college major. Furthermore, the correlations are smallest for workers ages 55-64, possibly consistent with some of these workers adopting strategies related to phased retirement such as changes of career fields and reductions in labor supply.

married, and 307 categories for local area[6] of current residence.[7] We also control for spousal characteristics, $\boldsymbol{S}_{it}$, including 50 categories for age,[8] a dummy for living in one's home state, 6 categories for race/ethnicity (white, black, American Indian or Alaska native, Asian, other races, and Hispanic), 51 categories for birth state, 4 categories for education level (bachelor's degree, master's degree, professional degree beyond a bachelor's, doctoral degree), 3 categories for times married (once, twice, thrice or more), 173 categories for detailed college major, 284 categories for occupation (3-digit), and 221 categories for industry categories (3-digit). τ_t is a set of year fixed effects. ε_{it} is an error term.

Reverse causality is a potential challenge to identifying the causal effect of interest, α. Having more children may reduce the time a woman spends on paid work and career investments and reduce her income. Additionally, there may exist some unobserved factors omitted from the regression equation that affect both fertility and income. For example, women from a conservative background may prefer a large family with a husband as the primary earner, while women with more liberal views may place greater emphasis on career and less on family size. Alternatively, some women may aim for both a high-income career and a large family; the high-income career may be viewed

---

[6] Local areas are defined as metropolitan statistical areas (MSAs) for individuals living in MSAs that are identified in the ACS. Individuals living in non-metropolitan areas and individuals in areas that cannot be specifically identified as part of a particular MSA are assigned to state-specific residual categories, i.e., one non-MSA group per state.

[7] For each characteristic, one group serves as the omitted reference category, so the number of dummy variables for that characteristic is one less than the number of categories for that characteristic.

[8] As shown in Table 1, the age range of the spouse spans from 24 to 94 years, but only includes 55 distinct age levels, specifically 24 to 75, 78, 81, and 94 years. Excluding the ages with only one observation, which are 24, 26, 72, 81, and 94 years, we controlled for 50 categories of spouse age in the regression analysis.

as necessary to finance significant child investments. Other unobservable factors could also be important. Unfortunately, we cannot observe all the factors affecting family size and cannot rule out bias in ordinary least squares (OLS) estimates. Furthermore, we cannot confidently sign the likely direction or magnitude of the overall bias *a priori*; the bias could be positive or negative and small or large.

In response to these challenges, we first employ a two-stage least squares (2SLS) instrumental variable (IV) estimation strategy. Specifically, we consider two instrumental variables for the endogenous regressor, namely, log annual earned income: mean log hourly income by industry during the previous three years (IV1) and mean log hourly income by college major during the previous three years (IV2).[9] We take a 3-year lag to construct the IVs and have two overlapping sample periods: 2009–2018 to construct the IVs and 2012–2019 for the analysis dataset.[10] To be clear, we do not have panel data; the ACS is a repeated cross-section. The instruments are constructed based on different women with the same industry and college major in the previous three years.

To be valid instruments, IV1 and IV2 should be strongly correlated with the endogenous regressor. This condition is testable. The major threat to instrument validity is the exclusion restriction, i.e., IV1 and IV2 should affect fertility only through log

[9] Hourly income is calculated using annual income divided by usual hours worked per week times 52 weeks for the sample of full-year workers.

[10] For example, we compute the 2009–2011 mean of the log hourly income by industry. Then, we match this lagged mean to the 2012 sample by industry. We repeat this process for the subsequent years and for college major. We use lagged mean log income to avoid mechanical endogeneity from having the same individuals form their own instrument. We use a three-year mean to smooth out sampling variation and short-term fluctuations. The college major variable in the ACS is first available in year 2009, which limits the time period for our analysis.

annual earned income. As IV1 and IV2 are constructed using lagged data and aggregated by industry and college major, respectively, it is unlikely that an individual's fertility would have a reverse causal effect on IV1 and IV2. However, IV1 and IV2 could still be jointly determined with the dependent variable. For example, a female with unobserved preferences for a high-income career and large family size may choose a high-income college major. After college graduation, she may sort into a high-income industry. Thus, IV1 and IV2 are likely to be imperfect instrumental variables (IIV).

To conduct reliable inference in this case, we utilize the procedure of Nevo and Rosen (2012) and use the imperfect IV1 and IV2 to estimate bounds and confidence intervals for the effect of female income on fertility relaxing the strict exclusion restriction condition. Nevo and Rosen (2012) show that under certain assumptions, analytical bounds can be calculated for endogenous parameters using IIV. They show that if the correlation between IIV and the endogenous regressor is negative, two-sided bounds can be obtained for the endogenous parameter. As we will show in the next section, both IV1 and IV2 are positively correlated with log annual total income. Thus, we cannot obtain two-sided bounds using IV1 or IV2 individually.

We follow Proposition 5 and Lemma 2 in Nevo and Rosen (2012) and simultaneously use IV1 and IV2 to construct two-sided bounds for the effect of female earned income on fertility. To produce two-sided bounds, the following assumptions and conditions in Nevo and Rosen (2012) need to be justified and/or tested:[11]

---

[11] Assumptions 1 (sampling process), 2 (exogenous control variables), and 5 (rank and order) in Nevo and Rosen (2012) are standard.

(1) Assumption 3: the correlation between the IIV and the error term should have the same direction as the correlation between the endogenous regressor and the error term. Two IIVs should both satisfy this normalization assumption. This assumption is not directly testable, but we can justify it using the OLS and 2SLS results below.

(2) New IIV $\omega(\gamma)$: we can construct a new IIV using the IIVs we have as $\omega(\gamma) = \gamma Z_2 - (1-\gamma)Z_1$, where $\gamma \in (0,1)$.[12] If $\omega(\gamma^*)$ also satisfies Assumption 3 and is negatively correlated with the endogenous regressor, it yields two-sided bounds for the parameter of interest. The existence of such a γ^* can be tested using Lemma 2 in Nevo and Rosen (2012).

(3) Existence of γ^*: following the second condition of Lemma 2, if we set $\gamma^* = 0.5$, the condition indicates that Z_1 is more correlated with the endogenous regressor and more exogenous than Z_2. The partial correlations between IIVs and the endogenous regressor is testable, and we will justify the relative exogeneity of our two IIVs in next section. Next, we test the third condition of Lemma 2 that $\sigma_{z_1\tilde{y}}\sigma_{\tilde{x}z_2} < \sigma_{z_2\tilde{y}}\sigma_{\tilde{x}z_1}$, where $\tilde{x}$ and $\tilde{y}$ are the residuals in the regressions of the endogenous regressor and dependent variable on exogenous covariates, respectively. If this condition holds, then $\omega(\gamma^* = 0.5)$ satisfies the assumptions and conditions described above in (2). We can use $\omega(\gamma^* = 0.5)$ to obtain two-sided bounds for the parameter of interest.

(4) Assumption 4: this assumption is not required by Proposition 5 and Lemma 2.

---

[12] As the order of the two IIVs matters, following Nevo and Rosen (2012), we use the notations Z_1 and Z_2 to distinguish the two IIVs. We will show that the IV1 and the IV2 are Z_1 and Z_2, respectively, in the next section.

Yet, if we can assume the two IIVs are more exogenous than the endogenous regressor, we can impose Assumption 4 to sharpen the bounds.

We will discuss in the next section the justification for our IIVs satisfying the assumptions and conditions discussed above. The summary statistics of the IIVs and the other variables are shown in Table 1.

3. Empirical Results

We first estimate the empirical model using ordinary least squares (OLS). Column (1) in Table 2 shows the results. A one unit increase in log annual earned income is significantly associated with a 0.091 decrease in the number of children.

Columns (2) to (4) in Table 2 report results from two-stage least squares (2SLS) estimation. Column (2) suggests that a one unit increase in log annual earned income yields a 0.103 decrease in fertility using mean log hourly income by industry (IV1) as the instrument. The first stage result is positive and strongly significant consistent with expectations; working in a high hourly income industry increases annual earned income.

Column (3) shows the 2SLS results using mean log hourly income by college major (IV2) as the instrument. The first-stage relationship is again positive and strongly significant, but the second-stage coefficient estimate (0.047) is noisily estimated and not statistically significant. Thus, despite their similarities in construction, IV1 and IV2 yield qualitatively different 2SLS results.

When we include both instruments simultaneously in 2SLS in Column (4), the second-stage results are similar to the 2SLS results using just IV1 in Column (2). The second-stage coefficient estimate in Column (4) is also identical to the OLS estimate in Column (1). The first-stage coefficients in Column (4) are both significantly positive but smaller than in Columns (2) and (3). However, the decreased first-stage coefficient for IV1 is modest while the decreased first stage coefficient for IV2 is more substantial. This implies that IV1 is more strongly correlated with female income than IV2, an issue that we return to more formally below.

As discussed above, IV1 and IV2 are both lagged and aggregated, so reverse causality problems are mitigated. However, we cannot rule out omitted variable bias as the number of children could be jointly affected by unobservable factors that also affect the choice of industry and/or college major. Neither instrument is likely to be perfectly exogenous. A likely explanation for the discrepancy in Columns (2) and (3) of Table 2 is that IV2 is more endogenous than IV1. A college graduate female who prefers a large family may choose a high-income college major to finance the considerable expenses of raising children, inducing a positive bias to the 2SLS estimate in Column (3). After college graduation, she may seek out a high-income industry, also inducing a positive bias to the 2SLS estimate in Column (2). However, she has less control over her industry than college major. Employers in high-paying industries may choose to not hire her, possibly due to discrimination. Even if she is initially hired in a high-paying industry, she may experience barriers to career advancement in the industry, especially as her preferences for a large family are revealed to employers. Thus, industry is less under her control than college major, and IV1 is likely less endogenous than IV2.

We next construct a set of Nevo and Rosen (NR) bounds for the causal effect of interest using IV1 and IV2. Since IV1 and IV2 are likely not perfectly exogenous, we follow Nevo and Rosen (2012) to relax the strict exclusion restriction. Because both IV1 and IV2 are positively correlated with the endogenous regressor, the NR procedure cannot provide two-sided bounds using one of them individually. Thus, we utilize Proposition 5 in Nevo and Rosen (2012) to construct two-sided bounds using both IV1 and IV2. The discussion below follows their Proposition 5 and Lemma 2 along with their notation. First, as discussed above, the potential endogeneity in the IIVs and in the endogenous regressor should have the same direction. The discussion above suggests that endogenous human capital investments and industry choice lead to positively biased 2SLS estimates, i.e., the IIVs are positively correlated with the error term. Furthermore, these endogenous choices are likely to drive positive bias in OLS estimates since human capital and industry are major determinants of earnings. Additionally, OLS results in Column (1) are more positive than 2SLS estimates in Column (2), so positive bias in Column (2) implies positive bias in Column (1), i.e., the endogenous regressor and both instruments are positively correlated with the error term. Thus, Assumption 3 is satisfied.

Next, to test whether the γ^* in Proposition 5 exists, we test the conditions in Lemma 2. In the second condition of Lemma 2, it is intuitive to set $\gamma^* = 0.5$, which implies that "the more relevant variable is also weakly better in terms of validity" (Nevo and Rosen 2012, 665). This seems to be the case in our analysis; IV1 is more relevant than IV2, and IV1 is arguably more exogenous than IV2 as discussed above. We formally test the partial correlations between IV1 and IV2 with the endogenous regressor, respectively, controlling for the exogenous covariates. Appendix Table A2 reports the results and shows that IV1 is more strongly correlated with the endogenous regressor than IV2. Thus, IV1 is the Z_1 in Nevo and Rosen (2012) and IV2 is the Z_2. Next, we test the third condition of Lemma 2. Table A1 shows that $\sigma_{z_1\tilde{y}}\sigma_{\tilde{x}z_2} < \sigma_{z_2\tilde{y}}\sigma_{\tilde{x}z_1}$, so that the condition is satisfied, and hence such a γ^* exists. Finally, Table A1 shows that the partial correlation between endogenous regressor and $\omega(\gamma^* = 0.5)$ is negative controlling for the exogenous covariates. Thus, we can use the NR procedure to obtain two-sided bounds.

Columns (5) and (6) in Table 2 report the NR bounds for the coefficient of interest using IV1 and IV2 as IIVs. Column (5) imposes Assumption 4 and indicates that a one unit increase in log annual earned income causes 0.117–0.174 decrease in fertility. Column (6) shows the results without Assumption 4. It also shows that a one unit increase in log annual earned income causes 0.117–0.174 decrease in fertility. Thus, the impact of income on fertility is strictly negative. The OLS estimate and 2SLS estimate using IV2 in Table 2 are positively biased. The 2SLS estimate using IV1 is more closely comparable to the bounds than the other estimates.

At this point, it is important to consider the estimated magnitudes. The effect is clearly negative, but is it large? The dependent variable is the number of children and has a mean of 1.925. The endogenous regressor is log annual earned income, so we can think of a 0.1 increase in the regressor as an approximately 10 percent increase in income. Thus, a 10 percent increase in income implies a reduction in fertility of 0.0117–0.0174. In other words, a 10 percent increase in average female income would decrease mean fertility from 1.925 to around 1.908–1.913. Additionally, dividing the coefficient bounds by mean fertility implies a negative elasticity magnitude smaller than 0.1.[13] Thus, the effect of female income on fertility is clearly negative and non-trivial, but the magnitude is not very large. Looking at the long-term changes in fertility outcomes, women's income is also not the dominant factor. The log change in the average real income of college-educated women aged 39 to 41 from 1970 to 2019 is 0.450199.[14] Multiplying 0.450199 by the estimated effects in Table 2, -0.117 and -0.174, we get impacts on the number of children during this period to be -0.05267 and -0.07833, respectively. The change in the number of children from 1970 to 2019 is -0.24028, so the change explained by our results is 21.9% to 32.6%. Therefore, women's income is not the dominant factor in explaining the long-term changes in fertility outcomes.

We also examine multiple alternative samples to consider the sensitivity of the main results in Columns (5) and (6) of Table 2. Specifically, Table 3 reports Nevo-Rosen bounds using our two IIVs with and without imposing Assumption 4 as before but with slightly different sample inclusion criteria. In Columns (1) and (2) of Table 3,

---

[13] Dividing -0.174 and -0.117 by 1.925 yields an elasticity range of -0.0904 to -0.0608.

[14] The data for 1970 is sourced from the Census 1970 1% Form 1 State sample, while the data for 2019 is derived from the ACS 2019 sample.

we examine the sample of women ages 42–44 but meeting all other main sample inclusion criteria. Similarly, Columns (3) and (4) examine women ages 45–47 meeting all other criteria. Columns (5) and (6) expand the main sample to include all native-born married college graduate women ages 39–41 with positive annual earned income whose spouses are native-born men but imposing no spousal education sample restriction; i.e., the main sample is expanded to include women with non-college educated spouses, so long as they meet all other main sample criteria. Finally, Columns (7) and (8) include all native-born college graduate women ages 39–41 with positive annual earned income, regardless of marital status and spousal characteristics.

As noted above, we focus the main analysis on ages 39–41 as the preferred sample because this is an age at which most women have completed their childbearing yet their older children are unlikely to have left home. However, it is also useful to consider slightly older age ranges given that some college graduate women do have children at later ages. The results in Columns (1) – (4) are qualitatively similar to the main results, though the bounds are actually somewhat tighter for the older age groups. Thus, the main conclusions are not considerably affected by focusing on ages 39–41.

We also discuss above that our primary interest is on highly educated women with highly educated spouses because of the frequently discussed tradeoffs between career and family for these women. Focusing on a more homogenous sample also reduces concerns about unobservable omitted variables that could distort results. Still, college graduate women with less educated spouses and unmarried women are also of potential interest, so we expand the sample in Columns (5) – (8) of Table 3 to consider the sensitivity of the main results. The bounds for these expanded samples are qualitatively similar to those for the main sample in Columns (5) and (6) of Table 2. Thus, we *a priori* chose to focus on the sample of college graduate women married to college graduate men but expanding the sample to those with less educated spouses or unmarried does not greatly alter the results.

Furthermore, we attempted to explore the heterogeneity of the fertility effect of women's income across different education groups. Based on technical feasibility, we made the following two attempts. First, we used the main IIV procedure and divided the sample into two subsamples: those with only a bachelor's degree and those with a graduate degree (master’s, professional, or doctorate). The reason for doing this is that the college major information for individuals with a bachelor's degree and above is available, allowing the analysis method to remain consistent with our main analysis. Columns (1) to (4) of Table A3 report these results. The estimated bounds of the two subsamples exhibit some differences in magnitudes, with the negative effect of income seeming to be slightly larger for the subsample with only a bachelor's degree; however, the 95% confidence intervals for the bounds (in parentheses) overlap considerably indicating that the results for the two samples are not statistically significantly different. Second, due to the lack of college major information for the group with less than a bachelor's degrees, we cannot construct IV2. Nevertheless, in order to examine the differences in income effects between high and low education groups, we attempted to use IV1 as a single instrumental variable for the three groups: above bachelor's (graduate degree), bachelor's, and below bachelor's.[15] All three groups have industry information, which can be used to construct IV1. Additionally, according to Table 2, the results using IV1 as an instrumental variable are the closest to our preferred NR bounds. Columns (5) to (7) of Table A3 report the results for the three subsamples: above bachelor's, bachelor's, and below bachelor's. It suggests that the negative effects exhibit an increasing pattern in terms of the absolute value of the estimates, but the standard errors indicate overlapping confidence intervals and that the coefficient estimates are not statistically significantly different from each other. Thus, we cannot provide precise evidence on differing impacts of female income on fertility, but there is some suggestive evidence of more negative effects for less educated women than for more educated women.

4. Conclusion

Improvements in female labor market opportunities have contributed to both higher household incomes and lower fertility levels. Fertility rates have fallen below replacement levels in many economies and further improvements in female income opportunities may lower fertility even more. However, the causal effect of female incomes on fertility is difficult to quantify because of potential reverse causality and omitted variable bias. We apply the Nevo and Rosen (2012) imperfect instrumental

[15] Sample inclusion criteria are otherwise the same as our main analysis: native married females ages 39–41 with positive annual earned income and whose spouses are native, male, and college graduates.

variable procedure to compute two-sided bounds for the effect of female earnings on fertility. Our results indicate that a 10 percent increase in a woman's income decreases her number of children by roughly 0.01 to 0.02 relative to a sample mean of 1.93. This effect is non-trivial but relatively small in magnitude. The negative effect is highly consistent with the existing literature (Jones, Schoonbroodt, and Tertilt 2010), but our estimated range appears to be much smaller than the elasticity obtained in existing papers that employ instrumental variable methods (Fleisher and Rhodes 1979). This discrepancy may be attributed to differences in samples and methodologies, and our estimated range should more accurately reflect the current reality. Our results also show that the increase in real income of college-educated women aged 39 to 41 from 1970 to 2019 explains 21.9% to 32.6% of the decline in fertility outcomes during this period, and while important, it does not appear to be a dominant factor. It should be noted that related studies, including the current one, typically investigate the impact of relative changes in lifecycle income or labor market conditions (Schultz 1997; Fleisher and Rhodes 1979), rather than general shifts in the overall wealth level. The effect of wealth levels on fertility outcomes may differ, which is a gap that needs to be filled in future research.

Overall, the relatively modest magnitudes that we estimate imply that the recent fertility slowdown in the U.S. is not primarily due to higher female incomes. Other factors such as social norms, childcare availability, and family policies may be more important in explaining the decreased fertility in recent years. Collins (2019) explores how intensive parenting norms exert pressure on working mothers, especially in countries with strong cultural expectations around motherhood. Regardless of the presence or absence of national-level family policies, this pressure for highly educated women is linked to the desire to achieve both career success and the ideal of perfect parenting, which can lead to work-family conflicts. This may help explain the low fertility rates among college-educated women, with future research needed to provide broader evidence.

References

Abeysinghe, Tilak, and Jiaying Gu. 2011. "Lifetime Income and Housing Affordability in Singapore." *Urban Studies* 48 (9): 1875–91. https://doi.org/10.1177/0042098010380956.

Attanasio, Orazio P., and Guglielmo Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 103 (6): 1121–57. https://doi.org/doi.org/10.1086/601443.

Buckles, Kasey, Melanie Guldi, and Lucie Schmidt. 2019. "Fertility Trends in the United States, 1980-2017: The Role of Unintended Births." w25521. NBER Working Paper. Cambridge, MA. https://doi.org/10.3386/w25521.

Collins, Caitlyn. 2019. *Making Motherhood Work: How Women Manage Careers and Caregiving*. Princeton: Princeton University Press. https://doi.org/https://doi.org/10.1515/9780691185156.

Craig, Lyn, and Brendan Churchill. 2021. "Working and Caring at Home: Gender Differences in the Effects of Covid-19 on Paid and Unpaid Labor in Australia." *Feminist Economics* 27 (1–2): 310–26. https://doi.org/10.1080/13545701.2020.1831039.

Cygan-Rehm, Kamila, and Miriam Maeder. 2013. "The Effect of Education on Fertility: Evidence from a Compulsory Schooling Reform." *Labour Economics* 25 (December): 35–48. https://doi.org/10.1016/j.labeco.2013.04.015.

DeCicca, Philip, and Harry Krashinsky. 2023. "The Effect of Education on Overall Fertility." *Journal of Population Economics* 36 (1): 471–503. https://doi.org/10.1007/s00148-022-00897-y.

Fleisher, Belton M, and George F Rhodes. 1979. "Fertility, Women's Wage Rates, and Labor Supply." *The American Economic Review* 69 (1): 14–24. http://www.jstor.org/stable/1802493.

Franck, Thomas. 2021. "Social Security Trust Funds Now Projected to Run out of Money Sooner than Expected Due to Covid, Treasury Says." CNBC News. 2021. https://www.cnbc.com/2021/08/31/social-security-trust-funds-set-to-be-depleted-sooner-than-expected.html.

Goldin, Claudia. 2006. "The Quiet Revolution That Transformed Women's Employment, Education, and Family." *The American Economic Review* 96 (2): 1–21. http://www.jstor.org/stable/30034606.

Hailemariam, Abebe. 2024. "Income and Differential Fertility: Evidence from Oil Price Shocks." *Journal of Demographic Economics* 90 (1): 31–54. https://doi.org/10.1017/dem.2022.8.

Jacob A. Mincer. 1974. *Schooling, Experience, and Earnings*. National Bureau of Economic Research. https://www.nber.org/books-and-chapters/schooling-experience-and-earnings.

Jones, Larry E., Alice Schoonbroodt, and Michèle Tertilt. 2010. "Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?" In *Demography* *and the Economy*, edited by John B. Shoven, 43–100. University of Chicago Press. http://www.nber.org/chapters/c8406.

Juhn, Chinhui, and Kristin McCue. 2017. "Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts." *Journal of Economic Perspectives* 31 (1): 183–204. https://doi.org/10.1257/jep.31.1.183.

Kim, Rae Yule. 2020. "An Unforeseen Story of Alpha-Woman: Breadwinner Women Are More Likely to Quit the Job in Work-Family Conflicts." *Applied Economics* 52 (55): 6009–21. https://doi.org/10.1080/00036846.2020.1781775.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2021. "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families." *American Economic Review: Insights* 3 (2): 183–98. https://doi.org/10.1257/aeri.20200260.

Kuehn, Bridget. 2019. "Declining US Fertility." *JAMA* 322 (10): 919. https://doi.org/10.1001/jama.2019.13200.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington. 2018. "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" w24740. NBER Working Paper. Cambridge, MA. https://doi.org/10.3386/w24740.

Miyazaki, Takeshi. 2018. "Examining the Relationship between Municipal Consolidation and Cost Reduction: An Instrumental Variable Approach." *Applied Economics* 50 (10): 1108–21. https://doi.org/10.1080/00036846.2017.1352077.

Nevo, Aviv, and Adam M. Rosen. 2012. "Identification With Imperfect Instruments."

Review of Economics and Statistics 94 (3): 659–71. https://doi.org/10.1162/REST_a_00171.

Orazio P. Attanasio, and Martin Browning. 1995. “Consumption over the Life Cycle and over the Business Cycle.” *American Economic Review* 85 (5): 1118–37. https://doi.org/www.jstor.org/stable/2950978.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2019. “IPUMS USA: Version 9.0 [Dataset].” MN: IPUMS. https://doi.org/https://doi.org/10.18128/D010.V9.0.

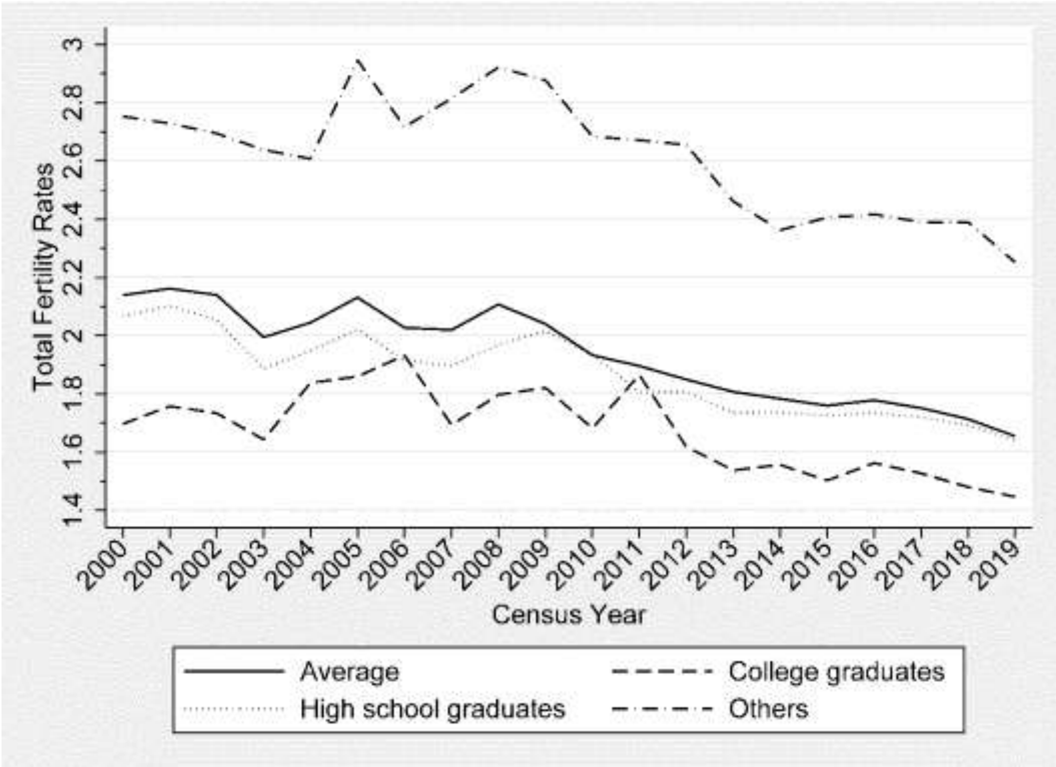
Schaller, J. 2016. “Booms, Busts, and Fertility: Testing the Becker Model Using Gender-Specific Labor Demand.” *Journal of Human Resources* 51 (1): 1–29. https://doi.org/10.3368/jhr.51.1.1.

Schultz, T. Paul. 1986. “The Value and Allocation of Time in High-Income Countries: Implications for Fertility.” *Population and Development Review* 12: 87. https://doi.org/10.2307/2807895.

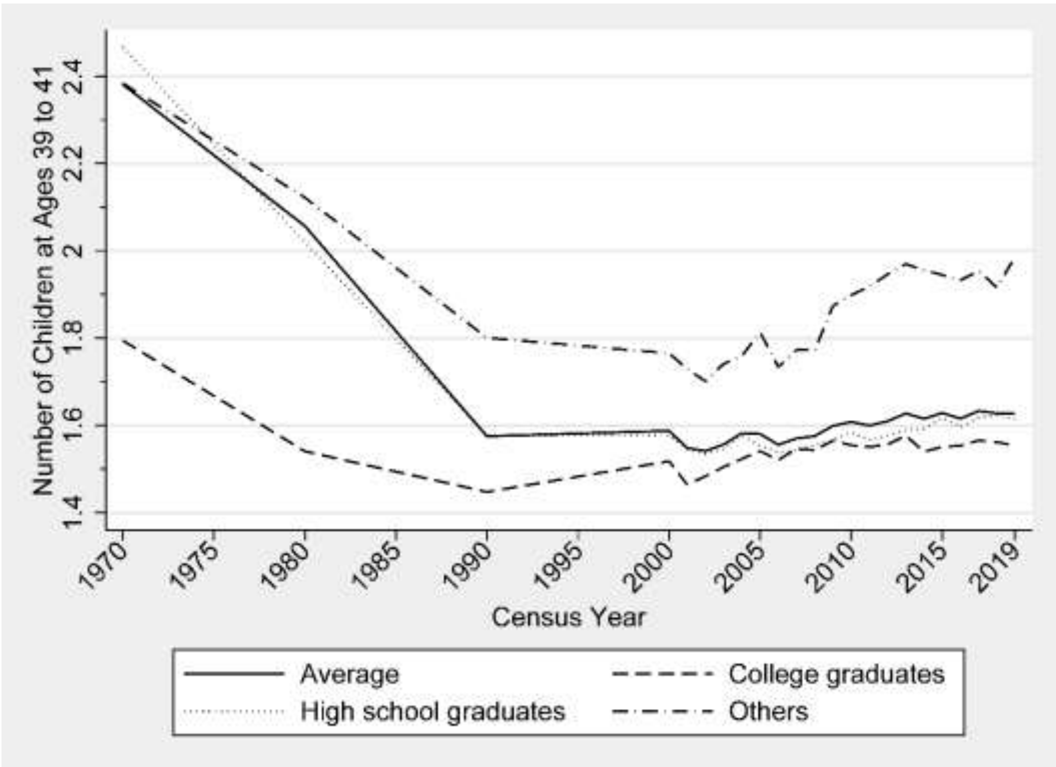
———. 1997. “Demand for Children in Low Income Countries.” In *Handbook of Population and Family Economics*, edited by M.R. Rosenzweig and O. Stark, 349–430.

Stone, Lyman. 2018. “Declining Fertility in America.” https://www.aei.org/wp-content/uploads/2018/12/Declining-Fertility-in-America.pdf.

Figure 1: Fertility Outcomes Among Women in the United States



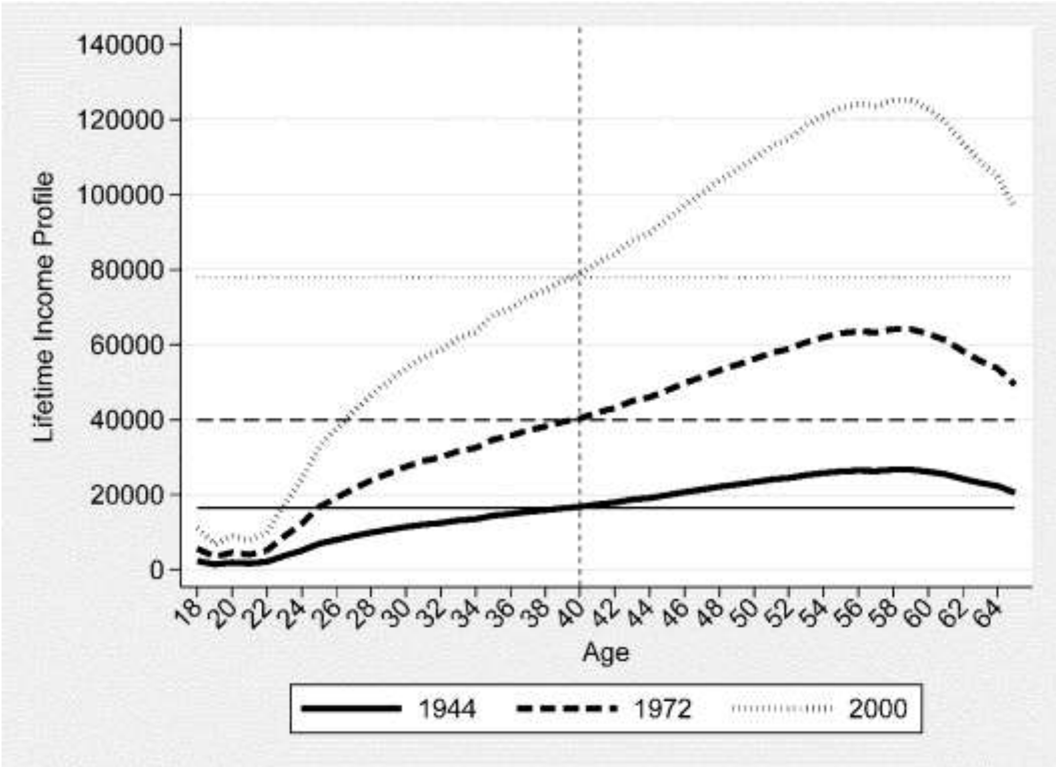
(a) Total Fertility Rates



(b) Number of Own Children at Ages 39 to 41

Notes: The analytical sample for (a) is derived from the ACS samples from 2000 to 2019. The analytical sample starts from the year 2000 because the key variables for calculating the total fertility rate became available from that year. The analytical sample for (b) is sourced from the 1% Form 1 State sample of the 1970 Census, the 1980 to 2000 Census 5% samples, and the ACS samples from 2001 to 2019. College graduates are defined as those who have at least obtained a bachelor's degree, including those with graduate experience or holding a graduate degree; high school graduates include those who have obtained a high school diploma but have not obtained a bachelor's degree, this encompasses those with an associate's degree and those who have attended some college without obtaining a degree; “others” indicates all groups without a high school diploma, including those who have received no education. The personal survey weights are used for the calculation.

Figure 2: Lifetime Income Profiles for Selected Cohorts



Notes: The analytical sample of the figure is restricted to college educated females from age 18 to 65. The thick solid, dash, and dotted lines are the lifetime income profiles for the oldest (1944), median (1972), and the youngest (2000) cohorts in our sample, respectively. The thin solid, dash, and dotted horizontal lines are lifetime average incomes. The personal survey weights are used for the calculation.

Table 1: Summary Statistics for Analytical Sample

	No. Obs	Mean	Std. Dev.	Min	Max
Female variables					
Number of children	49,834	1.925	1.085	0	9
Log annual total earned income	49,834	10.739	1.194	0.063	13.905
Mean log hourly income by industry	49,834	3.184	0.225	2.168	3.767
Mean log hourly income by major	49,834	3.192	0.160	2.696	3.820
Age	49,834	40.003	0.813	39	41
Home state resident	49,834	0.529		0	1
White	49,834	0.859		0	1
Black	49,834	0.052		0	1
American Indian or Alaska native	49,834	0.002		0	1
Asian	49,834	0.024		0	1
Other races	49,834	0.016		0	1
Hispanic	49,834	0.047		0	1
Bachelor's degree	49,834	0.525		0	1
Master's degree	49,834	0.356		0	1
Professional degree	49,834	0.077		0	1
Doctoral degree	49,834	0.042		0	1
Married once	49,834	0.884		0	1
Married twice	49,834	0.107		0	1
Married thrice (or more)	49,834	0.009		0	1
Spousal variables					
Age	49,834	41.904	4.015	24.000	94.000
Home state resident	49,834	0.517		0	1
White	49,834	0.873		0	1
Black	49,834	0.058		0	1
American Indian or Alaska native	49,834	0.002		0	1
Asian	49,834	0.016		0	1
Other races	49,834	0.014		0	1
Hispanic	49,834	0.037		0	1
Bachelor's degree	49,834	0.603		0	1
Master's degree	49,834	0.274		0	1
Professional degree	49,834	0.079		0	1
Doctoral degree	49,834	0.043		0	1
Married once	49,834	0.872		0	1
Married twice	49,834	0.115		0	1
Married thrice (or more)	49,834	0.014		0	1

Notes: Our analytical sample is restricted to native-born married college graduate females ages 39–41 with positive annual total earned income whose spouses are native, male, and college graduates. Summary statistics use survey weights to ensure national representativeness. Statistics for state of birth, years married, MSA, college major, spousal occupation, spousal industry, and year dummies are not reported to conserve space.

Table 2: Female Income and Fertility Results

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) IIV	(6) IIV
Log annual total earned income	-0.091***	-0.103***	0.047	-0.091**	[-0.174 -0.117]	[-0.174 -0.117]
	(0.006)	(0.034)	(0.099)	(0.040)	(-0.218 -0.085)	(-0.218 -0.080)
2SLS First Stage Results:						
Mean log hourly income by industry (IV1)		1.828***		1.731***		
		(0.131)		(0.144)		
Mean log hourly income by college major (IV2)			1.183***	0.482***		
			(0.078)	(0.145)		
2SLS First-Stage F-Statistics		194.932	229.040	114.949		
Endogeneity F-Statistic		0.167	1.781	1.169		
Endogeneity P-value		0.682	0.182	0.280		
Nevo-Rosen Assumption 4	N/A	N/A	N/A	N/A	Yes	No
N	49,834	49,834	49,834	49,834	49,834	49,834

Notes: The dependent variable is the number of own children in the household. Our analytical sample is restricted to native married college graduate females ages 39–41 with positive annual earned income whose spouses are native, male, and college graduates. For space conservation, we only report the estimated coefficients of the independent variable of interest, log total annual earned income (2019 Dollars). Though not reported, all columns include dummies for home state residential status (own and spousal), age (own and spousal), race (own and spousal), birthplace (own and spousal), detailed education (own and spousal), times married (own and spousal), survey year, MSA (and non-MSA), years of being married, detailed spousal college major, spousal occupation (3-digit), and spousal industry (3-digit). Numbers in brackets in Columns (5) and (6) are Nevo-Rosen bounds. The Nevo-Rosen bounds are estimated using both instruments in Column (4) as imperfect instrumental variables (IIV): mean log hourly income by industry and mean log hourly income by college major are used as the first and second IIVs to satisfy the condition (2) of Lemma 2 in Nevo and Rosen (2012), assuming $\gamma^* = 0.5$. Our Nevo-Rosen specifications are tested to be satisfied with condition (3) of Lemma 2, which indicates that γ^* exists. Columns (5) and (6) report the bounds with and without Assumption 4 in Nevo and Rosen (2012), respectively. Assumption 4 states that the correlation between the instruments and the unobserved error term is less than the correlation between the endogenous regressor and the error term. Standard errors in parentheses are clustered at detailed college major level in Columns (1) and (3), clustered at industry (3-digit) level in Column (2), and two-way clustered at detailed college major and industry (3-digit) level in Column (4). Two-way clustered standard errors at 3-digit industry and detailed college major level are used to calculate the 95% confidence intervals in parentheses for the Nevo-Rosen bounds in Columns (5) and (6). Individual survey weights are used.

** $p < 0.05$, *** $p < 0.01$.

Table 3: IIV Results for Alternative Samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ages 42–44		Ages 45–47		No Spouse Education Sample Restriction		No Marital Status Sample Restriction	
Log annual total earned income	[-0.213 -0.136]	[-0.213 -0.136]	[-0.161 -0.119]	[-0.161 -0.119]	[-0.213 -0.125]	[-0.213 -0.122]	[-0.216 -0.124]	[-0.216 -0.124]
	(-0.257 -0.111)	(-0.257 -0.095)	(-0.207 -0.095)	(-0.207 -0.089)	(-0.352 -0.091)	(-0.352 -0.086)	(-0.262 -0.088)	(-0.262 -0.082)
Assumption 4	Yes	No	Yes	No	Yes	No	Yes	No
N	50,171	50,171	50,208	50,208	80,106	80,106	93,568	93,568

Notes: The dependent variable is the number of own children in the household. The analytical samples are restricted as in Table 2 except Columns (1) and (2) are restricted to female ages 42–44 and 45–47, respectively, Column (3) drops the spousal education sample restriction, and Column (4) drops the marital status sample restriction and all spousal characteristic restrictions. Specifications are otherwise similar to Table 2 Columns (5) and (6). Two-way clustered standard errors at 3-digit industry and detailed college major level are used to calculate the 95% confidence intervals in parentheses for the Nevo-Rosen bounds. Individual survey weights are used.

Table A1: Hourly Income Correlation Matrix Across Ages by Industry or College Major

A. By industry	25-34	35-44	45-54	55-64
25-34	1			
35-44	0.9543	1		
45-54	0.9096	0.9365	1	
55-64	0.6872	0.6890	0.6032	1
N		224		
B. By major	**25-34**	**35-44**	**45-54**	**55-64**
25-34	1			
35-44	0.8443	1		
45-54	0.5135	0.6735	1	
55-64	0.4298	0.3658	0.2546	1
N		176		

Notes: The analytical sample is restricted to native college graduate females from 2009 to 2019 ACS samples. Average hourly incomes are calculated by industry (3-digit; Panel A) and by college major (3-digit; Panel B) at the ages of 25-34, 35-44, 45-54, and 55-64; individual survey weights are used.

Table A2: Auxiliary Tests for Nevo-Rosen Procedure in Table 2

	Mean log hourly income by industry (IV1)	Mean log hourly income by college major (IV2)
Partial correlation between endogenous regressor and IIV	0.347	0.158
$\sigma_{z_1\tilde{y}}\sigma_{\tilde{x}z_2} < \sigma_{z_2\tilde{y}}\sigma_{\tilde{x}z_1}$	Yes	
Partial correlation between endogenous regressor and $\omega(\gamma^* = 0.5)$	-0.226	

Notes: The partial correlations are controlling for the same set of covariates as in the main analysis. Our analytical sample is restricted to native married college graduate females ages 39–41 with positive annual earned income whose spouses are native, male, and college graduates. Individual survey weights are used.

Table A3: Female Income and Fertility Results

	(1) IIV: Graduate degree holder	(2) IIV: Graduate degree holder	(3) IIV: Bachelor's degree holder	(4) IIV: Bachelor's degree holder	(5) IV1: Graduate degree holder	(6) IV1: Bachelor's degree holder	(7) IV1: Without a bachelor's degree
Log annual total earned income	[-0.139 -0.090]	[-0.139 -0.074]	[-0.185 -0.131]	[-0.185 -0.118]	-0.074*	-0.118***	-0.165**
	(-0.219 -0.061)	(-0.219 -0.054)	(-0.234 -0.091)	(-0.234 -0.087)	(0.040)	(0.033)	(0.065)
2SLS First Stage Results:							
Mean log hourly income by industry (IV1)					1.739***	1.882***	1.429***
					(0.156)	(0.149)	(0.124)
2SLS first-stage F-statistics					124.979	160.115	132.204
Endogeneity Chi-squared-statistic					0.008	0.249	0.289
Endogeneity P-value					0.929	0.618	0.591
Nevo-Rosen Assumption 4	Yes	No	Yes	No	N/A	N/A	N/A
N	24,022	24,022	25,812	25,812	24,022	25,812	13,043

Notes: The dependent variable is the number of own children in the household. Our analytical samples are restricted to native married females ages 39–41 with positive annual earned income who attain the educational levels described in each table header and whose spouses are native, male, and college graduates. For space conservation, we only report the estimated coefficients of the independent variable of interest, log total annual earned income (2019 Dollars). Though not reported, all columns include dummies for home state residential status (own and spousal), age (own and spousal), race (own and spousal), birthplace (own and spousal), detailed education (own and spousal), times married (own and spousal), survey year, MSA (and non-MSA), years of being married, detailed spousal college major, spousal occupation (3-digit), and spousal industry (3-digit). Numbers in brackets in Columns (1) to (4) are Nevo-Rosen bounds. The Nevo-Rosen bounds are estimated using the same imperfect instrumental variables (IIV) in Table 2: mean log hourly income by industry and mean log hourly income by college major are used as the first and second IIVs to satisfy the condition (2) of Lemma 2 in Nevo and Rosen (2012), assuming $\gamma^* = 0.5$. The Nevo-Rosen specifications are tested to be satisfied with condition (3) of Lemma 2, which indicates that γ^* exists. Columns (1) and (3) report the bounds with Assumption 4 in Nevo and Rosen (2012), while Columns (2) and (4) report the bounds without Assumption 4. Assumption 4 states that the correlation between the instruments and the unobserved error term is less than the correlation between the endogenous regressor and the error term. Columns (5) to (7) reports the 2SLS estimates using IV1 as instrumental variable. Two-way clustered standard errors at 3-digit industry and detailed college major level are used to calculate the 95% confidence intervals in parentheses for the Nevo-Rosen bounds in Columns (1) to (4). Standard errors in parentheses are clustered at industry (3-digit) level in Columns (5) to (7). Individual survey weights are used.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.