Forthcoming

Economic Journal

The Effect of Fertility on Mothers' Labour Supply over the Last Two Centuries

Daniel Aaronson, Rajeev Dehejia, Andrew Jordan, Cristian Pop-Eleches, Cyrus Samii, and Karl Schulze

March 2020Abstract

Using a compiled dataset of 441 censuses and surveys between 1787 and 2015, representing 103 countries and 51.4 million mothers, we find that: (1) the effect of fertility on labour supply is typically indistinguishable from zero at low levels of development and large and negative at higher levels of development; (2) the negative gradient is stable across historical and contemporary data; and (3) the results are robust to identification strategies, model specification, and data construction and scaling. Our results are consistent with changes in the sectoral and occupational structure of female jobs and a standard labour-leisure model.

Aaronson: Federal Reserve Bank of Chicago (daaronson@frbchi.org); Dehejia: New York University (rajeev@dehejia.net); Jordan: University of Chicago (awjordan@uchicago.edu); Pop-Eleches: Columbia University (cp2124@columbia.edu); Samii: New York University (cds2083@nyu.edu); Schulze: Princeton University (kschulze@princeton.edu). We thank seminar participants at the NBER Summer Institute, SOLE, Williams, Cornell, the University of Washington, the University of Illinois Chicago, the University of Connecticut, and GREQAM and especially Quamrul Ashraf and Leah Boustan for invaluable suggestions. The views expressed in this paper are not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

I. Introduction

The relationship between fertility and female labour supply is widely studied in economics. For example, the link between family size and mothers' work decisions has helped explain household time allocation and the evolution of women's labour supply, particularly among rapidly growing countries in the second half of the 20th century (e.g. Angrist and Evans, 1998 ; Cristia, 2008). Development economists relate the fertility-work relationship to the demographic transition and study its implications on economic growth (Bloom *et al.*, 2001). Yet despite the centrality of these issues in the social sciences, the existing evidence is fragmentary and, as we discuss below, seemingly contradictory.

Our contribution is to provide unified evidence on whether the relationship between fertility and labour supply has evolved over time and with the process of economic development. Using data spanning not only a broad cross-section of countries at various stages of development but historical examples from currently developed countries dating back to the late 18th century, we show a strikingly consistent albeit evolving relationship between fertility and mothers' labour supply. To provide consistent estimates over time and space, we use two common instrumental variables strategies: (i) twin births introduced by Rosenzweig and Wolpin (1980) and (ii) the gender composition of the first two children (Angrist and Evans, 1998). We implement these estimators using four large databases of censuses and surveys: the International Integrated Public Use Micro Sample (IPUMS), the U.S. IPUMS, the North Atlantic Population Project, and the Demographic and Health Surveys. Together, the data cover 441 country-years, and 51.4 million mothers, stretching from 1787 to 2015 and, consequently, a large span of economic development.

A natural starting point in thinking about the fertility-labour supply relationship is Angrist and Evans (1998). Based on U.S. IPUMS data from 1980 and 1990, Angrist and Evans document a negative effect of fertility on female labour supply using both gender mix and twin births as instruments for subsequent children, a result also established by Bronars and Grogger (1994).[1] Alternative instruments that rely on childless mothers undergoing infertility treatments in the U.S. and Denmark (Cristia, 2008; Lundborg *et al.*, 2017) or natural experiments like the

[1] Bhalotra and Clarke (2016) and Clarke (2018) provide useful summaries of the validity of various fertility instruments and the broader empirical literature.

introduction of birth control pills (Bailey, 2013) or changes in abortion legislation (Bloom *et al.*, 2009) similarly conclude that children have a negative effect on their mother's labour supply or earnings. That the results are consistent across instruments is notable since each IV uses a somewhat different subpopulation of compliers to estimate a local average treatment effect, and therefore is suggestive of wide external validity.

However, we show that the negative relationship between fertility and mother's work behaviour holds only for countries at a later stage of economic development. At a lower level of income, including the U.S. and Western European countries prior to WWII, there is no causal relationship between fertility and mothers' labour supply. The lack of a negative impact at low levels of development aligns with Aguero and Marks' (2008, 2011) studies of childless mothers undergoing infertility treatments in 32 developing countries, Godefroy's (2017) analysis of changes to women's legal rights in Nigeria, and Heath (2017) who finds an economically small effect of fertility on women working using non-experimental evidence from urban Ghana. Strikingly, combining U.S historical censuses with data from a broad set of contemporary developing countries, we find that the negative gradient of the fertility-labour supply effect with respect to economic development is remarkably consistent across time and space. That is, women in the U.S. at the turn of the 20th century make the same labour supply decision in response to additional children as women in developing countries today. We show that the negative gradient is robust to a wide range of data, sampling, and specification issues, including alternative instruments, development benchmarks, sample specification criteria, conditioning covariates including those highlighted by Bhalotra and Clarke (2016), additional measures of mother's labour supply, and a variety of other adjustments to make our data historically consistent.

That said, our main results come with important qualifications, some of which we can address with additional assumptions or subsets of data and some of which we cannot. First, there are significant measurement concerns about female labour force participation in historical and modern developing country data. As we explain in detail below, our results are robust to excluding historical data and to using developing country samples where female labour participation is externally validated by the International Labour Organization (ILO), the most reliable outside source. Second, the complier population varies from developed to developing countries and over the two-century span of our data, as does the base rate of women's labour force participation. We can address this heterogeneity, in part, by weighting our results to a constant complier covariate profile or scaling by the complier outcome mean. Our results are robust to both methods, although each comes with assumptions. Third, exact dates of birth and complete birth histories are available only for a subset of our data. We show that our results are similar in this subset of the data. Fourth, our main results are based primarily on labour force participation rather than the intensive margin of hours worked; we present results on hours below, although they are based on much more limited samples.

There are two important issues our data do not allow us to consider. First, by construction, the twins and same gender instruments cannot be applied to the birth of first children. Indeed, we are only aware of two research strategies that focus on the effects of first children. The first uses longitudinal data in event studies of first birth (e.g. Angelov *et al.*, 2016; Kleven *et al.*, 2019a). These studies find large negative labour supply effects in several developed countries, though this strategy has not, to our knowledge, been applied in a developing country. The second approach to first births relies on the random success of in vitro fertilization (IVF), which is not classified in any of our datasets. That said, the contrast between Aguero and Marks' (2008, 2011) IVF-based finding of a zero effect in developing countries and Cristia's (2008) and Lundborg *et al.*'s (2017) large negative effect in developed countries is tellingly consistent with the patterns in our data. Moreover, we show a similar pattern, albeit with a monotonically declining magnitude, across all family size parities beyond one child, at least suggestive that the negative gradient is a general result. Second, our data are cross-sectional and therefore only allow identification of the short-run effect of fertility. As noted in Adda *et al.* (2017) among others, the life-cycle response is often attenuated compared to the short-run effect, and late-in-life (rather than early) shocks are more likely to have lasting impacts on fertility.

The empirical regularities we describe are consistent with a standard labour-leisure model augmented to include a taste for children. As wages increase during the process of development, households face an increased time cost of fertility but also experience increased income. With a standard constant elasticity of substitution utility function, the former effect dominates as countries develop, creating a negative gradient (Appendix A provides a sketch of the model).

Indeed, in exploring the mechanism behind our result, we document that the substitution effect falls from zero to negative and is economically important as real GDP per capita increases. We argue that the declining substitution effect arises from changes in the sectoral and occupational structure of female jobs, as in Goldin (1995) and Schultz (1991). As economies evolve, women's labour market opportunities transition from agricultural and self-employment to urban wage work. The latter tends to be less compatible with raising children and causes some movement out of the labour force. In support of this channel, we show that the negative gradient is steeper among mothers with young children that work in non-professional and non-agricultural wage-earning occupations (e.g., urban wage work). Moreover, a growing literature documents a causal relationship between access to child care or early education and the propensity of mothers to work (e.g. Baker *et al.*, 2008; Havnes and Mogstad, 2011), a finding that is consistent with leaving the workforce when labour market opportunities become less compatible with child rearing. We cannot rule out that the income effect from rising wages could also be playing a role in the negative gradient but the evidence is at best mixed. Other explanations, most notably the widespread adoption of modern contraceptives and shifting social norms about female work (Goldin, 1977; Boustan and Collins, 2014) could also be compatible with our results. While we can find little evidence consistent with these alternative mechanisms, our data do not allow us to rule them out.

Our main empirical findings have important implications both for understanding the historical evolution of women's labour supply and the relationship between the demographic transition and the process of economic development. As Goldin (1995) documents in her comprehensive study of women's work in the 20th century, women's labour supply follows a U-shape over the process of economic growth, first declining before eventually increasing (see also Mammen and Paxson, 2000). Our results suggest that declining fertility may have contributed to the upswing in women's labour supply in much of the developed world during the second half of the century. Moreover, family policies (Olivetti and Petrongolo, 2017) and childcare costs (Del Boca, 2015) likely played a role. At the other end of the economic development spectrum, our results suggest that the demographic transition to smaller families probably does not have immediate implications for women's labour supply and growth. This in turn reinforces a claim in the demographic transition literature (Bloom *et al.*, 2001) that family planning policies are unlikely to enhance growth through a labour supply channel, although such policies could still be desirable for other reasons.

The paper is organized as follows. Section II explains the empirical strategy, followed in section III by a description of the data. Section IV presents our main findings. Section V analyses potential channels for our results. Section VI briefly discusses a series of robustness checks. Section VII concludes.

II. Empirical Strategy

Our empirical analysis adopts the standard approach of exploiting twin births and gender composition as sources of exogenous variation in the number of children to identify the causal effect of an additional child on the labour force activity of women (Rosenzweig and Wolpin, 1980; Bronars and Grogger, 1994; Angrist and Evans, 1998; Black *et al.*, 2005; Caceres-Delpiano, 2006; Vere, 2011). In particular, for twin births, we consider a first stage regression of the form:

(1) $$z_{ijt} = \gamma S_{ijt} + w_{ijt}'\rho + \pi_{jt} + \mu_{ijt}$$

where z_{ijt} is an indicator of whether mother *i* in country *j* at time *t* had a third child, the instrument S_{ijt} is an indicator for whether the second and third child are the same age (twins), w_{ijt} is a $k \times 1$ vector of demographic characteristics that typically include the current age of the mother, her age at first birth, and an indicator for the gender of the first child, and π_{jt} are country-year fixed effects. γ measures the empirical proportion of mothers with at least two children who would not have had a third child in the absence of a multiple second birth.
The local average treatment effect (LATE) among mothers with multiple children is identified from a second stage regression:

(2) $$y_{ijt} = \beta z_{ijt} + w_{ijt}'\alpha + \theta_{jt} + \varepsilon_{ijt}$$

where y_{ijt} is a measure of labour supply for mother *i* in country *j* at time *t* and β is the IV estimate of the pooled labour supply response to the birth of twins for women with at least one prior child.[2] Our baseline twin estimates condition on one child prior to the singleton or twin so that all mothers have at least two children, as in Angrist and Evans (1998). This restriction provides a family-size-consistent comparison so that both the same-gender and twins IV study the effect of a family growing from two to three children.

While twins are a widely-used source of variation for studying childbearing on mothers'

[2] We also aggregate the results in a procedure that is analogous to a hierarchical Bayesian model with a flat prior. To identify the gradient, we use a local polynomial smoother with a bandwidth of $1,500, where each country-year point estimate is weighted by its precision. That has no impact on our inferences.

labour supply, it is by no means the only strategy in the literature. Perhaps the leading alternative exploits preferences for mixed gender families (Angrist and Evans, 1998). Angrist and Evans estimate a first-stage regression like equation (1) but, for S_{ijt}, substitute twin births for an indicator of whether the first two children of woman i are of the same gender (boy-boy or girl-girl). Again, the sample is restricted to women with at least two children and γ measures the likelihood that a mother with two same gendered children is likely to have additional children relative to a mother with a boy and a girl.

Both twins and same gender children have been criticized as valid instruments on the grounds of omitted variables biases. Twin births may be more likely among healthier and wealthier mothers and can consequently vary over time and across geographic location (Rosenzweig and Wolpin, 2000; Hoekstra *et al.*, 2007; Bhalotra and Clarke, 2016; Clarke, 2018). Rosenzweig and Zhang (2009) also argue that twin siblings may be cheaper to raise, leading to a violation of the exclusion restriction. While the same gender instrument has proven quite robust for the U.S. and other developed countries (Butikofer, 2011), there are many reasons to be cautious in samples of developing countries (Schultz, 2008). Among other factors, households may practice either sex selection or selective neglect of children based on gender (Ebenstein, 2010; Jayachandran and Pande, 2017).

We adopt the broad view of Angrist *et al.* (2010) that the sources of variation used in various IV strategies are different and, therefore, so are the biases. As such, each IV provides a specification check of the other. Besides the basic LATE estimates underlying the multiple instrument methodology of Angrist *et al.* (2010), we also report a) a third instrument introduced by Klemp and Weisdorf (2019), which relies on exogenous variation in the timing of first births; b) twin results at alternative family parities; c) estimates that control for education and health measures to the greatest extent possible, including height and body mass index that have been highlighted as key determinants of twin births (Bhalotra and Clarke, 2016); and d) estimates by same gender versus mixed gender twins.[3] All these specification checks (see Appendix B for details) are consistent with a declining labour supply gradient over development when they can be implemented across the GDP distribution.

[3] Monozygotic (MZ) twinning is believed to be less susceptible to environmental factors. Hoekstra *et al*. (2007) provides an excellent survey of the medical literature. Since we cannot identify MZ versus dizygotic (DZ) twins in our data, we take advantage of the fact that MZ twins are always the same gender, whereas DZ twins share genes like other non-twin siblings and therefore are 50 percent likely to be the same gender.

The literature analyses a number of measures of y_{ijt}, including whether the mother worked, the number of hours worked, and the labour income earned. These measures are sometimes defined over the previous year or at the time of the survey. In order to include as wide a variety of consistent data across time and countries as possible, we typically focus on the labour force participation (LFP) of mothers at the time of a census or survey. When LFP is unavailable, especially in pre-WWII censuses, we derive LFP based on whether the woman has a stated occupation. Appendix B discusses the robustness of the results to several alternative labour market measures, including mismeasurement of occupation-based LFP (Goldin, 1990).

In concordance with much of the literature (especially Angrist and Evans, 1998), our standard sample contains women aged 21 to 35 with at least two children, all of whom are 17 or younger. We exclude families where a child's age or gender or mother's age is imputed. We also drop mothers who gave birth before age 15, who live in group quarters, or whose first child is a multiple birth. It is worth emphasizing that the restrictions on mother's (21-35) and child's (under 18) age may allay concerns about miscounting children that have moved out of the household.[4] We also experiment with even younger mother and child age cut-offs, which additionally provide some inference about difference in the labour supply response to younger and older offspring. Further sample statistics, single sample estimates, as well as results when these restrictions are relaxed, are provided in the Appendix tables.

We present our results stratified by time, country, level of development, or some combination. The prototypical plot stratifies countries-years into seven real GDP per capita bins (in 1990 U.S. dollars): under \$2,500, \$2,500-5,000, \$5,000-7,500, \$7,500-10,000, \$10,000-15,000, \$15,000-20,000, and over \$20,000. To be concrete, in this example, all country-years where real GDP per capita are, say, under \$2,500 in 1990 U.S. dollars are pooled together for the purpose of estimating equations (1) and (2). Similarly, countries with real GDP per capita between \$2,500 and \$5,000 are also pooled together for estimation, and so on. The plots report weighted estimates of γ and β, and their associated 95 percent confidence interval based on country-year clustered standard errors, for each bin.[5]

[4] As a robustness check, we also use information about complete fertility when it is available.
[5] Household weights are supplied by the various surveys or censuses, normalized by the number of mothers in the final regression sample.

III. Data

We estimate the statistical model using four large databases of country censuses and surveys.

a. U.S. Census, 1860-2010

The U.S. is the only country for which historical microdata over a long stretch of time is *regularly* available. We use the 1 percent samples from the 1860, 1870, 1950, and 1970 censuses; the 5 percent samples from the 1960, 1980, 1990, and 2000 censuses; the 2010 American Community Survey (ACS) 5-year sample, which combines the 1 percent ACS samples for 2008 to 2012; and the 100 percent population counts from the 1880, 1900, 1910, 1920, 1930, and 1940 censuses.[6] Besides additional precision, the full count censuses allow us to stratify the sample (e.g. by states) to potentially take advantage of more detailed cross-sectional variation.

IPUMS harmonizes the U.S. census samples to provide comparable definitions of variables over time. However, there are unavoidable changes to some of our key measures. For example, the 1940 census is the first to introduce years of completed schooling and earnings; therefore, when we show results invoking education or earnings, we exclude U.S. data prior to 1940. Perhaps most important, the 1940 census shifted our labour supply measure from an indicator of reporting any "gainful occupation" to the modern labour force definition of working or looking for work in a specific reference week. Fortunately, there does not appear to be a measurable difference in our results between these definitions in 1940 when both measures are available. Nevertheless, there is concern that women's occupations (Goldin, 1990) as well as fertility (Moehling, 2002) could be systemically under- or over-reported, especially in U.S. census samples for 1910 and earlier. We present a number of robustness checks meant to isolate these mismeasurement issues in Appendix B, and in Section IV.e present results that exclude historical data.

For Puerto Rico, we use the 5 percent census samples from 1980, 1990, and 2000 and the 2010 Community Survey, which combines the 1 percent samples for 2008 to 2012. Censuses prior to 1980 are missing labour force data or reliable information about real GDP per capita.

---

[6] For information on the IPUMS samples, see Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek, ***Integrated Public Use Microdata Series: Version 5.0*** [Machine-readable database], Minneapolis: University of Minnesota, 2010. The 100 percent counts were generously provided to us by the University of Minnesota Population Center via the data collection efforts of ancestry.com. Those files have been cleaned and harmonized by IPUMS. The 1890 U.S. census is unavailable and U.S. censuses prior to 1860 do not contain labour force information for women. In some figures, we also report single-year estimates from the 1880 10 percent, 1900 and 1930 5 percent, as well as the 1910, 1920, and 1940 1 percent random IPUMS samples.

b. IPUMS International Censuses, 1960-2015

IPUMS harmonizes censuses from around the world, yielding measures of our key variables that are roughly comparable across countries and time. We use data from 212 of the 301 non-U.S. country-year censuses between 1960 and 2015 that were posted at the IPUMS-I website as of May 2017. Censuses are excluded if mother-child links or labour force status is unavailable (83 censuses) or age is defined by ranges rather than single-years (6 censuses).[7]

c. North Atlantic Population Project (NAPP), 1787-1911

The North Atlantic Population Project (NAPP) provides 18 censuses from Canada, Denmark, Germany, Great Britain, Norway, and Sweden between 1787 and 1911. As with IPUMS, these data are made available by the Minnesota Population Center.[8] For most samples, NAPP generates family interrelationship linkages. However, in a few cases (Canada for 1871 and 1881 and Germany in 1819) such linkages are not available. In those cases, we use similar rules developed to link mothers and children in the U.S. full count census. Also, consistent with the pre-1940 U.S. censuses, labour force activity is based on whether women report an occupation rather than the modern definition of working or seeking work within a specific reference period, and education is unavailable.

d. Demographic and Health Surveys (DHS), 1990-2014

We supplement the censuses with the Demographic and Health Surveys (DHS).[9] From the initial set of 254 country-year surveys, spanning 6 waves from the mid-1980s onward, we exclude samples missing age of mother, marital status of mother, current work status, whether the mother works for cash, birth history, and comparable real GDP per capita. These restrictions force us to drop the first wave of the DHS, leaving 692,923 mothers in 192 country-years.

The DHS includes a number of questions that are especially valuable for testing the

---

[7] Similar to the U.S., the international linking variables use relationships, age, marital status, fertility, and proximity in the household to create mother-child links. Sobek and Kennedy (2009) compute that these linking variables have a 98 percent match rate with direct reports of family relationships. However, we are not able to compute linkages that do not include relevant household information on relationship and surname similarity. Unfortunately, this affects some censuses from Canada and the U.K. Although the 1971 to 2006 Irish censuses use age ranges for adults, they do not for children under 20 (so we literally include Irish twins!).

[8] See Minnesota Population Center (2015), North Atlantic Population Project: Complete Count Microdata, Version 2.2 [Machine-readable database], Minneapolis: Minnesota Population Center.

[9] For additional information about the DHS files see ICF International (2015). The data is based on extracts from DHS Individual Recode files. See http://dhsprogram.com/Data/.

robustness of our census results. First, detailed health information allows us to control for characteristics that may be related to a mother's likelihood of twinning (Bhalotra and Clarke 2016). Second, we can use an indicator of whether children are in fact twins to test the accuracy of our coding of census twins.[10] To keep the DHS results comparable to the censuses, our baseline DHS estimates identify twins based on the census year-of-birth criterion and consider only living children who reside with the mother.

e. Real GDP per Capita

Real GDP per capita (in US$1990) is collected from the Maddison Project.[11] To reduce measurement error, we smooth each GDP series by a seven year moving average centred on the survey year. We are able to match 441 country-years to the Maddison data, leaving 51,449,770 mothers aged 21 to 35 with at least two children in our baseline sample.[12]

When we split the 1930 and 1940 full population U.S. censuses into the 48 states and DC, we bin those samples by state-specific 1929 or 1940 income-per-capita.[13] The income data are converted into 1990 dollars using the Consumer Price Index.

f. Summary Statistics

Table 1 provides summary statistics separately for the U.S. and non-U.S. samples and by real GDP per capita bins. Although the first bin (less than $2,500 GDP per capita) is dominated by DHS samples, most bins have a large number of mothers for both U.S. and non-U.S. samples. Appendix Table A1 provides additional descriptive statistics and estimates by individual country-year datasets.

---

[10] Appendix Figure A1 illustrates the high degree of correspondence between twinning rates when we define twins using "real" multiple births and those imputed for children sharing the same birth-year. The DHS has a number of labour force variables but none that directly compare to those in the censuses. We chose to use an indicator of whether the mother is currently working since it is most correlated with the IPUMS labour force measures (see Appendix Figure A2).

[11] See http://www.ggdc.net/maddison/maddison-project/home.htm.

[12] In a few minor cases, we were not able to match a country to a specific year but still left the census in our sample because we did not believe it would have impacted their placement in a real GDP per capita bin. Specifically, the censuses of Denmark in 1787 and 1801 are matched to real GDP per capita data for Denmark in 1820 and Norway in 1801 is matched to data for Norway in 1820. Excluding these country-years has no impact on our results. More importantly, the Maddison data ends in 2010 and therefore censuses or surveys thereafter are assigned their most recently available real GDP per capita data.

[13] http://www2.census.gov/library/publications/1975/compendia/hist_stats_colonial-1970/hist_stats_colonial-1970p1-chF.pdf.

IV. Results

a. OLS Estimates

We begin with estimates from OLS regressions of the labour supply indicator on the indicator for a third child and the controls described above. These results do not have a clear causal interpretation, but they are useful for establishing key data patterns. In Figure 1, we plot the coefficients for the U.S., the non-U.S. countries, and the combined world sample (labelled "All"), binned into the seven ranges of real GDP per capita reported on the x-axis ($0-2,500, $2,500-5,000, etc.). Point estimates and country-year clustered standard errors are provided in Table 2. The three samples exhibit a similar pattern. At low levels of real GDP per capita, the OLS estimate of the effect of children on mother's labour supply is negative and statistically significant at the 5 percent level but economically small in magnitude (e.g. -0.022 (0.005) in the lowest GDP bin). As real GDP per capita increases, the effect becomes more negative, ultimately flattening out between -0.15 and -0.25 beyond real GDP per capita of $15,000.

Figure 2 plots the U.S.-only OLS results over time. [14] Blue circles represent IPUMS samples and red diamonds represent full population counts. These estimates start out negative, albeit relatively small (e.g. -0.011 (0.004) in 1860 and -0.008 (0.0004) in 1910), decrease from 1910 to 1980, at which point the magnitude is -0.177 (0.001), and flatten thereafter.

Appendix Figure A3 plots the OLS estimates by real GDP per capita separately by time periods (pre-1900, 1900-1949, 1950-1989, and 1990+). Years prior to 1950 combine U.S. census and NAPP data. Years thereafter include all four of our databases. The same general pattern appears *within* time periods.[15] The effect of fertility on labour supply tends to be small at low levels of GDP per capita but increases as GDP per capita rises.

b. Twins IV

The left panel of Figure 3 shows the first-stage effect, γ in equation (1), of a twin birth on our fertility measure, the probability of having three or more children. For the U.S., non-U.S., and combined world samples, there is a positive and concave pattern, with the first-stage increasing with higher real GDP per capita up to $15,000 or so and flattening thereafter. Note

---

[15] Relative to Figure 1, we combined some real GDP per capita bins because of small sample sizes within these tight time windows.

that the regression specification controls for the mother's age, but does not, indeed cannot, control for the number of children or target fertility. Therefore, the positive gradient over real GDP per capita reflects the negative impact of income on target fertility and hence the heightened impact of a twin birth on continued fertility relative to a non-twin birth.[16] In all cases, the instrument easily passes all standard statistical thresholds of first-stage relevance, including among countries with low real GDP per capita and high fertility rates. [17]

The right panel of Figure 3 (and Table 2) plots β, the instrumental variables effect of fertility on mother's labour supply. In the world sample, β is mostly statistically indistinguishable from zero among countries with real GDP per capita of $7,500 or less. Subsequently, β begins to decline and eventually flattens out between -0.05 and -0.10 at real GDP per capita of around $15,000 and higher. The results for the U.S. and non-U.S. samples are similar in that there is a notable negative gradient with respect to real GDP per capita. For example, above $20,000, the U.S. estimate is -0.070 (0.008)[18] while the non-U.S. estimate is -0.105 (0.003). The U.S. (non-U.S.) estimate implies that an extra child is associated with a decrease in a mother's labour supply of around 11 (14) percent, relative to an average base rate of 62.9 (73.6) percentage points (e.g. -0.070/0.629=-0.111).

In Figure 4, we show the results by time window. This gives us a sense of how much of the pattern we observe is due to differences in development instead of secular changes across time. The central message of this figure is that the results are very consistent across time periods at similar levels of GDP per capita.[19] We think it is particularly notable that the declining β

[16] The first stage coefficient, γ, is E{z=1|S=1,w} – E{z=1|S=0,w}. Mechanically, E{z=1|S=1,w}=1 because of the definition of twins. This means that if, for example, γ=0.6, then E{z=1|S=0,w}=0.4, implying that 40 percent of mothers would have a third child if their second child is a singleton. The increasing coefficient over real GDP per capita means having a third child after a singleton second child is declining with development. The reversal of this pattern at real GDP per capita of $10-15,000 in the U.S. represents the Baby Boom.

[17] The smallest first stage F-statistic for the results displayed in Figure 3 is 170 for the non-U.S.-only results for countries between $2,500-5,000 GDP per capita.

[18] By comparison, Angrist and Evans (1998) report a twins IV estimate of -0.079 for the 1980 U.S. census. Vere (2011) estimates twins IV coefficients for a third child of -0.086, -0.095, and -0.078 for 1980, 1990, and 2000, respectively.

[19] In Appendix Figures A4 and A7, we present U.S. twin and same gender results by census decade. The pattern is broadly similar to the previous figure. The magnitude of the first stage is increasing over time, and the second-stage IV results begin to exhibit a pronounced negative gradient, particularly post-WWII. The same pattern arises within datasets (Appendix Figure A5 and Figure A8) and within geographic regions of the world (Appendix Figure A6 and Figure A9, although again with much noisier estimates for the same gender instrument).

appears prior to the wide-spread availability of modern fertility treatments like IVF in wealthy countries and after modern census questions on labour force participation and fertility were introduced in 1940. We further address these potential issues below.

c. Are There Positive Labour Supply Effects Among the Lowest Income Countries?

One surprising finding is that at low real GDP per capita levels, we sometimes estimate a positive labour supply response to childbearing. This result is particularly evident in the pre-WWI U.S. (displayed in Figure A4), but also periodically appears, although not always statistically significantly so, for some low-income, post-1990 countries. The positive U.S. results are not statistically different from zero for the early census samples (1860, 1870) but are for the full population counts of 1880 and 1910.

While these positive results are not artefacts in the statistical sense, it is worth noting that the underlying rates of labour force participation for U.S. women are very low at this time in history (e.g. 6.2 and 10.0 percent for 1880 and 1910 mothers, respectively). As such, a positive effect could reflect that low-income mothers are more likely to work after having children, for example because subsistence food and shelter are necessary, whereas childcare might be cheaply available.

To gain further insight into the low real GDP sample results, we split the U.S. 1930 and 1940 full population counts by state of residence and pool states into income-per-capita estimation bins (matching what we did with countries in previous figures). Figure 5 shows the now familiar upward sloping pattern to the first stage results by real income per capita. In the second stage, we see that the effect of fertility on labour supply is in general statistically indistinguishable from zero at low-income levels in 1930 and 1940 and overlaps with the low-income post-1990 non-U.S. results (shown in the green line). But we also find a small positive effect from the lowest income states in 1930, seemingly corroborating the positive estimates from a lower income U.S. prior to WWI.[20] These findings are directionally consistent with

[20] For the 1930 census, the states in that lowest bin ($2,000-3,000) are: Alabama, Arkansas, Georgia, Mississippi, North Carolina, North Dakota, New Mexico, South Carolina, and Tennessee.

Godefroy (2017) and Heath (2017).

d. Same Gender IV

Next, we discuss results, displayed in Figure 6 and Table 2, which use the same gender instrument. Like the twins IV, we estimate a positive gradient to the first stage with respect to real GDP per capita, although the interpretation of this pattern is different than for twins. In particular, the same-gender first-stage picks up the increased probability that a mother opts to have more than two children based on the gender mix of her children (rather than picking up the proportion of mothers with incremental fertility when the twin instrument is zero, i.e., for non-twin births).[21] Most importantly, we again see a negative gradient on the second stage IV estimates, from a close-to-zero effect among low GDP countries to a negative and statistically significant effect at higher real GDP per capita that flattens at around $15,000. As with the twins estimates, the negative estimates appear in the U.S. post-WWII (Appendix Figure A7).[22]

Our main intention is to highlight the similar shapes of the labour supply effect across the development cycle, despite using instruments that exploit different sources of variation. Indeed, when we combine all possible instrument variation into a singled pooled estimator, as in Angrist *et al.* (2010), our weighted average twin and same gender IV results also, unsurprisingly, shows the same strong negative gradient. That said, the magnitude of the same gender IV result is larger than the twin IV result at the high GDP per capita bins. For example, at the $20,000 and above bin, the twin estimate is -0.070 (0.008) for the U.S. sample and -0.105 (0.003) for the non-U.S. sample. By comparison, the same gender estimates are -0.121 (0.008) for the U.S. sample and -0.173 (0.019) for the non-U.S. sample. Since this is a local average treatment effect, this disparity suggests a greater effect of fertility on labour supply for those women encouraged to have an incremental child based either on son preference or the taste for a gender mix compared to those induced to higher fertility by a twin birth.

e. Measurement Concerns with Female Labour Force Participation

There are significant concerns with how female labour participation is measured in pre-1940 U.S. censuses and modern developing country surveys and censuses, especially relative to

[21] We find that the first stage of the same gender instrument is overall weaker than the twin instrument but passes the usual tests of relevance in binned samples, such as those in Figure 6. The one case with a weak first stage is the U.S. estimates with GDP less than $2,500, which is based on the 1860 and 1870 U.S. censuses. See Bisbee *et al*. (2017) for more details.

[22] Like the twins estimates, we also find systematic evidence of a positive fertility-labour supply effect at low levels of income, which are statistically significant for the 1900, 1930, and 1940 U.S. censuses (see Appendix Figure A7).

measurement in modern developed country censuses. With regard to the historical U.S., pre-1940 censuses use an occupation-based measure of labour force participation and introduce a number of miscodings highlighted in Goldin (1990). For developing countries, women's work may not be as clearly defined in informal settings, home production, and agriculture.

To address these concerns, Figure 7 compares our baseline estimates to results that exclude two sets of potentially mismeasured data. First, we throw out pre-1940 censuses and non-U.S. pre-1950 data.[23] Second, we exclude IPUMS and DHS samples where our measure of female labour force participation fails to adequately match female LFP that was independently validated by the International Labour Organization (ILO) (see https://www.ilo.org/ilostat). We identify 177 country-years where the ILO estimate of female LFP for 25 to 34 year-olds is within 4.8 percentage points (the median difference) of a comparable IPUMS or DHS estimate.[24]

The key patterns are the same as our baseline results: the negative effect of fertility on female labour force participation starts out small and becomes more negative over the process of development for both twin instrument (panel A) and the same gender instrument (panel B). One difference is that the two lowest GDP bins for the twins instrument (and the first and third bins for the same gender instrument) have statistically significant negative effects. However, the magnitude of the negative effect in the lowest GDP bins is small both relative to female labour force participation (56.9 percent) in these samples and to the point estimates at higher levels of GDP. The results are similar when we retain the best third or best two-thirds of ILO matches. In the former case, the ILO LFP rates are nearly identical to the IPUMS/DHS LFP rates.

Finally, in principle, we would like to analyse the effect of fertility on hours worked and participation separately. However, this would require instruments for the intensive and extensive margins. Nonetheless, as an exploratory analysis, Figure A10 plots twin and same gender instrumental variables results for the number of hours worked per week where those out of the labour force are coded as working zero hours. We include all country-years that contain a measure of hours worked, which unfortunately limits us to only 56 censuses -- eight from the

[23] In Appendix B, we also partially recode the miscoded occupations following Goldin (1990). Those results are also similar to the baseline estimates.

[24] Since our surveys do not always align with ILO's periodicity, when necessary we extrapolate or linearly interpolate between ILO estimates (up to a maximum of four years) to obtain an estimate of female labour force participation in the IPUMS and DHS years. In the end, we are able to match 355 of our 441 country-years to the ILO, with all samples based on 1950 or later – that is, historical data is excluded. The 177 country-years that we use in this exercise represent the best half of the country-year matches.

U.S. (1940-2010) and 48 from the International IPUMS (The DHS and NAPP do not contain hours worked per week). We continue to find a negative gradient to labour supply, with the difference between hours worked among mothers in low-income and high-income countries being about 1.3 for the twins instrument and 4.3 hours for same gender instrument. As a benchmark, all mothers work, on average, just under 23 hours per week in countries with real GDP per capita above $20,000, suggesting a roughly 4 to 18 percent average decline in hours as a result of an additional child, conditional on working.

V. Channels

This section explores some of the potential mechanisms that account for the remarkably robust negative income gradient of mother's labour supply response to children.

a. Accounting for a Changing Complier Participation

A key challenge in interpreting our results is that the complier population is likely to change across our data. The group of women induced to have more than two children because of an initial twin birth in a context where most women have more than three children (e.g., in a developing country or in historical data from developed countries) is presumably different from the women encouraged to have more than two children by a twin birth in a low fertility context. It is important to acknowledge that we cannot directly address this issue, at least without a stronger set of assumptions.

An indirect approach to capture variation from different sets of mothers is to condition on different family size parities (Angrist *et al.*, 2010). For example, one might expect that mothers with a large number of previous children would be less likely to adjust their labour supply in response to unexpected incremental fertility (for example, because of low incremental childcare costs for higher births). Indeed, as shown in Figure 8, we observe a stronger first stage effect for the sample that conditions on more children, especially at higher income levels. In the second stage, we see a notably, although not always statistically significantly, more negative effect in high-income countries for women starting with one child. The pattern of results is similar regardless of how many children are in the household when the twins are born. In all non-zero family size circumstances (up to three initial children), we continue to find no effect among low-income countries and an increasingly larger negative effect among higher income countries, flattening out around \$20,000 per capita.[25] The continued robustness of the negative gradient to family parity suggests that the key patterns in our results may not solely be driven by changes in the complier population, although as noted above this is at best indirect evidence.[26]

More direct evidence requires stronger assumptions. Using the approach suggested by Angrist and Fernandez-Val (2013) (see Bisbee *et al.*, 2017 for a related application), we can adjust for changes in the complier population by reweighting our IV estimate to a constant complier profile. This adjustment assumes a constant treatment effect conditional on a covariate profile, in other words that heterogeneity in IV causal effects is driven by observable changes in complier characteristics. We use Abadie's (2003) kappa function to recover the covariate profile of compliers in a target year. We then compute covariate-specific IV treatments in other years, and reweight these to match the twin IV complier covariate profile in the 1980 U.S. Specifically, given a $k \times 1$ vector of covariates that have been de-meaned by the means of the target complier population, $\widetilde{w}_{ijt}$, we augment the standard 2SLS framework by estimating the following second-stage equation and reporting estimates of β:[27]

$$y_{ijt} = \beta z_{ijt} + z_{ijt}\widetilde{w}_{ijt}{}'\delta + \widetilde{w}_{ijt}{}'\alpha + \theta_{jt} + \varepsilon_{ijt} \tag{3}$$

This procedure involves estimating $k + 1$ corresponding first-stage equations for $\{z_{ijt}, z_{ijt}\widetilde{w}_{ijt}{}'\}$ using $\{S_{ijt}, S_{ijt}\widetilde{w}_{ijt}{}'\}$ as instruments. Reweighting by age and education bins significantly impact the first stage at low levels of GDP per capita, but there are no significant changes in the IV estimates (see Figure 9).

Together, these results lead us to postulate – albeit with significant qualifications on the

[25] Additionally, we restrict the DHS sample to mothers whose report their ideal number of children as less than three (or four) and obtain nearly identical point estimates. This test loosely addresses concern that the parities we consider would not be binding and, consequently, have no labour supply effect in high-fertility, low-income countries.

[26] Unfortunately, by construction, the twin and same gender instruments are unable to identify the labour supply effect from an unexpected first child. Causal evidence on the impact of first births sometimes uses childless mothers undergoing in vitro fertilization (IVF) treatments. Interestingly, Cristia (2008) and *Lundborg et al*. (2017) find large negative labour supply responses to successful IVF treatment in the U.S. and Denmark, respectively. By contrast, Aguero and Marks (2008, 2011) find no impact among 32 developing countries. While, we cannot replicate these findings with our data, the patterns seem to further validate a negative labour supply gradient across all family parities. See also Angelov *et al*. (2016), *Kuziemko et al*. (2018), Kleven *et al*. (2019a and 2019b) for an event study approach in developed countries. For comparable results across family size parities, see Bronars and Grogger (1994), Angrist and Evans (1998), Cruces and Galiani (2007), Maurin and Moschion (2009), Vere (2011), and Lundborg *et al*. (2017).

[27] See propositions 1 and 2 of Bisbee *et al*. (2017). To be as flexible as possible, we discretize our baseline covariates into dummy variables of mother's age (3-year bins), age at first birth (3-year bins), first child gender, and education (<8, 8-11, 12-15, 16+ years of schooling where applicable). Note that we include country-year fixed effects as usual.

available evidence – that the key patterns in our results are not driven by changes in the complier population over the process of development.

b. Accounting for Changing Base Rates of Labour Force Participation

A related possibility is that the negative gradient is driven by the changes in the base rate of labour force participation. A lower base rate of labour force participation would imply less scope for a negative fertility effect on labour supply. This mechanically limits the scale of any average causal effect of fertility. We can account for this possibility by rescaling estimates to the relevant base rate (as in Angrist *et al.*, 2013). The rescaling relies on the assumption that effects tend to be monotonic in the population under study. That is, write the average effect in population s as

(3) $$\beta_s = E_s[Y_1 - Y_0],$$

where Y_1 and Y_0 are potential labour outcomes (with support $\{0,1\}$) under the condition of three or more children and less than three children, respectively. Effect monotonicity implies $Y_1 \leq Y_0$, which also means

(4) $$E_s[Y_1 - Y_0 | Y_0 = 0] = 0.$$

This further implies that

(5) $$\beta_s = E_s[Y_1 - Y_0 | Y_0 = 1]\mathrm{E_s}[Y_0],$$

in which case the average effect of having three or more children *among those for which there can be an effect* is given by

(6) $$\beta_s^r = E_s[Y_1 - Y_0 | Y_0 = 1] = \frac{\beta_s}{\mathrm{E_s}[Y_0]}.$$

Comparing trends in β_s versus β_s^r allows us to assess the influence of base participation rates.[28]

Given that we are estimating complier LATEs via IV, the populations indexed by s correspond to the compliers in our various country years. As such, the relevant base rate, $E_s[Y_0]$, corresponds to the labour force participation rate among compliers with instrument values equal

[28] This rescaling recovers a meaningful effect in populations for which the monotonicity assumption is reasonable. Rescaling would not be valid in country-years, such as those described in Section IV.c, where we estimate statistically significant positive fertility effects. Our figures are based on samples that include positive estimates, except for the pre-1920 U.S. which shows the most consistently positive results. If we apply our rescaling strategy to country-year samples for which we observe either negative or (statistically indistinguishable from) zero fertility effects, we still recover a comparable negative gradient, although, unsurprisingly, labour supply responses at all real GDP per capita levels become more negative.

to 0. We compute these complier-specific rates using the IV approach of Angrist *et al.*, (2013).[29]

Figure 10 shows the rescaled baseline twins estimates (rescaled estimates for the ILO-restricted sample are shown in Figure A11). For the U.S., the rescaling results in a substantial flattening past $7,500 per capita. For the non-U.S. populations, the rescaled estimates are consistent (taking into account the uncertainty in the estimates) with a flattening after $10,000 per capita. However, a negative gradient is still evident over lower levels of income. This indicates that the decline in the labour supply effect of an additional child is not solely driven by increases in the base rate of mother's LFP and motivates further analysis into the channel driving the negative gradient, particular over income levels under $10,000 per capita. The analyses below examine results both with and without the base-rate rescaling.

It is worth noting that this procedure does not adjust for changing selectivity into the complier population, which the literature (e.g., Olivetti and Petrongolo, 2008) suggests is likely to be occurring.

c. Changes to the Income and Substitution Effect Across Stages of Development

We believe much of the remaining negative gradient is due to a declining substitution effect, in combination with a mostly unchanging income effect, resulting from increasing wages for women during the process of economic development.

We identify the substitution effect primarily through changes in job opportunities. This exercise is motivated by previous work that documents a U-shape of female employment with development in the U.S. and across countries (Goldin, 1995; Schultz, 1991; Mammen and Paxson, 2000). Schultz (1991) shows that the U-shape is not observed within sector. Rather, it is explained by changes in the sectoral composition of the female labour force. Specifically, women are less likely to participate in unpaid family work (mostly in agriculture) and self-employment and more likely to be paid a wage in the formal sector in the later stages of the development process. In addition, we have reason to believe that the types of jobs that women have over time might change in a way that is less suitable to raising children. For example, in rural, agricultural societies, women can work on family farms while simultaneously taking care of children, but the transition to formal urban wage employment is less compatible with

[29] Specifically, we stack the two-stage estimation used in Angrist *et al.* (2013) to calculate the complier-control mean with our baseline two-stage least squares regression to get the covariance between the base rate and the labour supply effect.

providing care at home (Jaffe and Azumi, 1960; McCabe and Rosenzweig, 1976; Kupinsky, 1977; Goldin, 1995; Galor and Weil, 1996; Edwards and Field-Hendrey, 2002; Szulga, 2014).

Given that consistent information on occupations and sectors across our many samples is limited, we rely on two coarse indicators of job type that can be consistently measured in almost all of our data. First, we try to capture the distinction between urban/rural and formal/informal occupations by changing the outcome to be whether women work for a wage or work but are unpaid. These results, unscaled (left) and scaled (right), are presented in Figure 11 (results for the ILO-restricted sample are presented in Figure A12). The unscaled results show that the changing relationship between fertility and labour supply is driven by women who work for wages. The scaled results are consistent with this finding, in that the effect is greater for wage workers than non-wage workers, so that the gradient is driven by changes in the sectoral composition of the labour force toward wage workers.

A second proxy of sectoral shifts is whether women work in agricultural or non-agricultural sectors (Figure 12 for the main sample, and Figure A13 for the ILO-restricted sample). Although the scaled results presented in the right plot are unfortunately noisy for agricultural labour, the labour supply response of women in non-agricultural sectors becomes clearly more negative as real GDP per capita rises. We also observe in Figure 13 (Figure A14 for the ILO-restricted sample) that fertility has almost no differential effect across the development cycle on female labour supply in professional occupations, despite the fact that these occupations tend to have higher wages.[30] Instead, the changing gradient seems to be driven entirely by women who work in non-professional occupations, suggesting either that education and professional status are poor proxies for the substitution effect or that the opportunity differences they capture are small in comparison to the sectoral shifts out of agricultural and non-wage work. This is consistent with an implication of the model laid out in Appendix A, which predicts that the negative gradient will be sharper among lower-skilled women.[31]

---

[30] Professional occupations are defined somewhat differently across data sources. For the U.S., we define professionals as Professional, Technical, or Managers/Officials/Proprietors. This definition corresponds to 1950 occupation codes 0-99 and 200-290. In all non-U.S. sources, we define professionals as close as possible to the U.S. For IPUMS-I, we use the International Standard Classification of Occupations (ISCO) occupation codes. For the NAPP, we use the Historical ISCO codes, except for 1911 Canada where we use 1950 U.S. occupation codes. We dropped the 1851 and 1881 U.K censuses due to difficulty convincingly identifying professionals.

[31] The fertility response literature has long used a woman's education to proxy for the type of jobs and wages available to her. While Gronau (1986) documents several results finding education is correlated with a fertility response, this correlation appears to reverse once Angrist and Evans (1998) apply instrumental variables. While we

By contrast, we believe that the income effect of rising wages on fertility is likely small and invariant to the stage of development, as in Jones and Tertilt (2008), although the evidence is admittedly somewhat mixed. We further investigate the relevance of income effects in two ways. First, we examine the husband's labour supply response to children using the same twin IV estimator. A long literature, tracing back to classic models of fertility such as Becker (1960) and Willis (1973), uses the husband's labour supply response as a proxy of the income effect, since the substitution effect is likely to be smaller for men, who typically spend less time rearing children than women. In Figure 14, we return to the unscaled estimates and show that the husband's labour supply response is economically indistinguishable from zero and invariant to the level of real GDP per capita. Second, we examine the 1940 to 2010 U.S. censuses, which contain hourly wages of husbands, to measure the differential labour supply response of married women throughout the hourly wage distribution of their spouse. Although we continue to see a negative gradient over time, there are some, not always statistically significant, cross-sectional differences among women based on their husband's wage (see Figure A15). In particular, the negative gradient appears to be more pronounced among women whose spouses are high wage. Thus, we cannot rule out that the income effect could be playing a (smaller) role in the negative gradient as well.

d. Child Care Costs

A key factor driving the relationship between mother's labour supply and children is the time cost of raising kids (e.g. see equation A.7 in Appendix A). One simple indication that child care costs could be a relevant channel is visible in Figure 15, which stratifies the samples by six year age bins of the oldest child (similar results by the age of youngest child are presented in Figure A16). Regardless of kids' ages, we find a negative gradient, with the labour supply elasticity declining at real GDP per capita around \$7,000 to \$15,000. However, the gradient is monotonically sharper for families with younger children who typically require more care, and especially among mothers in non-professional occupations with younger children (Table 3).[32] In

are able to replicate their results, we find that this education gradient is sensitive to instrument and the sample used. Overall, we find no strong heterogeneity by education (Appendix Figure A22).

[32] There is a monotonic relationship between age of children and time spent on child care. For example, in the U.S. Time Use Survey, 21-35 year old women with two children at home where one was under 6 spent 2.9 hours per day, on average, on child care (plus an additional 2.5 hours per day on other household activities). By comparison, when the youngest child is 6 to 11 or 12 to 17, mothers spend 1.8 and 1.3 hours per day, respectively, on child care. For the subset of mothers who are not working, child care takes up 6.8 (youngest child under 6), 5.4 (6 to 11), and 4.7 (12 to 17) hours per day.

particular, among mothers with a child under 6, the impact of a child on working in a non-professional occupation falls by -0.067 (0.010) in countries with real GDP per capita above \$10,000 relative to countries below \$10,000.[33] By comparison, the non-professional gradient falls to -0.053 (0.011) and -0.020 (0.021) for mothers with a youngest child between 6 to 11 and 12 to 17. Strikingly, the labour supply gradient among professional occupations is invariant to the age of the youngest child. These results are at least suggestive that non-professional mothers, who are most exposed to sectoral shifts over the development cycle, may also be least likely to be able to pay for childcare costs through formal wage work.

Ideally, we would test the importance of child care costs using exogenous variation across countries or over time. Unfortunately, we are not aware of such variation that spans our data. There is, however, a growing literature that uses quasi-experimental variation in access to child care or early education to study mother's labour supply in individual countries, including the U.S. (Cascio, 2009; Fitzpatrick, 2012; Herbst, 2017), Argentina (Berlinski and Galiani, 2007), Canada (Baker *et al.*, 2008), and Norway (Havnes and Mogstad, 2011).[34] Summarizing this literature, Morrisey (2017) concludes that the availability of child care and early education generally increases the labour supply of mothers, although there is some response heterogeneity across countries. We view this literature as at least consistent with the possibility that the negative labour supply gradient may be amplified if child care costs increase because jobs become less conducive to child rearing, and, if so, this dynamic could be stronger among lower wage mothers with less flexibility to provide child care to young children (Blau and Winkler, 2019).

e. Other Explanations and Robustness

The evidence from the U.S. shows that mothers' labour supply response to children likely fell in the decades immediately after WWII.[35] This is a period in which at least two important developments may have impacted female labour force participation: the introduction and wide-

[33] For exposition and due to sample size concerns that arise when dividing samples too finely, country-years in Table 3 are sorted into two real GDP per capita bins: above and below \$10,000. The bottom row, labeled "gradient," is the difference.

[34] To take one example, Herbst (2017) is based on the WWII-era U.S. Lanham Act that provided childcare services to working mothers with children under 12. State variation of funding offered a natural experiment in a period when we find the aggregate labour supply response of mothers to additional children was close to 0. Herbst reports that additional Lanham Act child care funding raised mother's labour force participation.

[35] The evidence from other countries for which we have data spanning the development cycle (Canada, France, Ireland, the United Kingdom) suggests a similar pattern (see Table A1).

spread usage of modern contraceptives and shifts in the social norms of female work.

To explore the importance of birth control pills, we exploit differences in the timing in which U.S. states allow access to the pill among 18 to 21 year-olds (Bailey *et al.*, 2012). Using mothers in the 1970 and 1980 censuses and a difference-in-difference design, we could not find evidence that access to birth control impacted the labour supply decisions of mothers with either of our main instruments. Combined with a robust cross-sectional negative mother labour supply gradient over the last couple of decades, when much of the world has access to oral contraceptives, we do not see support for changing access to birth control as an important explanation of our main findings.

We looked at two exercises for evidence on the role of changing social norms. Our first attempt borrows an idea from the important work of Goldin (1977), who traced persistent differences in black-white female labour force participation to different social norms about female work by race that arose during slavery. Boustan and Collins (2014) further show that this disparity persisted into the mid-20th century through the intergenerational transmission of work norms between mothers and daughters. Following them, we looked for differences in the labour supply gradient in the U.S. over time by race. We find that the gradients for whites and blacks follow the same general pattern, with the black labour supply gradient enduring a steeper decline in the 1950 and 1960 censuses (Figure A17). While interesting in its own right, the lack of any economic or statistical difference in the pre-WWII period when the labour supply effect of children is zero indicates that race-specific social norms about female work cannot explain the increasing costliness of a second child over development, at least in the U.S.

Secondly, we looked more directly at female work norms using a question from the General Social Survey (GSS): "Do you approve or disapprove of a married woman earning money in business or industry if she has a husband capable of supporting her?" We show (Figure A18) that the negative gradient across real GDP/capita is similar in economic magnitude for the bottom, middle, and top terciles of state-census years ranked by the share of respondents who do not approve of married women working outside the home within each year. That is, there is a declining labour supply elasticity between 1970 and 1980 that flattens out thereafter for each of the three "women work norm" tercile samples. Consequently, although these tests are limited to the U.S. experience, we see no compelling evidence to claim that evolving social norms influence our main results during this narrow time period.

We perform a wide range of robustness checks, examining the consequence of omitted variables bias, alternative benchmarks of development, and a variety of variable definition, specification, and sampling considerations. In particular, we examine the robustness of our instrumental variables strategy by trying an alternative instrument (time to first birth; see Klemp and Weisdorf, 2019), by including additional controls suggested by Bhalotra and Clarke (2016), and by splitting the results by same gender versus different gender twins. We present results that look at alternative development benchmarks on the x-axis, including average female wages (for the 1940 to 2010 U.S. censuses), and the average education level of women. Finally, we examine the robustness of our results to the choice of sample and specification.

The full set of results are described in detail in Appendix B. Among these, one result to highlight is the robustness of our results to the use of a more precise date of birth when available. In order to maximize data coverage, our main results define twins as being born in the same calendar year. However, for a subset of our data we also observe the month or quarter of birth, allowing us to rule out so-called Irish twins. The key patterns in our result are robust to the choice of sample and the more precise definition of twinning (Appendix Figure A19).

VI. Conclusion

In her classic monograph of the evolution of women's work in the United States, Goldin (1995) documents a U-shaped evolution of women's labour supply over the 20th century. At the same time, she notes the paucity of historical causal evidence on the link between fertility and labour supply. A parallel literature in development economics has investigated the implications of evolving patterns of fertility in developing countries on economic growth (and implicitly labour supply). While there have been many notable and pioneering studies on the effect of fertility on labour supply in developing countries, they naturally tend to focus on single countries or non-causal evidence.

Using a twin birth and same gender of the first two children as instruments for incremental fertility, this paper links these two literatures by examining causal evidence on the evolution of the response of labour supply to additional children across a wide swath of countries in the world and over 200 years of history. Our paper has two robust findings. First, the effect of fertility on labour supply is small, indeed typically indistinguishable from zero, at low levels of income and both negative and substantially larger at higher levels of income. Second, the magnitude of these effects is remarkably consistent across the contemporary cross-section of countries and the historical time series of individual countries, as well as across demographic and education groups.

The results are consistent with an increased time cost of looking after children, which seems to arise from changes in the sectoral and occupational structure of female jobs, in particular the rise of formal, non-professional, and non-agricultural wage work that flourishes with development. We also show that the negative gradient is steeper among mothers with young children that work in non-professional occupations and argue that access to child care subsidies may attenuate the negative gradient, suggesting that the affordability of child care costs may play a key role in declining LFP during the development cycle.

It is important to note that our findings are also consistent with and complementary to other explanations. Over the two-century-plus horizon we examine, there have been significant shifts in social norms regarding both work and fertility, parenting styles, and wide-spread adoption of modern contraceptives, among other plausible changes in the response of mother's work to fertility (Mammen and Paxson, 2000). While we have provided indirect evidence from the U.S. against some of these mechanisms, our data does not allow us to fully disentangle these plausible channels.

In discussing the evolution of female labour force participation in the United States, Goldin (1990) notes that "… women on farms and in cities were active participants [in labour] when the home and workplace were unified, and their participation likely declined as the marketplace widened and the specialization of tasks was enlarged." In examining the relationship between labour supply and fertility over the process of development, we arrive at a parallel conclusion. The declining female labour supply response to fertility is especially strong in wage work that is likely the least compatible with concurrent childcare.

We see three implications of our results. First, in thinking about the U-shaped pattern of labour force participation that has been widely documented in the economic history literature, our results suggest that decreases in fertility play an explanatory role. That is, as fertility rates have declined over the latter half of the 20th century, the responsiveness of labour supply to fertility has increased, contributing to increases in female labour force participation. Second, among developing countries, our results however suggest that changes in fertility (such as those documented in Chatterjee and Vogl, 2018) tend not to have a large impact on labour force participation, arguing against fertility-reduction policies specifically motivated by women's labour force participation and its contribution to growth. Third, our results provide an interesting example of the external validity of a diverse and seemingly different set of results on fertility and labour supply across the development, labour, and economic history literatures.

Bibliography

Abadie, A. (2003). 'Semiparametric Instrumental Variables Estimation of Treatment Response Models', *Journal of Econometrics*, vol. 133(2), pp. 231-263.

Adda, J., Dustman, C. and Stevens, K. (2017). 'The Career Cost of Children', *Journal of Political Economy*, vol. 125(2), pp. 293-337.

Agüero, J. and Marks, M. (2008). 'Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks', *American Economic Review*, vol. *98*(2), pp. 500–504.

Agüero, J. and Marks, M. (2011). 'Motherhood and Female Labor Supply in the Developing World Evidence from Infertility Shocks', *Journal of Human Resources*, vol. 46(4), pp. 800–826.

Angelov, N., Johansson, P. and Lindahl, E. (2016). 'Parenthood and the Gender Gap in Pay', *Journal of Labor Economics,* vol. 34(3), pp. 545-579.

Angrist, J. and Evans, W. (1996). 'Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size', NBER Working Paper no. 5778.

Angrist, Jand Evans, W. (1998). 'Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size', *American Economic Review*, vol. 88(3), pp. 450–477.

Angrist, J. and Fernández-Val, I. (2013) 'ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework', in (D. Acemoglu, M. Arellano and E. Dekel, eds.), *Advances in Economics and Econometrics: Tenth World Congress,* Econometric Society Monographs, pp. 401-434, New York: Cambridge University Press.

Angrist, J., Lavy, V. and Schlosser, A. (2010). 'Multiple Experiments for the Causal Link between the Quantity and Quality of Children', *Journal of Labor Economics* vol. 28(4), pp. 773-824.

Angrist, J., Pathak, P. and Walters, C. (2013). 'Explaining Charter School Effectiveness', *American Economic Journal: Applied Economics*, vol. 5(4), pp. 1-27.

Baker, M., Gruber, J. and Milligan, K. (2008). 'Universal Child-Care, Maternal Labor Supply, and Family Well-Being', *Journal of Political Economy*, vol. 116, pp. 709–45.

Bailey, M., Hershbein, B. and Miller, A. (2012). 'The Opt-In Revolution? Contraception and the Gender Gap in Wages', *American Economic Journal: Applied Economics*, vol. 4(3), pp. 225-254.

Bailey, M. (2013). 'Fifty Years of Family Planning: New Evidence on the Effects of Increasing Access to Contraception', *Brookings Papers on Economic Activity,* vol. 46(1), pp. 341-409.

Basso, O., Juul, S. and Olsen, J. (2000). 'Time to pregnancy as a correlate of fecundity: differential persistence in trying to become pregnant as a source of bias,' *International Journal of Epidemiology*, vol. 29(5), pp. 856-861.

Becker, G. (1960). 'An economic analysis of fertility', in (Universities-National Bureau of Economic Research Conference Series 11), *Demographic and Economic Change in Developed Countries,* pp. 209-231, New Jersey: Princeton NBER.

Berlinski, S., and Galiani, S. (2007). 'The Effect of a Large Expansion of Pre-Primary School Facilities on Preschool Attendance and Maternal Employment', *Labour Economics,* vol. 14, pp. 665–680.

Bisbee, J., Dehejia, R., Pop-Eleches, C. and Samii, C. (2017). 'Local Instruments, Global Extrapolation: External Validity of the Labor Supply-Fertility Local Average Treatment Effect', *Journal of Labor Economics*, vol. 35(S1), pp. 99-147.

Bhalotra, S.and Clarke, D. (2016). 'The Twin Instrument,' Working Paper, IZA.

Black, S., Devereux, P. and Salvanes, K. (2005). 'The More the Merrier? The Effect of Family Composition on Children's Education', *Quarterly Journal of Economics*, vol. 120(2), pp. 669-700.

Blau, F. D. and Winkler, A.E. (2019). 'Women, Work, and Family,' in (S. Averett, L. Argys, and S. Hoffman eds.), *The Oxford Handbook of Women and the Economy*, pp. 395-424, New York: Oxford University Press.

Bloom, D., Canning, D. and Sevilla, J. (2001). 'Economic Growth and Demographic Transition', NBER Working Paper no. 8685.

Bloom, D., Canning, D., Fink, G. and Finlay, J. (2009). 'Fertility, Female Labor Force Participation, and the Demographic Dividend', *Journal of Economic Growth*, vol. 14(2), pp. 79-101.

Bongaarts, J. (1975). 'A Method for the Estimation of Fecundability,' *Demography,* vol. 12, pp. 645-660.

Boustan, L.P. and Collins, W. (2014) 'The Origin and Persistence of Black-White Differences in Women's Labor Force Participation,' in (L.P. Boustan, C. Frydman and R.A. Margo eds.), *Human Capital in History: The American Record*, pp. 205-240, Chicago and London: University of Chicago Press.

Bronars, S., and Grogger, J. (1994). 'The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment', *American Economic Review*, vol. 84(5), pp. 1141-1156.

Labor Butikofer, A. (2011). 'Sibling Sex Composition and the Cost of Children,' Working Paper.

Caceres-Delpiano, J. (2006). 'The Impact of Family Size on Investment in Child Quality,' *Journal of Human Resources,* vol. 41(4), pp. 738-754.

LaborCascio, E. (2009). ‘Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools,’ *Journal of Human Resources,* vol. 44, pp. 140–70.

Chatterjee, S. and Vogl, T. (2018). ‘Escaping Malthus: Economic Growth and Fertility Change in the Developing World’, *American Economic Review*, vol. 108(6), pp. 1440-1467.

Clarke, D. (2018). ‘Children and their Parents: A Review of Fertility and Causality’, *Journal of Economic Surveys*, vol. 32(2), pp. 518-540.

Cristia, J. (2008). ‘The Effect of a First Child on Female Labor Supply: Evidence from Women Seeking Fertility Services’, *Journal of Human Resources*, vol. 43(3), pp. 487–510.

Cruces, G. and Galiani, S. (2007). ‘Fertility and female labor supply in Latin America: New causal evidence’, *Labour Economics*, vol. 14(3), pp. 565-573.

Dehejia, R., Pop-Eleches, C. and Samii, C. (2019). ‘From Local to Global: External Validity in a Fertility Natural Experiment’, *Journal of Business and Economic Statistics*, forthcoming.

Del Boca, D. (2015). ‘Child Care Arrangements and Labor Supply,’ Working Paper, Inter-American Development Bank.

Ebenstein, A. (2010). ‘The "Missing Girls" of China and the Unintended Consequences of the One Child Policy’, *Journal of Human Resources*, vol. 45(1), pp. 87-115.

Edwards, L. and Field-Hendry, E. (2002). ‘Home-Based Work and Women’s Labor Force Decisions’, *Journal of Labor Economics,* vol. 20(1), pp. 170-200.

Feng, W. and Quanhe, Y. (1996). ‘Age at Marriage and the First Birth Interval: The Emerging Change in Sexual Behavior Among Young Couples in China,’ *Population and Development Review,* vol. 22(2), pp. 299-320.

Fitzpatrick, M. (2012). ‘Revising our Thinking about the Relationship Between Maternal Labor Supply and Preschool’, *Journal of Human Resources,* vol. 47, pp. 583–612.

Galor, O. and Weil, D. (1996). ‘The Gender Gap, Fertility, and Growth’, *American Economic Review,* vol. 86(3), pp. 374-387.

Godefroy, R. (2017). ‘How Women’s Rights Affect Fertility: Evidence from Nigeria’, *The Economic Journal,* vol. 129(3), pp. 1247-1280.

Goldin, C. (1977). ‘Female Labor Force Participation: The Origin of Black and White Differences, 1870-1880’, *Journal of Economic History*, vol. 37(1), pp. 87-108.

Goldin, C. (1995). ‘The U-Shaped Female Labor Force Function in Economic Development and Economic History’, in (T.P. Schultz, ed.), *Investment in Women’s Human Capital and Economic Development*, pp. 61-90, Chicago: University of Chicago Press.

Goldin, C. (1990). *Understanding the Gender Gap: An Economic History of American Women*, New York: Oxford University Press.

Gronau, R. (1986). 'Home Production -- A Survey', in (O. Ashenfelter and R. Layard, eds.), *Handbook of Labor Economics*, vol. 1, pp. 273-304 Amsterdam: North-Holland.

Havnes, T. and Mogstad, M, (2011). 'Money for Nothing? Universal Child Care and Maternal Employment', *Journal of Public Economics,* vol. 95, pp. 1455–1465.

Heath, R. (2017). 'Fertility at Work: Children and Women's Labor Market Outcomes in Urban Ghana', *Journal of Development Economics*, vol. 126, pp. 190-214.

Herbst, C. (2017). 'Universal Child Care, Maternal Employment, and Children's Long-Run Outcomes: Evidence from the U.S. Lanham Act of 1940', *Journal of Labor Economics,* vol. 35(2), pp. 519-564.

Hoekstra, C., Zhao, Z.Z., Lambalk, C., Willemsen, G., Martin, N., Boomsma, D. and Montgomery, G. (2007). 'Dizygotic Twinning', *Human Reproduction Update*, vol. 14(1), pp. 1-11.

ICF International, 2015, Demographic and Health Surveys (various) [Datasets], Calverton, Maryland: ICF International [Distributor].

International Labour Organization, https://www.ilo.org/ilostat, retrieved February 8, 2019.

Jaffe, A. J. and Azumi, K. (1960). 'The birth rate and cottage industries in under-developed countries', *Economic Development and Cultural Change,* vol. 9, pp. 52-63.

Jayachandran, S. and Pande, R. (2017). 'Why Are Indian Children So Short? The Role of Birth Order and Son Preference', *American Economic Review*, vol. 107(9), pp. 2600-2629.

Jones, L. and Tertilt, M. (2008). 'An Economic History of Fertility in the United States: 1826-1960', in (P. Rupert, ed.), *Frontiers of Family Economics*, pp. 165-230, United Kingdom: Emerald Publishing.

Juul, S., Karmaus, W. and Olsen, J. (1999). 'Regional differences in waiting time to pregnancy: pregnancy-based surveys from Denmark, France, Germany, Italy and Sweden', *Human Reproduction,* vol. 14, pp. 1250-1254.

Klemp, M. and Weisdorf, J. (2019). 'Fecundity, Fertility and the Formation of Human Capital', *Economic Journal*, vol. 129(618), pp. 925-960.

Kleven, H., Landais, C. and Søgaard, J.E. (2019a). 'Children and Gender Inequality: Evidence from Denmark', *American Economic Journal: Applied Economics*, vol. 11(4), pp. 181-209.

Kleven, H., Landais, C., Posch, J., Steinhauer, A. and Zweimüller, J. (2019b). 'Child Penalties Across Countries: Evidence and Explanations', NBER Working Paper 25524.

Kupinsky, S. (1977). *The Fertility of Working Women: A Synthesis of International Research*, New York: Praeger Publishers.

Kuziemko, I., Pan, J., Shen, J. and Washington, E. (2018). 'The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?', NBER Working Paper 24740.

Lundborg, P., Plug, E. and Rasmussen, A.W. (2017). 'Can Women Have Children and a Career? IV Evidence from IVF Treatments', *American Economic Review*, vol. 107(6), pp. 1611-1637.

The Maddison-Project, http://www.ggdc.net/maddison/maddison-project/home.htm, 2013 version.

Mammen, K. and Paxson, C. (2000). 'Women's Work and Economic Development', *Journal of Economic Perspectives*, vol. 14(4), pp. 141-164.

Maurin, E. and Moschion, J. (2009). 'The Social Multiplier and Labor Market Participation of Mothers', *American Economic Journal: Applied Economics*, vol. 1(1), pp. 251-272.

McCabe, J. and Rosenzweig, M.R. (1976). 'Female Labor Force Participation, Occupational Choice, and Fertility in Developing Countries', *Journal of Development Economics,* vol. 3, pp. 141-160.

Minnesota Population Center, 2015, North Atlantic Population Project: Complete Count Microdata, Version 2.2 [Machine-readable database], Minneapolis: Minnesota Population Center.

Moehling, C. (2002). 'Broken Homes: The 'Missing' Children of the 1910 Census', *Journal of Interdisciplinary History*, vol. 33(2), pp. 205-233.

Morrissey, T. (2017). 'Child Care and Parent Labor Force Participation: a Review of the Research Literature', *Review of Economics of the Household*, vol. 15, pp. 1–24.

Olivetti, C. and Petrongolo, B. (2008). 'Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps', *Journal of Labor Economics*, vol. 26(4), pp. 621-654.

Olivetti, C. and Petrongolo, B. (2017). 'The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries', *Journal of Economic Perspectives*, vol. 31(1), pp. 205-230.

Rosenzweig, M. and Wolpin, K. (1980). 'Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment', *Econometrica,* vol. 48(1), pp. 227–240.

Rosenzweig, M. and Wolpin, K. (2000). 'Natural 'Natural Experiments' in Economics', *Journal of Economic Literature,* vol. 38(4), pp. 827-874.

Rosenzweig, M. and Zhang, J. (2009). 'Do Population Control Policies Induce More Human Capital Investment? Twin, Birth Weight and China's 'One-Child' Policy', *The Review of Economic Studies*, vol. 76(3), pp. 1149-1174.

Ruggles, S., Alexander, J.T., Genadek, K., Goeken, R., Schroeder, M.B., and Sobek, M. (2010). Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database], Minneapolis: University of Minnesota.

Schultz, T.P. (1991). 'International Differences in Labor Force Participation in Families and Firms', Working Paper, Yale Economic Growth Center Discussion Paper No. 634.

Schultz, T.P. (2008). 'Population Policies, Fertility, Women's Human Capital, and Child Quality', in (P.T. Schultz and J.A. Strauss, eds.), *Handbook of Development Economics, Volume Four*, pp. 3249–3303, Amsterdam: Elsevier Science B.V.

Sobek, M. and Kennedy, S. (2009). 'The Development of Family Interrelationship Measures for International Census Data', Working Paper, University of Minnesota Population Center.

Szulga, R. (2014). 'A Dynamic Model of Female Labor Force Participation Rate and Human Capital Investment', *Journal of Economic Development*, vol. 39(3), pp. 81-114.

United States Census Bureau, 1970, "National Income and Wealth," Series F 297-314, p. 243, retrieved from: http://www2.census.gov/library/publications/1975/compendia/hist_stats_colonial-1970/hist_stats_colonial-1970p1-chF.pdf

Vere, J. (2011). 'Fertility and Parents' Labour Supply: New Evidence from US Census Data', *Oxford Economic Papers*, vol. 63(2), pp. 211-231.

Willis, R. (1973). 'A New Approach to the Economic Theory of Fertility Behavior', *Journal of Political Economy*, vol. 81, pp. S14-S64.

Appendix A: A Sketch of a Model

We show that the differential female labour supply response to children over the development cycle can be explained within a standard labour-leisure model. Consider a constant elasticity of substitution (CES) utility function defined over consumption *c*, leisure *d*, and fertility *n*:

(A.1) $$U(c,d,n) = \left[\gamma(c+c_0)^\rho + \alpha d^\rho + \beta\left(\frac{n}{N}\right)^\rho\right]^{1/\rho}$$

where c_0 <0 is subsistence consumption and utility from fertility is relative to potential reproductive capacity *N*. Equation (1) is a CES variant of the model used by Bloom et al. (2009). Total time (normalized to 1) is allocated between leisure *d*, childcare *bn* (where *b* is the time cost per child), labour *l*, and non-market household work ε:

(A.2) $$1 = l + d + bn + \varepsilon$$

Assuming households do not save, consumption is derived directly from earned income:

(A.3) $$c = wl.$$

Substituting equations (2) and (3) into (1), we obtain the household utility function:

(A.4) $$V(l,n) = \left[\gamma(wl + c_0)^\rho + \alpha(1 - l - bn - \varepsilon)^\rho + \beta\left(\frac{n}{N}\right)^\rho\right]^{1/\rho}.$$

The first order conditions are:

(A.5) $$\partial V/\partial l = \frac{1}{\rho}v^{\left(\frac{1}{\rho}-1\right)}[\rho\gamma w(wl + c_0)^{\rho-1} - \alpha\rho(1 - l - bn - \varepsilon)^{\rho-1}] = 0$$

$$\partial V/\partial n = \frac{1}{\rho}v^{\left(\frac{1}{\rho}-1\right)}[-\alpha\rho b(1 - l - bn - \varepsilon)^{\rho-1} + \beta\rho N^{-\rho}n^{\rho-1}] = 0$$

where $v \equiv \left[\gamma(wl + c_0)^\rho + \alpha(1 - l - bn - \varepsilon)^\rho + \beta\left(\frac{n}{N}\right)^\rho\right]$. Re-arranging yields:

(A.6) $$l = \frac{(\alpha^\theta - \alpha^\theta\epsilon - w^\theta\gamma^\theta c_0) - \alpha^\theta bn}{w^{\theta+1}\gamma^\theta + \alpha^\theta}$$

$$n = \frac{\alpha^\theta b^\theta(1 - \varepsilon - l)}{\beta^\theta N^{-\rho\theta} + \alpha^\theta b^{\theta+1}},$$

where $\theta \equiv 1/(\rho - 1)$. Note that in the solution:

(A.7) $$\frac{\partial l}{\partial n} = -\frac{\alpha^\theta b}{w^{\theta+1}\gamma^\theta + \alpha^\theta} < 0$$

and $\partial^2 l/\partial n\partial w < 0$ if $\rho \in (0,1)$ or the elasticity of substitution is between (0,∞). Of note, the model predicts the effect of fertility on labour supply becomes more negative as the wage increases. As the wage increases, the agent experiences both a substitution and income effect. The former arises because an increase in the wage causes the price of leisure and the time-cost of children to also increase, leading to a substitution into labour and out of children. Higher wages also increase income, which moves households away from labour and toward children. When the elasticity of substitution is positive, the substitution effects tends to dominate, increasing the responsiveness of labour to fertility as the wage goes up.[36]

In a small number of low-income countries, including pre-WWI U.S., we estimate a

[36] We also considered the consequences of changing wages using the model in Angrist and Evans (1996). That model finds that for parent $i \in \{1,2\}$, the change in work in response to fertility can be expressed as $\frac{\partial t_i}{\partial n} = -\left(\frac{\partial h_i}{\partial n} + \frac{\partial l_i}{\partial n}\right)$ where t_i is work time, h_i is home time, l_i is leisure time, and n is number of children. We note that this derivative can be further decomposed as $\frac{\partial h_i}{\partial n} = w_i A_i$ and $\frac{\partial l_i}{\partial n} = w_i \frac{\partial l_i}{\partial \lambda}\frac{\partial \lambda}{\partial n}$ where w_i is the wage of parent i, A_i is a function of choice variables and parameters that do not include w_i, and λ is the marginal value of income. Note that the terms inside the parentheses of $\frac{\partial t_i}{\partial n} = -w_i\left(A_i + \frac{\partial l_i}{\partial \lambda}\frac{\partial \lambda}{\partial n}\right)$ do not depend on w_i since neither $\frac{\partial l_i}{\partial \lambda}$ nor $\frac{\partial \lambda}{\partial n}$ include w_i. Angrist and Evans (1996) show that the total effect of fertility on work time is ambiguous. However, their result is invariant to the sign of the effect; regardless of the sign, increasing the wage will amplify the response. Since nearly all empirical work has established that $\frac{\partial t_i}{\partial n} \leq 0$, we should expect to find that $\frac{\partial^2 t_i}{\partial n\partial w_i} \leq 0$ as well.

modest positive labour supply response to children. While equation (7) predicts a negative response, a positive result is possible with a simple extension of the model. Suppose there is a consumption (e.g., food) cost to children so $c = wl - kn$, and for simplicity set c_0 and ε to zero. The first-order condition with respect to labour, with rearrangement, now becomes:

(A.8) $$l = \frac{\alpha^\theta + n(w^\theta \gamma^\theta k - \alpha^\theta b)}{w^{\theta+1}\gamma^\theta + \alpha^\theta}.$$

In this case $\partial l/\partial n > 0$ is consistent with $k > \alpha^\theta b/\gamma^\theta w^\theta$. An increase in fertility implies an increased time cost but also a reduction in consumption, making increased labour more valuable. Since $\theta < 0$, if the wage or the time cost of children are sufficiently low relative to the consumption cost, mothers optimally increase labour. In this case, $\partial^2 l/\partial n \partial w < 0$ without further assumptions, so we would continue to expect a negative gradient of the fertility-labour relationship with respect to the wage.[37]

[37] Note $\mathrm{sgn}(\partial^2 l/\partial n \partial w) = \mathrm{sgn}\left(-\gamma^\theta k \gamma w^\theta + \theta k w^{-1}\alpha^\theta + (\theta+1)\alpha^\theta\right) = -1$ if $\rho \in (0,1)$.

Appendix B: Robustness Checks

In this appendix, we present a series of detailed robustness checks for our results.

a. Omitted Variables and Alternative Sources of Identification

Twin and same gender instruments are susceptible to omitted variables biases. These biases are likely to differ across instrument, suggesting that the twins and same gender IV estimates can be viewed as specification checks of each other (Angrist, Lavy, and Schlosser 2010). However, in this subsection, we push this idea further by describing three other sets of estimates that exploit alternative sources of instrument variation or control for observable characteristics that are known to explain variation in the treatment.

First, we examine a third instrument for fertility – the time that elapses between the parents' marriage and the couple's first birth ("time to first birth" or TFB) – introduced by Klemp and Weisdorf (2019). A long line of research in demography and medicine (Bongaarts 1975) uses birth spacing, not necessarily limited to first births, as an indicator of fecundity. While there is mixed evidence on the extent to which spacing is idiosyncratic (Feng and Quanhe 1996; Basso, Juul, and Olsen 2000; and Juul, Karmaus, and Olsen 1999), Klemp and Weisdorf argue that TFB is especially hard to predict based on observable characteristics outside of parent age and consequently is a valid indicator of ultimate family size. Because TFB requires marriage and birth dates, which are only available in the DHS, we cannot replicate the negative gradient across the development cycle. However, we do find that the TFB IV estimates are economically small and positive and statistically similar to twin IV and same gender estimates at the same low real GDP per capita level.[38]

Second, it has been noted by many researchers, most recently Bhalotra and Clarke (2016), that mothers of twins may be positively selected by health and wealth.[39] We provide two additional pieces of evidence that this selection process is not driving the negative labour supply gradient. When we control for the observable characteristics that have been highlighted by Bhalotra and Clarke (2016), such as mother's education, medical care availability, and mother's health, our results are statistically identical to the baseline estimates without these controls.

[38] The TFB IV estimates using the DHS data are: 0.031 (0.018), 0.050 (0.015), and 0.043 (0.014) for the $0-2,500, $2,500-5,000, and $5,000-10,000 GDP per capita bins, respectively.
[39] Relatedly, Rosenzweig and Zhang (2009) argue twins are less costly to raise than two singleton births spaced apart. While we cannot fully address this concern, we can restrict the analysis to mothers with close birth-spacing. Appendix Figure A20 shows that this restriction has little impact on our results.

Appendix Figure A21 plots the results with and without mother's education covariates using all available censuses and the DHS. We are also able to roughly replicate Bhalotra and Clarke's association between twinning and doctor availability, nurse availability, prenatal care availability, mother's height, mother's BMI (underweight and obese dummies), and infant mortality prior to birth. These Health measures are available only in the DHS. When we specifically control for these measures, our labour supply IV estimates are identical to the baseline for the <$2,500 bin and only slightly larger but statistically and economically indistinguishable for the $2,500-$5,000 bin (-0.006 (0.031) versus 0.012 (0.028)) and $5,000 and over bin (-0.075 (0.042) versus -0.044 (0.039)). [40]

Third, a strand of the medical literature argues that the proportion of dizygotic twins is affected by environmental and genetic factors of the type discussed by Bhalotra and Clarke (2016). By contrast, the proportion of monozygotic twins appears to be relatively constant over time and less affected by their omitted variables bias concern.[41] We find (in Figure A23) that results are statistically indistinguishable across same and opposite gender twins, lending additional credence to the view that our results are not driven by omitted variable bias with respect to twinning.

b. Alternative Development Benchmarks

The labour supply patterns we have documented thus far are based on an economy's real GDP per capita. The key model prediction, however, is based on the substitution and income effects arising from changes to a woman's wage. Unfortunately, data limitations make it difficult to show world results stratified by female (or overall) wages. However, for the 1940 to 2010 U.S. censuses, we can compute average female real wage rates by state and census year.[42] Results are

[40] In addition to controlling for mothers' education, we split the results by mother's education (Appendix Figure A22). There is no statistical or economic difference by mother's education at any level of GDP per capita.

[41] We cannot identify monozygotic and dizygotic twins in our data but we can exploit the fact that monozygotic twins are always same gender, whereas dizygotic twins are an equal mix of same and opposite gender (like non-twin siblings). The rate of monozygotic twinning is approximately 4 per 1000 births and is constant across various subgroups (Hoekstra et al. 2007). Under the standard assumption that dizygotic twins have a 50 percent chance of being the same gender, approximately 43 to 59 percent of same-gender twins are monozygotic across the various GDP bins. Notably, the proportion of monozygotic twins will be highest in low-GDP countries, where Bhalotra and Clarke (2016) find the potential for the omitted variable bias is greatest.

[42] There is no wage data prior to 1940. For all persons aged 18 to 64, we calculate the average hourly wage rate as annual earned income divided by weeks worked times hours worked per week. The age range overlaps with the cohort of mothers used in our baseline sample but we do not condition on gender or motherhood. The results are robust to using the average wage rate of men or women only as well. Wages are inflation adjusted using the consumer price index to 1990 dollars and winsorized at the 1st and 99th percentiles in each census prior to taking means.

presented in Figure A24, stratifying observations into four real hourly wage bins, ranging from under \$6 to over \$12 per hour, based on the average wage in the state at that time. Similar to the GDP per capita results, we find no labour supply effect at the lowest real wage levels and larger negative effects as the real hourly wage rises. Second, again for a subset of the sample, we can stratify by the average education level of women aged 21 to 35 (Appendix Figure A25).[43] We again find no effect at low education levels (below 9 years) but decreasing negative effects thereafter. Third, and perhaps more directly tied to Schultz (1991), we find the same pattern by agricultural employment. In this case, the negative gradient begins when agricultural employment drops below 15 percent.

c. Other Data, Specification, and Modelling Issues

Several variable definition choices that we make in our baseline estimates could conceivably be problematic, including a) using calendar year to identify twins, b) using occupation to define LFP in historical censuses, and c) counting non-biological children. We discuss each of these issues in turn.

Since few censuses record multiple births or the birth month/quarter, out of necessity we label siblings of the same age as twins. Naturally, this classification raises the risk that two births in the same calendar year could be successive rather than twins (so-called Irish twins). Fortunately, for a subset of our data, quarter or month of birth or direct measures of multiple births are available. Appendix Figure A19 presents results using both definitions of twins. By and large, we see a very similar negative gradient despite notably nosier estimates from a smaller sample of country-years with month or quarter of birth.[44]

Second, our historical results (in the U.S., 1930 and earlier) use an occupation-based measure of labour force participation. Post-1940, we switch to the modern LFP definition based on whether the person is working or searching for work at the time of the survey. When both LFP measures are available, initially and most prominently in the 1940 U.S. census, changing LFP definitions has no impact on our results. Using the full population 1940 U.S. census, we find

[43] Again, data availability limits our analysis to 1940 and later. We also exclude 29 country-years where years of education are not provided. By 1940, U.S. women in their twenties and thirties had, on average, at least 9 years of education. Consequently, the U.S. is included only in the two highest education bins (9 to 12 and 12+ years).
[44] By comparing the baseline and year-of-birth twin lines which both use the year-of-birth twin definitions but run regressions on different samples, it appears that the low-income country-years with month and quarter of birth are biased away from zero whereas the opposite is the case for high-income countries. Nevertheless, the line with twins defined by month or quarter of birth still exhibits a negative gradient.

a 0.94 *cross-state* correlation between the two measures and a 0.85 *cross-state* correlation of the IV results. More generally, Appendix Figure A26 illustrates the same general pattern of results when using: a) an occupation-based LFP for all censuses (U.S. and non-U.S.) that contain occupation, b) an indicator of whether the mother is employed at the time of the census/survey or c) an indicator of whether the mother worked over the prior year.

Despite the correspondence between the modern definition of LFP and the historical occupation-based results, there is still valid concern that specific women's occupations are misreported prior to 1940 and therefore could bias our results. In particular, Goldin (1990) highlights the mismeasurement of agricultural women workers in cotton growing states, an undercount of women in manufacturing, and mismeasurement of boarding-house keepers. While it is not possible to directly address the issues raised by Goldin, Appendix Figure A27 presents pre-1940 results that individually and simultaneously adjust the sample or outcome variable for each of these concerns.[45] Again, the findings are qualitatively similar to our baseline.

Another measurement concern relates to non-biological children and children who have left the household. Data identifying biological children are not consistently available across censuses. However, when we have information on the number of children to which a mother has given birth, we find that restricting our sample to mothers where this number matches the total number of children in the household has little impact on the results (see Appendix Figure A28). This restriction addresses concerns resulting from infant mortality, older children moving out the household, and complications resulting from stepchildren and children placed into foster care (Moehling 2002).

More broadly, we find it reassuring that the key pattern in the data is preserved when excluding the lower quality, pre-1940 data altogether. Namely, the female labour supply response to children in 1940 was economically small (Figures A4 and A7) and only turns statistically significant post-1940. This pattern suggests that our main findings are not driven by inconsistent historical data and sampling. In addition, our various robustness checks suggest that data issues are not the reason for the relatively constant labour supply response to children in the half century or so leading up to WWII.

[45] That is, we exclude women in cotton growing states and who list their industry as manufacturing. As an upper bound for boardinghouse keeper employment, we recode women as employed if the household has any members who identify their relationship to the household head as a boarder.

Finally, our findings are robust to a number of other reasonable tweaks to our specification, variable definitions, and sample selection. For example, we find larger negative effects among single (relative to married) mothers and children, especially in countries with higher GDP per capita (see Appendix Figure A29). (We find a similar result for younger mothers relative to older mothers.) All these cases exhibit the same negative gradient across the development cycle. We also find that specification and modelling choices – such as weighting each sample equally or using a Bayesian hierarchical model to smooth each country-year estimate – have no substantive impact on the results.

Table 1: Sample summary statistics by real GDP/capita bin

	Mothers	Samples	In labor force	3 or more children	2nd child is multiple birth	First 2 children are same gender	Children in household	Mother's age at survey	Mother's age at first birth
U.S.									
$0 - 2,500$	32,531	2	5.12%	62.47%	0.74%	49.48%	3.27	29.02	21.04
$2,500 - 5,000$	5,530,793	2	6.30%	62.47%	0.89%	50.27%	3.28	29.10	21.06
$5,000 - 7,500$	12,899,725	3	8.68%	55.75%	0.81%	50.37%	3.10	29.29	21.15
$7,500 - 10,000$	4,724,927	2	10.68%	47.07%	0.87%	50.50%	2.88	29.48	20.94
$10,000 - 15,000$	470,378	1	22.85%	55.09%	1.70%	50.38%	2.99	29.30	21.40
$15,000 - 20,000$	692,165	2	44.95%	40.85%	1.31%	50.48%	2.62	29.62	21.03
$20,000 - 35,000$	1,312,550	3	62.90%	36.64%	1.46%	50.58%	2.50	30.28	21.85
Non-U.S.									
$0 - 2,500$	9,676,791	213	43.33%	57.20%	1.28%	50.22%	3.06	29.07	20.66
$2,500 - 5,000$	7,617,815	103	36.14%	50.66%	1.05%	50.34%	2.96	29.82	21.19
$5,000 - 7,500$	4,192,823	52	36.77%	45.95%	1.22%	50.39%	2.77	29.43	20.46
$7,500 - 10,000$	2,184,583	20	34.95%	43.88%	1.25%	50.65%	2.69	29.54	20.66
$10,000 - 15,000$	614,503	19	37.90%	36.34%	1.19%	50.57%	2.61	29.99	21.63
$15,000 - 20,000$	415,161	10	56.06%	30.65%	1.19%	50.53%	2.41	30.73	22.61
$20,000 - 35,000$	1,085,025	9	73.66%	28.99%	1.44%	50.58%	2.38	31.23	24.00

Notes: This table displays summary statistics for the baseline sample of mothers by real GDP/capita bins. The sample consists of all two-child mothers aged 21 to 35 that were at least 15 when they had their first child, their oldest child is younger than 18, they do not live in group quarters, their first child is not a multiple birth, and mother and child have no imputations on age and gender. A twin is defined as the second and third birth being the same age. The samples directly correspond to those used in Table 2 and Figures 1, 3, and 6

Table 2: Baseline estimates by real GDP/capita bin

	Mothers	Samples	LFP	OLS	Twin FS	Twin 2S	Same-Gender FS	Same-Gender 2S
US: $0-2,500$	32,531	2	5.12%	-0.018***	0.345***	0.119***	0.015*	-0.068
				(0.006)	(0.018)	(0.005)	(0.007)	(0.162)
US: $2,500-5,000$	5,530,793	2	6.30%	-0.026***	0.363***	0.022***	0.009***	0.064***
				(0.002)	(0.013)	(0.008)	(0.000)	(0.013)
US: $5,000-7,500$	12,899,725	3	8.68%	-0.033***	0.451***	0.012	0.014***	0.032***
				(0.010)	(0.017)	(0.010)	(0.002)	(0.004)
US: $7,500-10,000$	4,724,927	2	10.68%	-0.064***	0.540***	-0.017***	0.021***	0.072***
				(0.001)	(0.002)	(0.001)	(0.000)	(0.002)
US: $10,000-15,000$	470,378	1	22.85%	-0.117***	0.452***	-0.033***	0.035***	-0.084**
				(0.001)	(0.002)	(0.010)	(0.001)	(0.034)
US: $15,000-20,000$	692,165	2	44.95%	-0.166***	0.575***	-0.059***	0.049***	-0.121***
				(0.015)	(0.065)	(0.018)	(0.006)	(0.007)
US: $20,000-35,000$	1,312,550	3	62.90%	-0.149***	0.636***	-0.070***	0.049***	-0.121***
				(0.010)	(0.007)	(0.008)	(0.001)	(0.008)
Non-US: $0-2,500$	9,676,791	213	43.33%	-0.022***	0.411***	-0.005	0.028***	-0.046**
				(0.005)	(0.018)	(0.009)	(0.007)	(0.019)
Non-US: $2,500-5,000$	7,617,815	103	36.14%	-0.058***	0.473***	-0.014	0.030***	-0.018
				(0.007)	(0.036)	(0.011)	(0.007)	(0.012)
Non-US: $5,000-7,500$	4,192,823	52	36.77%	-0.088***	0.545***	-0.003	0.035***	-0.037***
				(0.012)	(0.020)	(0.015)	(0.002)	(0.013)
Non-US: $7,500-10,000$	2,184,583	20	34.95%	-0.113***	0.548***	-0.033***	0.032***	-0.001
				(0.004)	(0.023)	(0.011)	(0.001)	(0.029)
Non-US: $10,000-15,000$	614,503	19	37.90%	-0.138***	0.604***	-0.089***	0.035***	-0.061*
				(0.023)	(0.064)	(0.016)	(0.004)	(0.035)
Non-US: $15,000-20,000$	415,161	10	56.06%	-0.276***	0.719***	-0.127***	0.042***	-0.205***
				(0.035)	(0.038)	(0.036)	(0.002)	(0.020)
Non-US: $20,000-35,000$	1,085,025	9	73.66%	-0.247***	0.706***	-0.105***	0.038***	-0.173***
				(0.009)	(0.003)	(0.003)	(0.001)	(0.019)

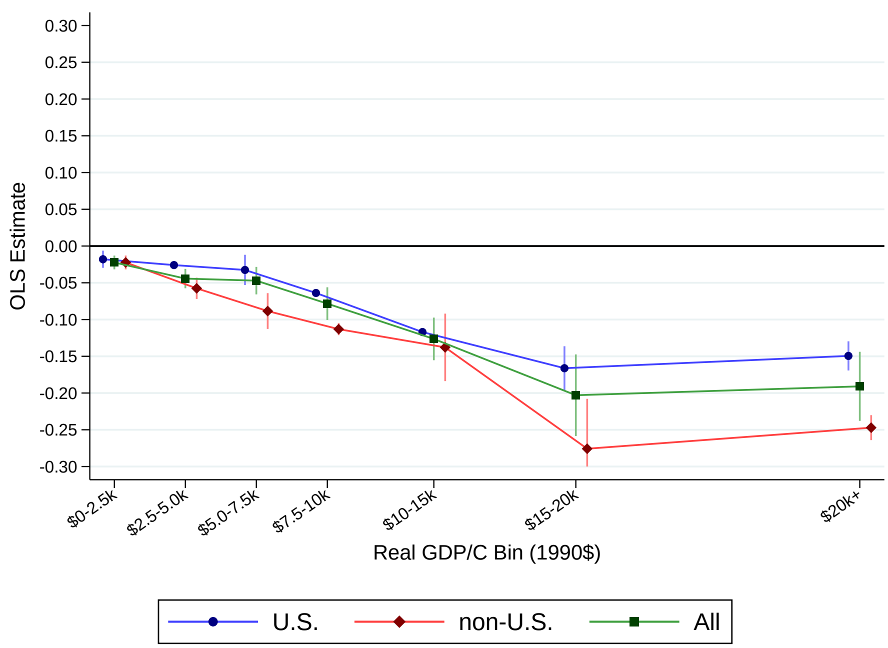
Notes: This table displays OLS, same gender and twin first stage (FS) and second stage (2S) IV estimates of the effect of a third birth on mother's labor force participation using the baseline sample of mothers described in the text and Table 1. Regressions control for mother's age, age at first birth, gender of first child (and second child for same gender IV), and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. These estimates are plotted in Figures 1, 3, and 6.

Table 3: Estimates by mother's professional status by the age of youngest child

	Mom occupation is professional			*Mom occupation is non-professional*		
	0 to 5	6 to 11	12 to 17	0 to 5	6 to 11	12 to 17
$\leq$ 10k	-0.007***	-0.005***	-0.006**	0.004	-0.008	-0.007
	(0.002)	(0.002)	(0.003)	(0.006)	(0.005)	(0.015)
$>$ 10k	-0.025***	-0.015***	-0.024***	-0.064***	-0.061***	-0.027*
	(0.004)	(0.005)	(0.006)	(0.008)	(0.009)	(0.015)
Gradient	-0.019***	-0.009*	-0.018***	-0.067***	-0.053***	-0.020
	(0.004)	(0.005)	(0.006)	(0.010)	(0.011)	(0.021)

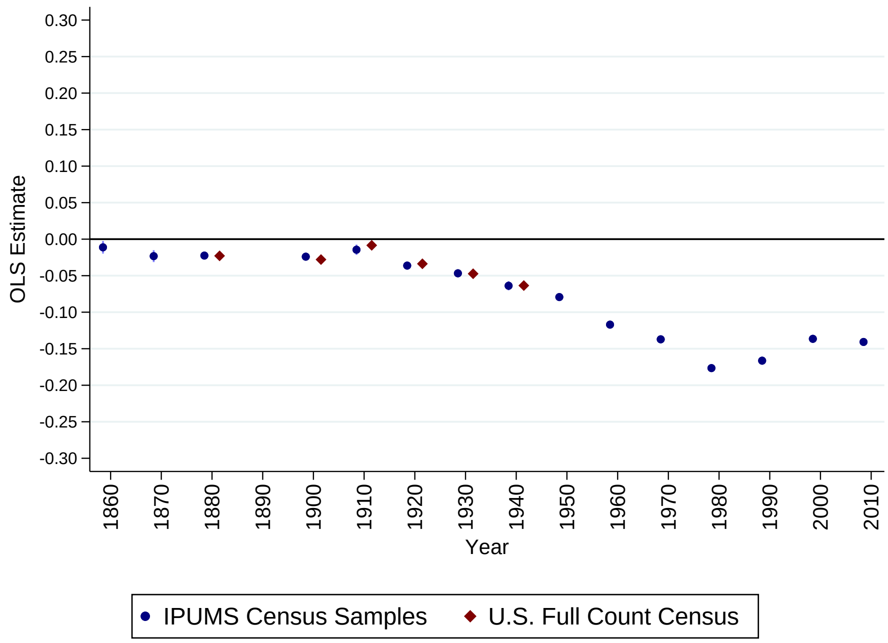
Notes: This table displays second-stage, twin IV estimates of the effect of a third birth on the occupational status of mothers using the baseline sample who also report occupational status. The samples are stratified by the age of the youngest child (0-5, 6-11, and 12-17). ''Gradient" refers to the difference between row 2 (countries with real GDP per capita of at least $10,000$ in 1990$) and row 1 (countries with real GDP per capita under $10,000$ in 1990$). Regressions control for mother's age, age at first birth, gender of first child (and second child for same gender IV), and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. See footnote 28 in the text for a description of the definition of professional and non-professional occupations in each data source.

Figure 1: OLS, by real GDP/capita



Notes: This figure displays OLS estimates of the relationship between having a third birth and mothers' labor force participation using the baseline sample of mothers in each GDP/capita bin. Matching OLS estimates for U.S. and non-U.S. samples are reported in Table 2. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals are displayed but may not always be visible at the scale of the figure.

Figure 2: OLS, U.S. by time

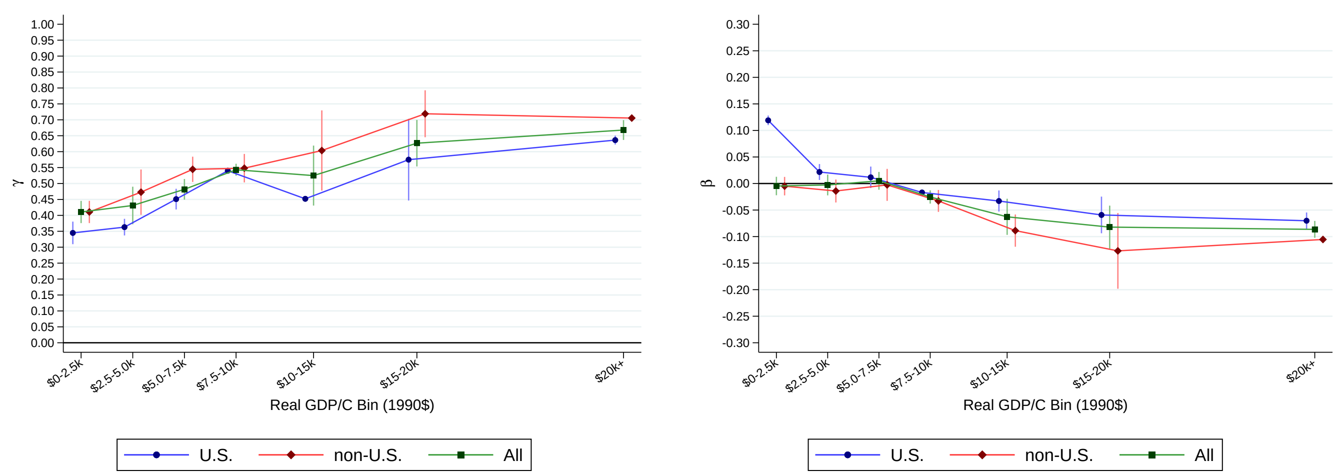


*Notes:*This figure displays OLS estimates of the relationship between having a third birth and mothers' labor force participation, binned by census year. It uses the baseline sample of mothers for the US only. Regressions control for mother's age, age at first birth, and gender of first child. Standard errors are robust to heteroskedasticity. 95 percent confidence intervals are displayed but may not always be visible at the scale of the figure. The estimates from this figure are reported in Appendix Table A1.

Figure 3: Twin IV, by real GDP/capita

(a) First Stage: third birth

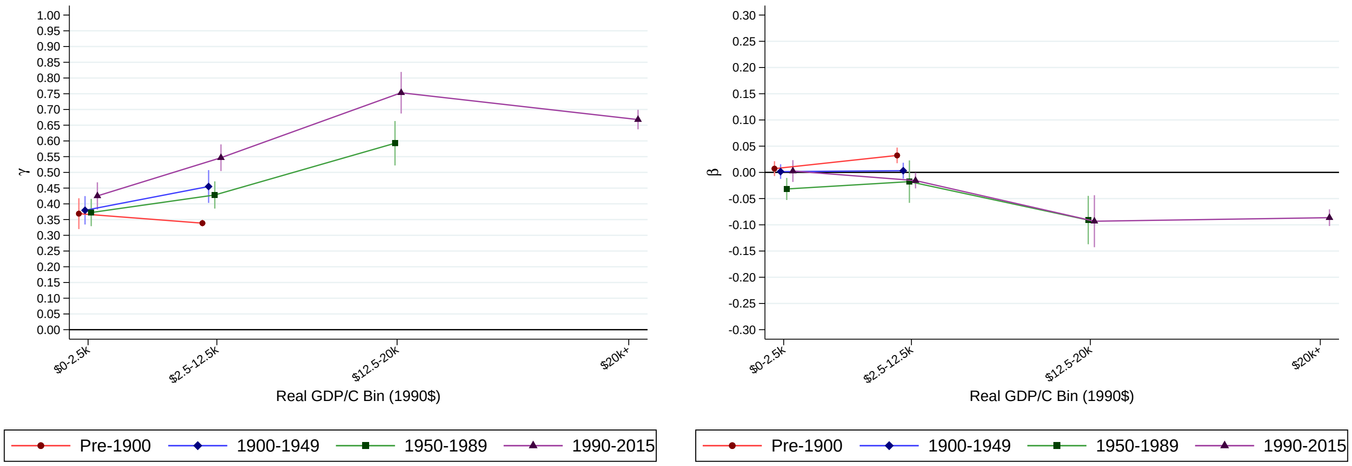
(b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates using the baseline sample of mothers for each each real GDP/capita bin. Panel (a) shows the first-stage estimates of the relationship between twins and having a third birth. Panel (b) shows the second-stage estimates of the relationship between having a third birth and mothers' labor force participation. These estimates are also reported in Table 2. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 4: Twin IV, by time and real GDP/capita bin

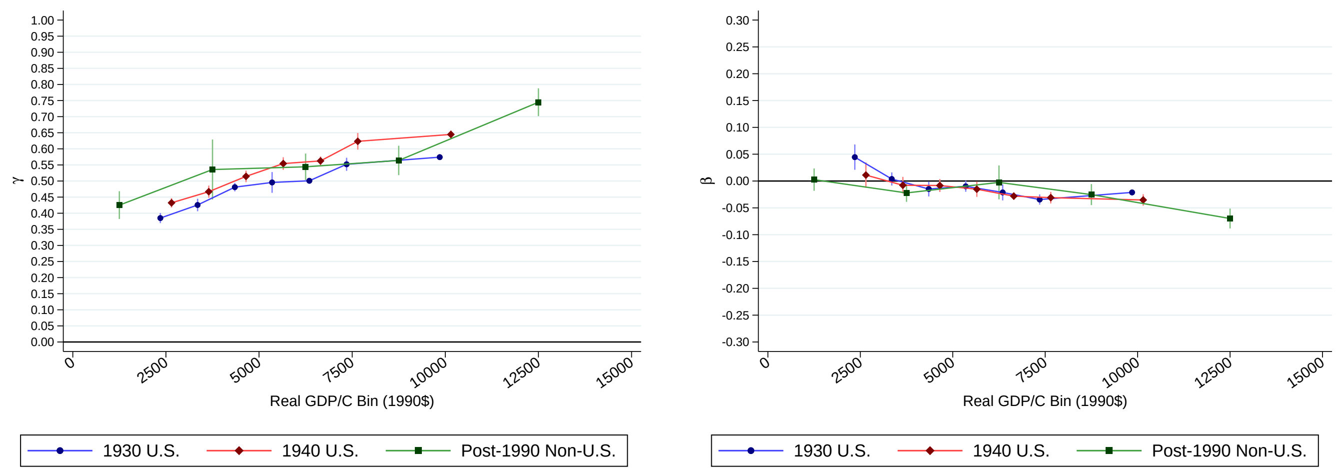
(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



*Notes:*This figure presents twin IV estimates stratified by year. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 5: Twin IV by 1930 and 1940 U.S. state compared to modern non-U.S countries

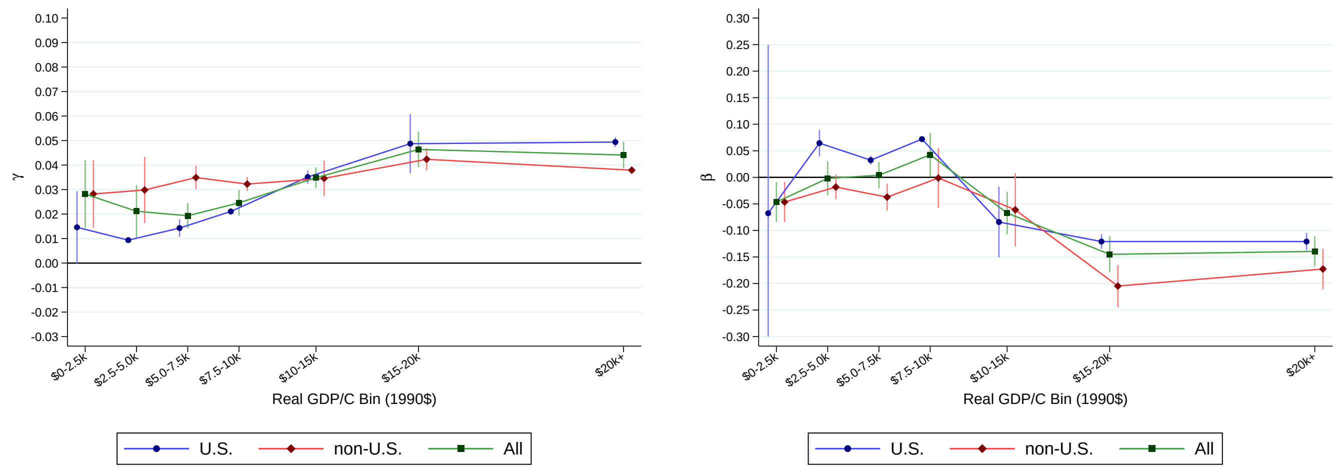
(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates from the 1930 and 1940 full count censuses, binned by state real income per capita. For comparison, we also plot the post-1990 non-U.S. estimates over the same real GDP/capita range. Income/capita for U.S. states is taken from the U.S. Census Bureau (see http://www2.census.gov/library/publications/1975/compendia/hist˙stats˙colonial-1970/hist˙stats˙colonial- 1970p1-chF.pdf). Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 6: Same gender IV, by real GDP/capita

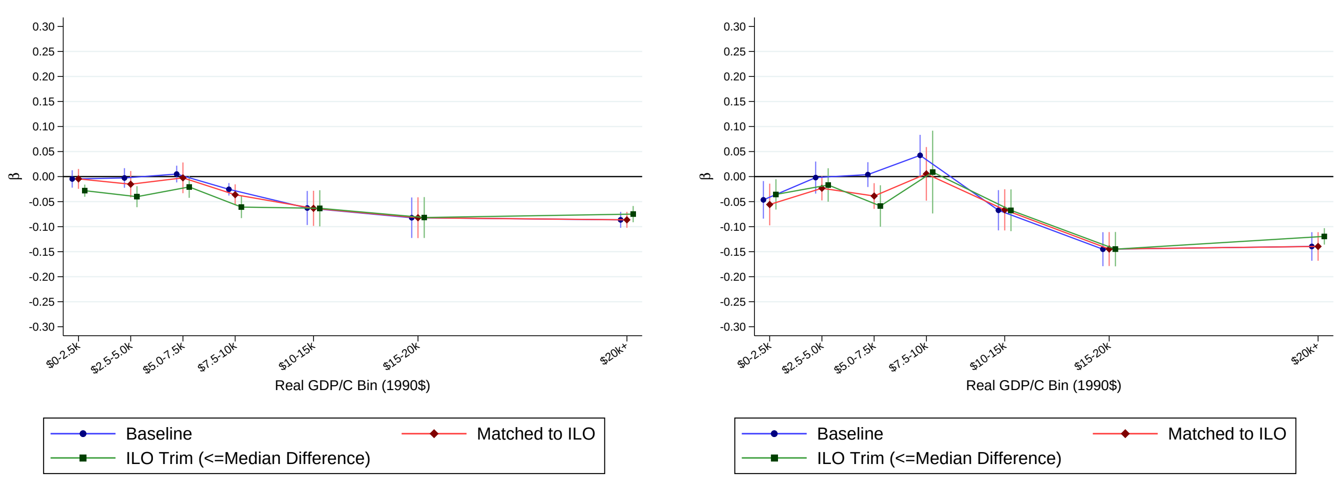
(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



Notes: This figure displays same gender IV estimates using the baseline sample of mothers for each each real GDP/capita bin. Analogous to figure 3, Panel (a) shows the first-stage estimates of the relationship between same gender children and having a third birth and Panel (b) shows the second-stage estimates of the relationship between having a third birth and mothers' labor force status. These estimates are also reported in Table 2. Regressions control for mother's age, age at first birth, gender of first two children, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 7: Second stage estimates matched to ILO statistics

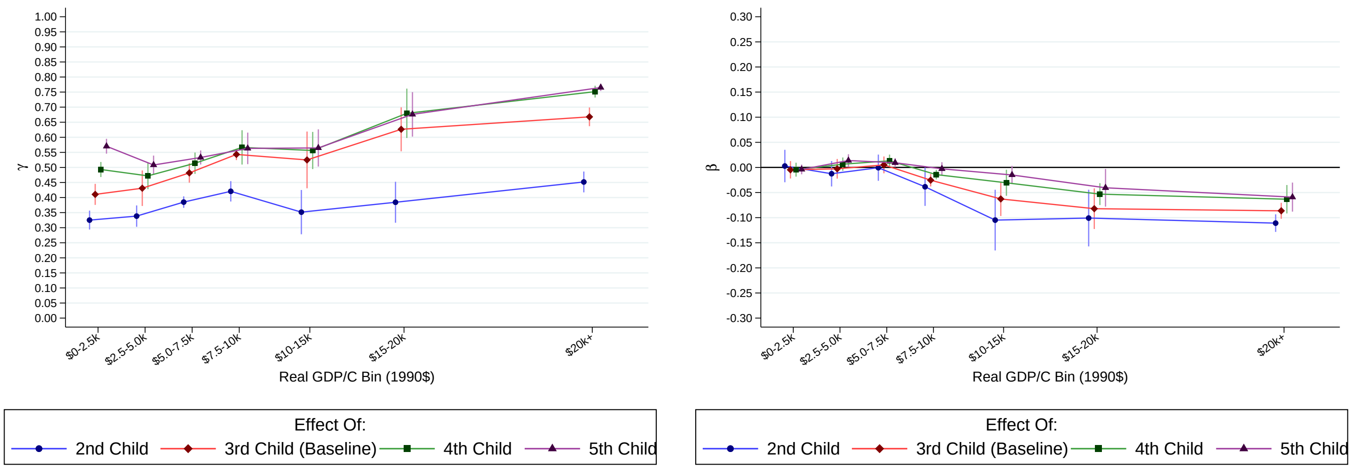
(a) Twin IV (b) Same Sex IV



Notes: This figure compares our baseline second-stage twin and same sex results to results that restrict to surveys that match well to ILO female labor supply statistics. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 8: Twin IV estimates at different family sizes

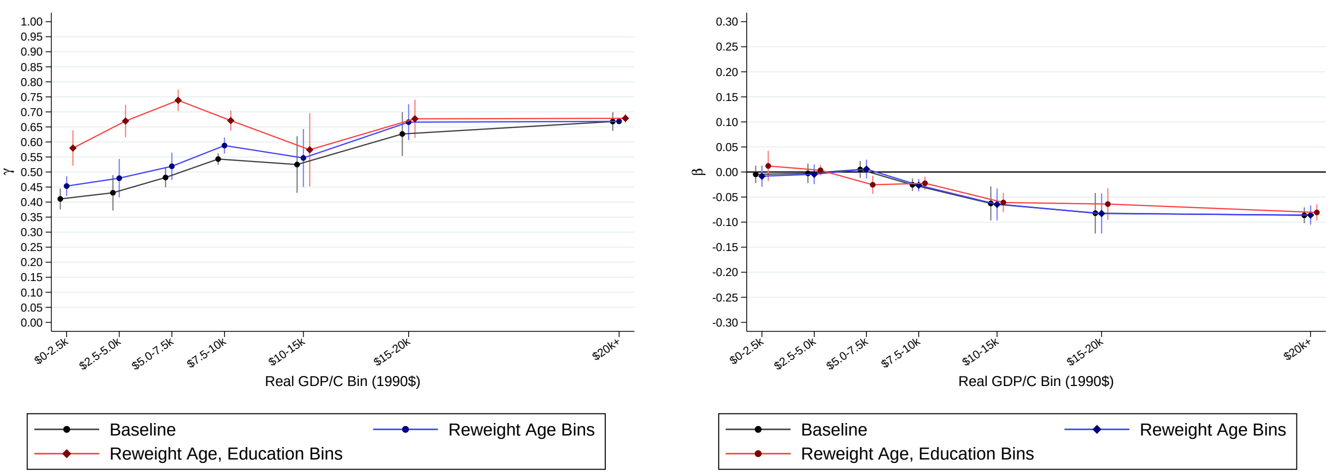
(a) First Stage: Additional Birth (b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates by the size of the family when the twins were born. For example, the line labeled "2nd child" includes mothers with at least one child and where twins are the first and second child born. The line labeled "3rd child" is our baseline. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

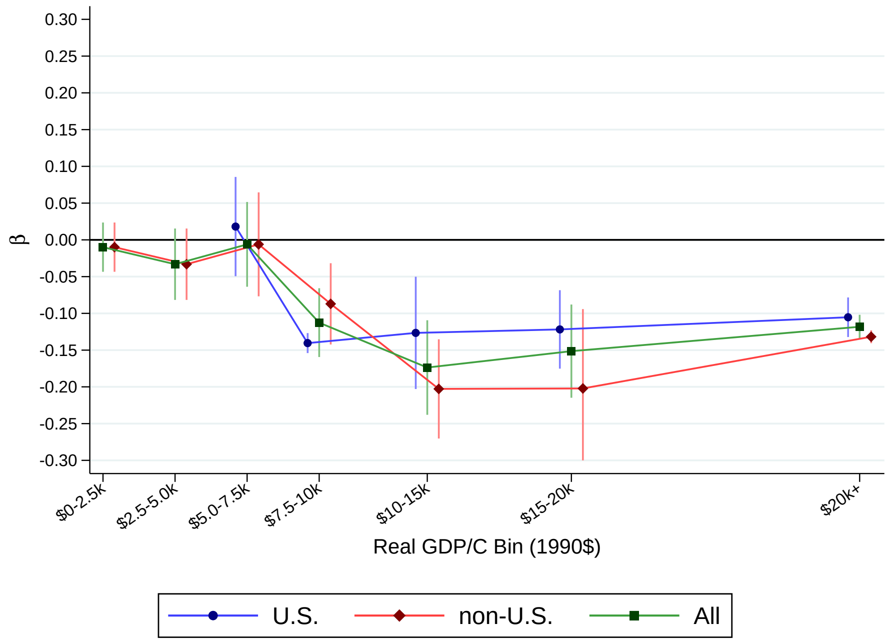
Figure 9: Reweight covariates to 1980 U.S. compliers, twin IV

(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



Notes: This figure adjusts for changes in the twin IV complier population by reweighting IV estimates to the U.S. 1980 complier profile (see Angrist and Fernandez-Val (2010) and Bisbee et al. (2017)). The sample is restricted to the set of mothers who report education. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

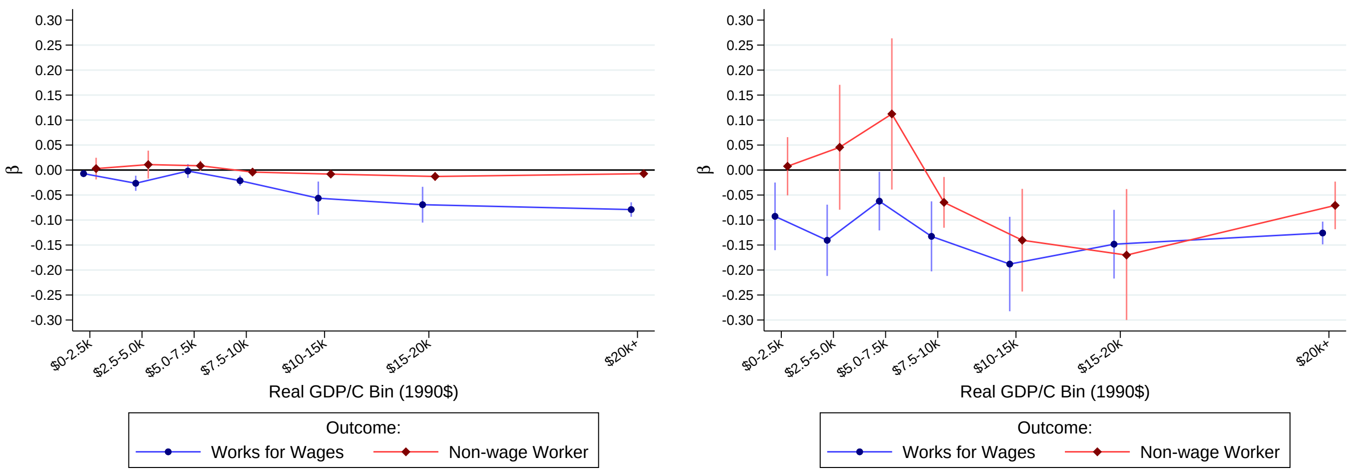
Figure 10: Second Stage twin IV estimates, rescaled by the complier-control outcome mean



Notes: This figure rescales the baseline, second-stage twin IV estimates by the complier-control mean of mothers' labor force status. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 11: Twin IV estimates by class of worker

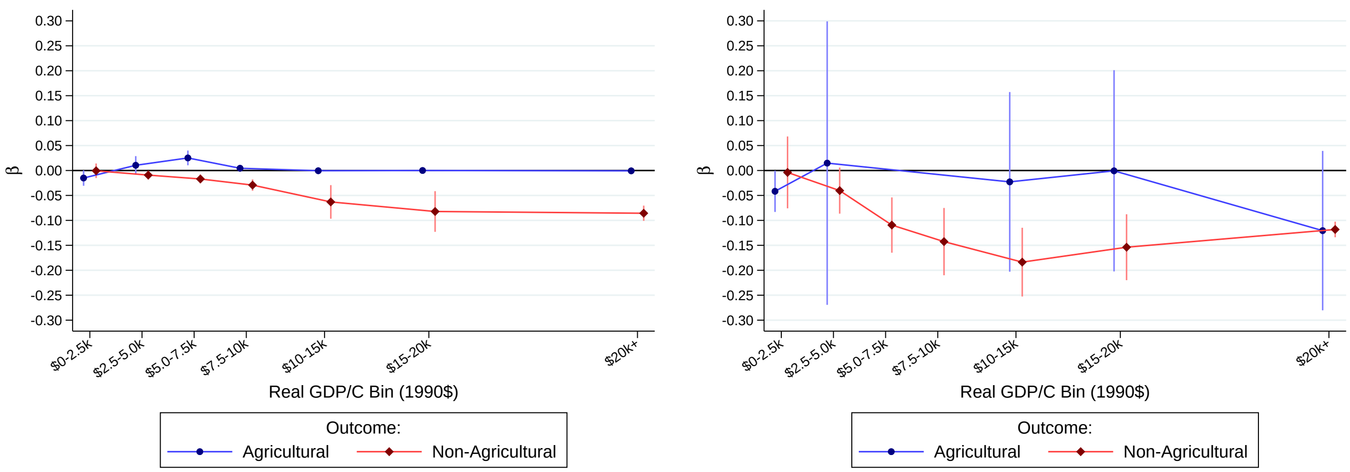
(a) Unscaled (b) Scaled by Complier-Control Mean



Notes: This figure displays second-stage twin IV estimates, unscaled (panel A) and scaled by the complier-control mean (panel B). The outcome for the blue line (circles) is an indicator of whether the mother works for wages. The outcome for the red line (triangles) is an indicator of whether a mother works but not for wages. The sample is restricted to the set of mothers with nonmissing data on wage work and held constant across panels. We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 12: Twin IV estimates by agricultural occupation

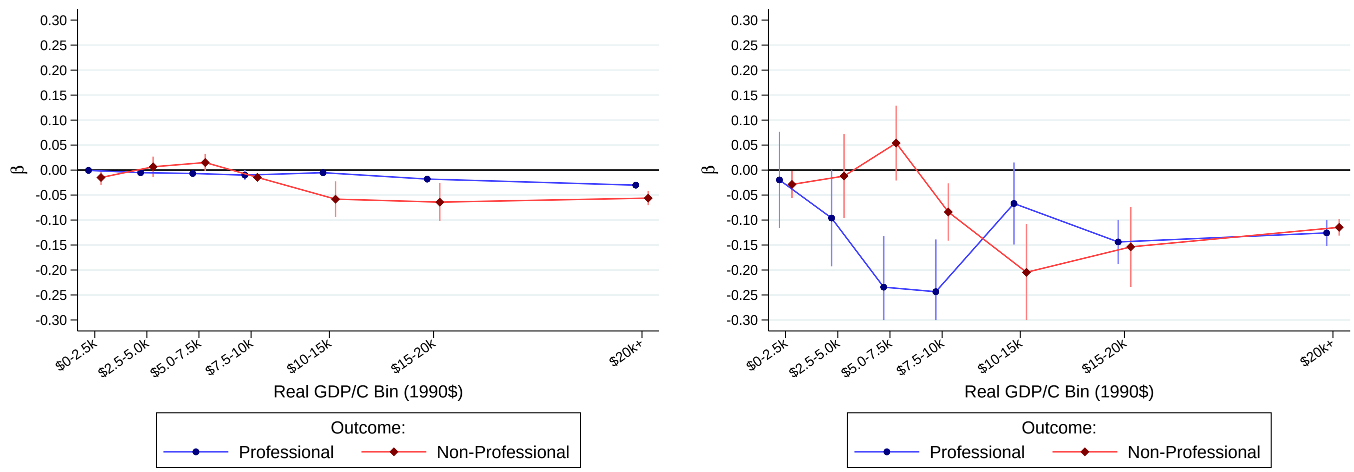
(a) Unscaled (b) Scaled by complier-control mean



Notes: This figure displays second-stage twin IV estimates unscaled (panel A) and scaled by the complier-control mean (panel B). The outcome for the blue line (circles) is an indicator of whether the mother works in agriculture (defined as a farm laborer, tenant, manager, or owner). The outcome for the red line (triangles) is an indicator of whether a mother works but not in agriculture. The sample is restricted to the set of mothers with nonmissing data on occupation and held constant across panels. We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure. The point estimates for $5-7.5kand7.5$-10k in the agricultural subsample of panel B are not displayed because the denominator is very small, and the point estimate does not fit on the figure.

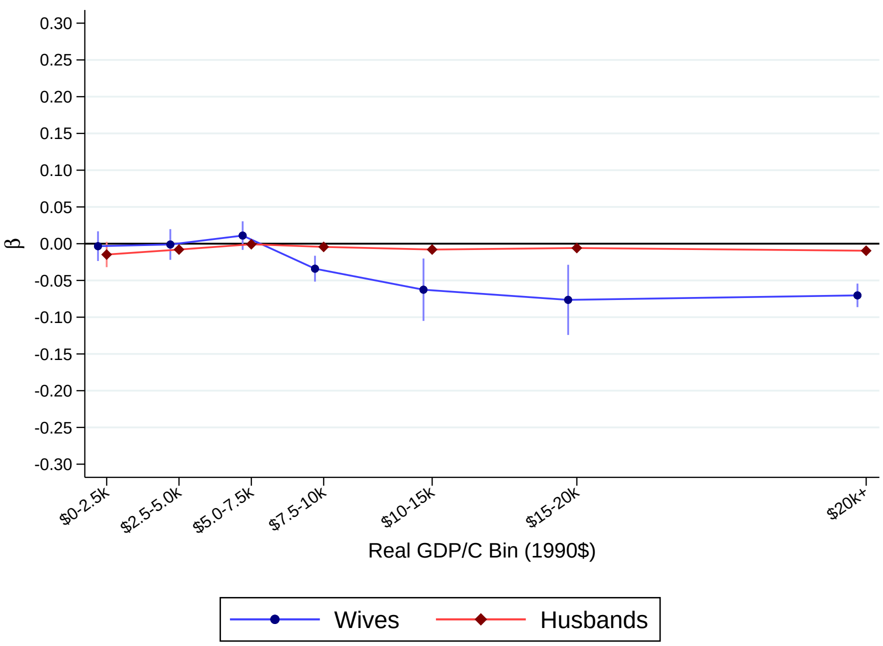
Figure 13: Twin IV estimates by professional occupation

(a) Unscaled | (b) Scaled by complier-control mean



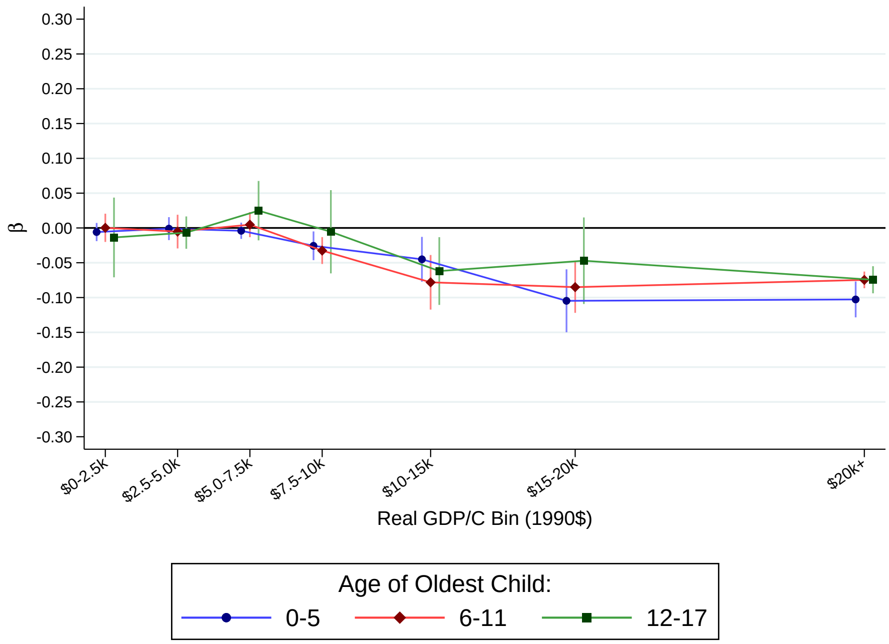
Notes: This figure displays second-stage twin IV estimates unscaled (panel A) and scaled by the complier-control mean (panel B). The outcome for the blue line (circles) is an indicator of whether the mother works in a professional occupation. The outcome for the red line (triangles) is an indicator of whether a mother works but not in a professional occupation. The sample is restricted to the set of mothers with nonmissing data on occupation and held constant across panels. We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 14: Twin IV estimates for fathers



Notes: This figure displays second-stage twin IV estimates for fathers living in the same household as mothers. The blue line (circles) shows our baseline mother labor supply estimates, restricted to those where the father also lives in the same household. The red line (trianges) shows the analaogous estimate for fathers. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure 15: Twin IV estimates by the age of the oldest child



Notes: This figure displays second-stage twin IV estimates, stratified by the age of the oldest child in the household. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Table A1: Country-year statistics and estimates

0 to 2,500 real GDP/Capita bin

	Year (num. samples)	Source	N	Percent of bin	Mean real GDP/C	In labor force	3 or more children	2nd child is twin	Educa-tion?	Birth Quarter?	OLS	FS Twin IV	2S Twin IV	FS Same gender IV	2S Same gender IV
Pooled	215		9,709,322			43.20%	57.22%	1.28%			-0.022***	0.411***	-0.005	0.028***	-0.046**
											(0.005)	(0.018)	(0.009)	(0.007)	(0.019)
Bangladesh	1991	IPUMS-I	702,804	7.24%	647	4.12%	62.25%	1.09%	X		-0.026***	0.429***	0.023***	0.027***	-0.027
											(0.001)	(0.003)	(0.006)	(0.001)	(0.017)
Bangladesh	1993	DHS	3,703	0.04%	684	17.17%	59.64%	0.39%	X	X	-0.069***	0.550***	0.003	0.044***	-0.231
											(0.015)	(0.059)	(0.199)	(0.015)	(0.288)
Bangladesh	1996	DHS	3,272	0.03%	749	38.57%	57.34%	0.41%	X	X	-0.048**	0.370***	-0.872***	0.050***	0.237
											(0.020)	(0.071)	(0.285)	(0.016)	(0.366)
Bangladesh	1999	DHS	3,590	0.04%	827	22.70%	54.81%	0.46%	X	X	-0.068***	0.515***	0.071	0.067***	-0.274
											(0.017)	(0.036)	(0.239)	(0.015)	(0.230)
Bangladesh	2001	IPUMS-I	754,996	7.78%	885	8.22%	50.91%	1.01%	X		-0.022***	0.495***	0.084***	0.037***	-0.192***
											(0.001)	(0.002)	(0.008)	(0.001)	(0.018)
Bangladesh	2004	DHS	3,825	0.04%	991	23.53%	52.07%	0.51%	X	X	-0.060***	0.574***	-0.054	0.054***	-0.324
											(0.017)	(0.055)	(0.159)	(0.016)	(0.295)
Bangladesh	2007	DHS	3,438	0.04%	1125	34.33%	46.18%	0.64%	X	X	-0.099***	0.414***	0.438	0.091***	-0.149
											(0.021)	(0.049)	(0.309)	(0.018)	(0.207)
Bangladesh	2011	IPUMS-I	466,242	4.80%	1276	5.80%	40.00%	0.65%	X		-0.021***	0.623***	0.003	0.067***	-0.023**
											(0.001)	(0.003)	(0.007)	(0.001)	(0.010)
Bangladesh	2011	DHS	5,606	0.06%	1276	11.50%	40.73%	0.41%	X	X	-0.057***	0.624***	0.153	0.089***	0.137
											(0.010)	(0.035)	(0.155)	(0.014)	(0.111)
Benin	1996	DHS	1,620	0.02%	1195	92.19%	59.48%	1.00%	X	X	-0.025	0.318***	-0.588	0.019	1.100
											(0.016)	(0.083)	(0.458)	(0.020)	(1.436)
Benin	2001	DHS	1,741	0.02%	1302	92.12%	58.57%	1.21%	X	X	0.004	0.456***	0.171***	0.009	0.188
											(0.017)	(0.049)	(0.021)	(0.019)	(1.465)
Benin	2006	DHS	5,847	0.06%	1360	88.05%	60.97%	1.63%	X	X	-0.005	0.449***	-0.038	0.007	-0.413
											(0.011)	(0.031)	(0.080)	(0.011)	(1.450)
Bolivia	1992	IPUMS-I	33,935	0.35%	2265	44.33%	62.15%	0.78%	X		-0.043***	0.376***	-0.002	0.016***	0.237
											(0.006)	(0.014)	(0.081)	(0.005)	(0.345)
Bolivia	1994	DHS	2,391	0.02%	2354	57.92%	63.73%	0.87%	X	X	-0.058**	0.336***	-0.442	0.032*	0.777
											(0.026)	(0.059)	(0.326)	(0.019)	(0.855)
Brazil	1960	IPUMS-I	164,570	1.69%	2296	8.46%	67.95%	0.74%	X		-0.037***	0.331***	0.018	0.014***	0.024
											(0.002)	(0.006)	(0.025)	(0.002)	(0.096)
Brazzaville (Congo)	2005	DHS	1,651	0.02%	2091	70.71%	51.88%	1.47%	X	X	0.029	0.442***	0.106	-0.001	0.476
											(0.030)	(0.050)	(0.212)	(0.025)	(18.771)
Burkina Faso	1996	IPUMS-I	58,935	0.61%	885	76.89%	65.06%	1.59%	X	X	0.007*	0.301***	0.067	0.005	0.019
											(0.004)	(0.009)	(0.045)	(0.003)	(0.658)
Burkina Faso	2006	IPUMS-I	80,012	0.82%	1122	66.60%	61.90%	1.89%	X	X	0.032***	0.368***	-0.030	0.002	-2.061
											(0.004)	(0.007)	(0.033)	(0.003)	(3.243)
Burkina Faso	1993	DHS	1,982	0.02%	833	61.51%	65.03%	0.60%	X	X	-0.028	0.419***	-0.168	0.015	-0.917
											(0.031)	(0.053)	(0.399)	(0.019)	(1.970)
Burkina Faso	1998	DHS	1,870	0.02%	934	70.91%	61.65%	0.70%	X	X	0.009	0.436***	0.070	0.017	1.487
											(0.027)	(0.085)	(0.295)	(0.019)	(2.107)
Burkina Faso	2003	DHS	3,569	0.04%	1046	92.00%	61.80%	0.76%	X	X	0.038**	0.370***	-0.325	0.022	0.531
											(0.016)	(0.045)	(0.269)	(0.016)	(0.617)
Burkina Faso	2010	DHS	5,722	0.06%	1234	79.72%	62.72%	0.97%	X	X	0.000	0.454***	-0.029	0.012	0.468
											(0.015)	(0.038)	(0.132)	(0.011)	(1.061)
Cambodia	1998	IPUMS-I	65,026	0.67%	1183	82.72%	60.85%	0.62%	X		0.010***	0.380***	-0.140**	0.028***	0.073
											(0.004)	(0.013)	(0.055)	(0.003)	(0.108)
Cambodia	2000	DHS	3,705	0.04%	1325	72.32%	60.01%	0.39%	X	X	-0.005	0.430***	0.385*	0.034**	-0.238
											(0.020)	(0.047)	(0.202)	(0.016)	(0.506)
Cambodia	2005	DHS	3,619	0.04%	1929	64.46%	50.55%	0.53%	X	X	-0.078***	0.453***	0.060	0.056***	-0.353
											(0.022)	(0.062)	(0.272)	(0.017)	(0.347)
Cambodia	2008	IPUMS-I	63,509	0.65%	2316	87.61%	45.88%	0.88%	X		0.005*	0.547***	-0.041	0.048***	0.070
											(0.003)	(0.010)	(0.027)	(0.003)	(0.055)
Cambodia	2010	DHS	3,761	0.04%	2450	69.99%	41.95%	0.19%	X	X	-0.073***	0.439***	0.008	0.064***	-0.002
											(0.022)	(0.050)	(0.275)	(0.017)	(0.291)
Cambodia	2014	DHS	4,031	0.04%	2450	71.47%	38.55%	0.51%	X	X	-0.129***	0.667***	-0.127	0.042**	-0.313
											(0.021)	(0.039)	(0.211)	(0.017)	(0.424)
Cameroon	1976	IPUMS-I	32,831	0.34%	1058	49.06%	63.87%	2.20%	X	X	-0.025***	0.376***	-0.009	-0.008*	0.107
											(0.006)	(0.008)	(0.050)	(0.005)	(0.675)
Cameroon	1987	IPUMS-I	47,169	0.49%	1472	48.71%	66.16%	2.89%	X		-0.036***	0.377***	-0.023	0.003	-1.401
											(0.005)	(0.006)	(0.036)	(0.004)	(2.149)
Cameroon	1991	DHS	1,061	0.01%	1154	65.97%	71.65%	1.16%	X	X	-0.111***	0.377***	-0.696**	0.015	0.991
											(0.038)	(0.073)	(0.291)	(0.025)	(2.771)
Cameroon	1998	DHS	1,300	0.01%	1033	78.25%	64.50%	1.31%	X	X	0.004	0.364***	0.209	0.030	-0.304
											(0.028)	(0.065)	(0.203)	(0.023)	(0.803)
Cameroon	2004	DHS	2,434	0.03%	1139	71.28%	62.25%	1.13%	X	X	-0.022	0.446***	-0.017	-0.016	0.652
											(0.024)	(0.037)	(0.195)	(0.018)	(1.405)
Cameroon	2005	IPUMS-I	83,411	0.86%	1149	48.89%	68.04%	6.83%	X	X	-0.017***	0.470***	0.043***	-0.013***	-0.085
											(0.004)	(0.003)	(0.015)	(0.003)	(0.263)
Cameroon	2011	DHS	3,690	0.04%	1179	72.97%	62.72%	1.74%	X	X	-0.023	0.428***	0.124	0.017	0.518
											(0.020)	(0.030)	(0.122)	(0.015)	(1.097)
Canada	1871	NAPP	2,014	0.02%	1718	1.05%	71.89%	0.36%			-0.015	0.215***	0.125	0.022	-0.034
											(0.010)	(0.072)	(0.196)	(0.021)	(0.238)
Canada	1881	NAPP	178,949	1.84%	1955	2.20%	68.49%	0.69%		X	-0.013***	0.298***	-0.011	0.009***	0.036
											(0.001)	(0.006)	(0.013)	(0.002)	(0.082)
Canada	1891	NAPP	14,506	0.15%	2343	6.92%	66.94%	0.41%		X	0.003	0.338***	-0.042	0.002	2.140

											(0.006)	(0.032)	(0.110)	(0.008)	(9.979)
Central African Republic	1994	DHS	1,514	0.02%	568	83.35%	63.85%	0.45%	X	X	0.014	0.424***	0.186	-0.014	-1.296
											(0.023)	(0.097)	(0.234)	(0.022)	(2.499)
Chad	1996	DHS	2,348	0.02%	448	45.10%	69.36%	0.94%	X	X	-0.027	0.424***	0.542**	-0.041**	-0.648
											(0.029)	(0.059)	(0.241)	(0.016)	(0.590)
Chad	2004	DHS	1,874	0.02%	643	77.00%	68.21%	0.18%	X	X	0.009	0.465***	-0.035	-0.028	0.532
											(0.030)	(0.089)	(0.319)	(0.023)	(0.895)
China	1982	IPUMS-I	570,519	5.88%	1224	87.04%	48.10%	0.46%	X		-0.024***	0.566***	-0.024**	0.068***	-0.043***
											(0.001)	(0.006)	(0.012)	(0.001)	(0.013)
China	1990	IPUMS-I	614,197	6.33%	1955	89.33%	30.45%	0.85%	X	X	0.002*	0.712***	-0.023***	0.123***	-0.011*
											(0.001)	(0.002)	(0.006)	(0.001)	(0.006)
Comoros	1996	DHS	631	0.01%	625	43.90%	67.51%	1.27%	X	X	0.008	0.351***	0.849*	-0.036	-1.249
											(0.050)	(0.077)	(0.435)	(0.031)	(1.554)
Congo	2007	DHS	2,729	0.03%	240	74.84%	64.03%	1.12%	X	X	0.023	0.396***	-0.565**	0.020	-0.711
											(0.029)	(0.065)	(0.274)	(0.022)	(1.349)
Congo	2013	DHS	5,657	0.06%	260	77.06%	67.04%	1.07%	X	X	-0.012	0.309***	-0.103	-0.009	-0.085
											(0.020)	(0.048)	(0.275)	(0.014)	(1.818)
Denmark	1787	NAPP	24,456	0.25%	1274	2.54%	51.79%	0.53%			-0.015***	0.492***	0.010	0.012**	0.431
											(0.002)	(0.020)	(0.031)	(0.006)	(0.274)
Denmark	1801	NAPP	27,372	0.28%	1274	1.90%	52.85%	0.53%			-0.015***	0.455***	0.018	0.015***	-0.052
											(0.002)	(0.020)	(0.030)	(0.005)	(0.114)
Dominican Republic	1981	IPUMS-I	22,567	0.23%	2368	28.17%	65.52%	3.23%	X		-0.097***	0.357***	-0.013	0.013*	0.809
											(0.009)	(0.010)	(0.055)	(0.007)	(0.719)
Egypt	1986	IPUMS-I	394,535	4.06%	2449	10.27%	66.30%	1.91%	X	X	-0.081***	0.326***	-0.020*	0.025***	-0.044
											(0.001)	(0.003)	(0.010)	(0.001)	(0.038)
El Salvador	1992	IPUMS-I	27,018	0.28%	2285	31.80%	56.70%	1.36%	X	X	-0.148***	0.436***	0.063	0.019***	0.013
											(0.006)	(0.011)	(0.058)	(0.006)	(0.297)
Ethiopia	2000	DHS	3,712	0.04%	562	57.07%	64.46%	0.25%	X	X	-0.071**	0.445***	0.086	0.002	-7.698
											(0.028)	(0.145)	(0.483)	(0.017)	(62.193)
Ethiopia	2005	DHS	3,747	0.04%	672	25.51%	71.56%	0.45%	X	X	-0.112***	0.380***	0.173	0.001	2.517
											(0.024)	(0.058)	(0.405)	(0.016)	(34.880)
Ethiopia	2007	IPUMS-I	73,510	0.76%	771	74.61%	69.55%	0.98%	X		-0.009**	0.341***	0.048	0.004	-0.233
											(0.004)	(0.008)	(0.047)	(0.003)	(0.789)
Ethiopia	2011	DHS	4,461	0.05%	935	37.12%	67.83%	0.93%	X	X	-0.045	0.244***	0.311	0.011	-1.494
											(0.028)	(0.070)	(0.533)	(0.018)	(3.104)
Germany	1819	NAPP	2,062	0.02%	986	3.72%	55.01%	0.72%		X	-0.022	0.464***	0.116	0.029	0.152
											(0.014)	(0.061)	(0.186)	(0.028)	(0.406)
Ghana	1993	DHS	1,355	0.01%	1133	85.39%	56.68%	1.18%	X	X	-0.033	0.459***	-0.111	0.017	1.125
											(0.022)	(0.063)	(0.199)	(0.023)	(1.936)
Ghana	1998	DHS	1,153	0.01%	1282	86.53%	53.87%	1.33%	X	X	-0.010	0.452***	0.140	0.003	-6.209
											(0.023)	(0.067)	(0.142)	(0.026)	(48.365)
Ghana	2000	IPUMS-I	72,394	0.75%	1353	86.66%	55.59%	2.98%	X		0.008***	0.446***	-0.002	0.003	-0.721
											(0.003)	(0.005)	(0.017)	(0.003)	(1.364)
Ghana	2003	DHS	1,316	0.01%	1471	90.44%	52.27%	0.91%	X	X	-0.024	0.440***	-0.106	-0.026	0.675
											(0.021)	(0.091)	(0.202)	(0.026)	(0.961)
Ghana	2008	DHS	1,043	0.01%	1767	90.28%	49.86%	1.26%	X	X	-0.003	0.507***	-0.237	-0.003	11.178
											(0.023)	(0.071)	(0.235)	(0.030)	(121.874)
Ghana	2010	IPUMS-I	99,670	1.03%	1922	86.06%	55.57%	2.98%	X		-0.008***	0.434***	-0.014	0.011***	-0.147
											(0.003)	(0.004)	(0.015)	(0.003)	(0.205)
Ghana	2014	DHS	2,053	0.02%	1922	83.15%	52.13%	1.80%	X	X	-0.028	0.511***	-0.194	0.028	-0.012
											(0.022)	(0.050)	(0.183)	(0.023)	(0.710)
Guinea	1983	IPUMS-I	20,684	0.21%	539	49.68%	52.29%	3.55%	X		-0.030***	0.456***	-0.100**	0.004	-2.530
											(0.008)	(0.008)	(0.041)	(0.006)	(3.992)
Guinea	1996	IPUMS-I	37,807	0.39%	555	71.66%	61.92%	2.36%	X	X	0.006	0.376***	-0.095**	-0.000	-33.935
											(0.005)	(0.008)	(0.042)	(0.004)	(2071.208)
Guinea	1999	DHS	2,027	0.02%	587	84.48%	62.19%	1.21%	X	X	0.008	0.355***	-0.002	0.031*	-0.777
											(0.019)	(0.051)	(0.199)	(0.018)	(0.706)
Guinea	2005	DHS	2,243	0.02%	615	87.03%	60.74%	1.36%	X	X	0.021	0.486***	0.003	0.015	0.685
											(0.019)	(0.040)	(0.128)	(0.019)	(1.328)
Haiti	1971	IPUMS-I	18,141	0.19%	966	69.99%	58.58%	1.19%	X		-0.045***	0.449***	-0.023	0.002	-8.420
											(0.008)	(0.016)	(0.071)	(0.007)	(36.650)
Haiti	1982	IPUMS-I	4,195	0.04%	1171	54.99%	53.92%	1.74%	X		-0.090***	0.422***	-0.088	0.003	0.429
											(0.017)	(0.028)	(0.139)	(0.014)	(6.100)
Haiti	1994	DHS	1,051	0.01%	800	45.55%	60.35%	1.34%	X	X	-0.007	0.295***	-0.317	0.029	-0.159
											(0.036)	(0.069)	(0.404)	(0.026)	(1.067)
Haiti	2000	DHS	2,002	0.02%	746	55.17%	58.43%	0.43%	X	X	-0.104**	0.568***	-0.539**	0.045	0.580
											(0.044)	(0.061)	(0.214)	(0.035)	(1.046)
Haiti	2003	IPUMS-I	29,838	0.31%	708	53.34%	55.15%	1.58%	X		-0.045***	0.444***	-0.064	0.011**	0.459
											(0.006)	(0.010)	(0.052)	(0.005)	(0.556)
Haiti	2005	DHS	1,932	0.02%	690	54.14%	53.88%	0.61%	X	X	-0.048	0.498***	0.006	0.010	0.292
											(0.032)	(0.045)	(0.292)	(0.025)	(3.036)
Honduras	2005	DHS	5,219	0.05%	2113	39.43%	54.28%	0.43%	X	X	-0.146***	0.522***	-0.082	0.034**	0.222
											(0.018)	(0.060)	(0.177)	(0.014)	(0.501)
India	1983	IPUMS-I	41,910	0.43%	1026	34.38%	61.60%	0.55%	X		-0.020***	0.375***	-0.080	0.013**	-0.297
											(0.007)	(0.027)	(0.109)	(0.005)	(0.469)
India	1987	IPUMS-I	45,884	0.47%	1166	33.89%	60.60%	0.55%	X		-0.003	0.432***	0.004	0.025***	-0.792***
											(0.006)	(0.018)	(0.137)	(0.005)	(0.281)
India	1992	DHS	33,928	0.35%	1377	32.68%	61.17%	0.36%	X	X	-0.026***	0.429***	-0.069	0.011**	-0.178
											(0.007)	(0.021)	(0.112)	(0.006)	(0.549)
India	1993	IPUMS-I	39,508	0.41%	1430	39.47%	56.22%	0.43%	X		-0.003	0.481***	-0.145	0.027***	-0.425*
											(0.007)	(0.020)	(0.095)	(0.006)	(0.244)
India	1998	DHS	34,272	0.35%	1755	37.25%	56.88%	0.32%	X	X	-0.027***	0.458***	-0.060	0.025***	0.385
											(0.007)	(0.021)	(0.113)	(0.006)	(0.262)
India	1999	IPUMS-I	41,373	0.43%	1819	38.43%	54.55%	0.44%	X		0.020***	0.491***	-0.025	0.037***	0.027
											(0.007)	(0.020)	(0.091)	(0.006)	(0.182)
India	2004	IPUMS-I	41,618	0.43%	2315	42.27%	48.98%	0.35%	X		-0.006	0.545***	-0.093	0.027***	0.103

											(0.008)	(0.021)	(0.096)	(0.007)	(0.270)
India	2005	DHS	32,970	0.34%	2457	37.80%	52.02%	0.45%	X	X	-0.006	0.511***	0.012	0.033***	-0.185
											(0.007)	(0.020)	(0.096)	(0.006)	(0.204)
Indonesia	1971	IPUMS-I	37,598	0.39%	1294	32.10%	67.41%	0.31%	X		-0.057***	0.319***	0.159	-0.006	1.362
											(0.012)	(0.024)	(0.193)	(0.009)	(2.917)
Indonesia	1976	IPUMS-I	16,776	0.17%	1635	46.35%	66.21%	0.60%	X	X	-0.110***	0.397***	-0.276*	0.011	-1.580
											(0.012)	(0.030)	(0.166)	(0.008)	(1.410)
Indonesia	1980	IPUMS-I	436,461	4.50%	1833	32.85%	62.30%	0.69%	X	X	-0.069***	0.385***	-0.088***	0.010***	0.159
											(0.002)	(0.004)	(0.022)	(0.001)	(0.144)
Iraq	1997	IPUMS-I	106,406	1.10%	1062	7.43%	72.06%	2.24%	X	X	-0.034***	0.295***	-0.047***	0.010***	0.179
											(0.002)	(0.005)	(0.017)	(0.002)	(0.168)
Ivory Coast	1994	DHS	2,193	0.02%	1312	78.37%	60.69%	0.88%	X	X	-0.004	0.438***	-0.233	-0.006	5.826
											(0.023)	(0.073)	(0.232)	(0.018)	(17.379)
Ivory Coast	1998	DHS	589	0.01%	1377	85.12%	54.32%	1.30%	X	X	-0.039	0.284**	0.250	-0.047	1.376
											(0.035)	(0.121)	(0.212)	(0.039)	(1.318)
Ivory Coast	2011	DHS	2,500	0.03%	1195	73.83%	52.39%	1.46%	X	X	0.011	0.523***	0.015	0.005	-2.805
											(0.027)	(0.048)	(0.149)	(0.021)	(11.767)
Kenya	1989	IPUMS-I	61,498	0.63%	1080	78.79%	70.40%	1.96%	X		0.012***	0.295***	0.085**	-0.000	12.790
											(0.004)	(0.007)	(0.038)	(0.003)	(200.978)
Kenya	1993	DHS	2,362	0.02%	1051	56.89%	69.76%	0.75%	X	X	0.007	0.265***	0.355	0.017	-0.081
											(0.030)	(0.058)	(0.464)	(0.018)	(1.382)
Kenya	1998	DHS	2,229	0.02%	1029	60.69%	61.11%	0.95%	X	X	-0.044	0.399***	0.283	0.034*	-0.645
											(0.028)	(0.070)	(0.286)	(0.020)	(0.764)
Kenya	1999	IPUMS-I	79,020	0.81%	1026	79.86%	61.55%	1.45%	X		-0.021***	0.391***	0.093***	0.008***	-0.306
											(0.003)	(0.008)	(0.028)	(0.003)	(0.372)
Kenya	2003	DHS	2,158	0.02%	1032	65.67%	61.92%	1.26%	X	X	-0.065**	0.486***	-0.060	0.036*	-1.103
											(0.026)	(0.048)	(0.201)	(0.020)	(0.841)
Kenya	2008	DHS	2,350	0.02%	1116	64.39%	60.68%	0.44%	X	X	-0.085***	0.484***	-0.124	0.005	-3.896
											(0.032)	(0.057)	(0.261)	(0.029)	(23.857)
Kenya	2009	IPUMS-I	224,868	2.32%	1121	78.80%	61.42%	1.46%	X		-0.019***	0.403***	0.009	0.012***	-0.174
											(0.002)	(0.004)	(0.018)	(0.002)	(0.147)
Kenya	2014	DHS	4,289	0.04%	1141	69.95%	56.90%	1.18%	X	X	-0.126***	0.429***	-0.060	0.027	0.583
											(0.021)	(0.041)	(0.204)	(0.019)	(0.803)
Kyrgyz Republic	1999	IPUMS-I	29,660	0.31%	2107	78.65%	52.69%	0.89%	X	X	0.019***	0.455***	-0.050	0.061***	-0.060
											(0.005)	(0.014)	(0.058)	(0.005)	(0.078)
Lesotho	2004	DHS	1,296	0.01%	1669	40.91%	46.20%	0.49%	X	X	-0.152***	0.663***	-0.549***	-0.016	-0.327
											(0.036)	(0.073)	(0.111)	(0.028)	(1.996)
Liberia	2007	DHS	1,715	0.02%	778	69.13%	49.48%	1.78%	X	X	0.004	0.527***	0.106	-0.016	0.750
											(0.030)	(0.063)	(0.204)	(0.027)	(2.105)
Liberia	2008	IPUMS-I	14,661	0.15%	802	57.85%	56.74%	2.38%	X		0.023**	0.430***	-0.091	0.006	2.394
											(0.009)	(0.012)	(0.063)	(0.007)	(3.077)
Madagascar	1992	DHS	1,575	0.02%	722	80.53%	65.84%	0.60%	X	X	-0.024	0.330***	0.615***	-0.036*	0.888
											(0.024)	(0.055)	(0.107)	(0.022)	(0.801)
Madagascar	1997	DHS	1,836	0.02%	676	82.01%	61.43%	0.61%	X	X	0.048**	0.216**	-1.461	0.053***	0.184
											(0.023)	(0.096)	(1.028)	(0.020)	(0.353)
Madagascar	2003	DHS	2,066	0.02%	671	84.60%	59.17%	0.46%	X	X	0.039	0.342***	0.014	-0.020	-0.664
											(0.027)	(0.092)	(0.279)	(0.026)	(1.411)
Madagascar	2008	DHS	4,664	0.05%	702	92.02%	62.48%	0.80%	X	X	0.028**	0.450***	0.193***	0.057***	-0.134
											(0.012)	(0.047)	(0.019)	(0.015)	(0.173)
Malawi	1987	IPUMS-I	42,881	0.44%	567	80.11%	58.04%	1.52%	X		-0.035***	0.394***	-0.043	0.007*	0.715
											(0.004)	(0.010)	(0.041)	(0.004)	(0.713)
Malawi	1992	DHS	1,389	0.01%	536	26.16%	61.07%	1.02%	X	X	0.070**	0.431***	0.037	0.053**	0.063
											(0.031)	(0.062)	(0.282)	(0.025)	(0.500)
Malawi	1998	IPUMS-I	51,847	0.53%	602	84.41%	56.42%	1.92%	X		-0.008**	0.407***	-0.007	0.003	0.088
											(0.004)	(0.008)	(0.028)	(0.004)	(1.020)
Malawi	2000	DHS	3,803	0.04%	598	59.61%	58.24%	0.94%	X	X	-0.057**	0.483***	-0.015	0.036**	-0.375
											(0.023)	(0.039)	(0.199)	(0.015)	(0.531)
Malawi	2004	DHS	3,989	0.04%	587	58.96%	57.81%	1.28%	X	X	0.020	0.307***	0.291	-0.030**	0.871
											(0.024)	(0.063)	(0.305)	(0.015)	(0.753)
Malawi	2008	IPUMS-I	87,562	0.90%	662	77.93%	60.08%	1.66%	X	X	-0.006*	0.383***	-0.041	0.005*	0.644
											(0.003)	(0.007)	(0.029)	(0.003)	(0.719)
Malawi	2010	DHS	8,215	0.08%	728	59.56%	62.75%	1.16%	X	X	-0.023	0.393***	-0.140	0.021*	-0.950
											(0.018)	(0.029)	(0.172)	(0.011)	(0.825)
Malaysia	1970	IPUMS-I	9,724	0.10%	2126	34.04%	73.17%	1.14%	X		-0.059***	0.242***	-0.281	0.018**	0.378
											(0.012)	(0.019)	(0.179)	(0.008)	(0.571)
Mali	1987	IPUMS-I	40,230	0.41%	713	51.31%	63.66%	1.48%	X		0.002	0.349***	-0.112*	0.004	1.016
											(0.006)	(0.010)	(0.059)	(0.004)	(1.455)
Mali	1995	DHS	3,161	0.03%	796	55.32%	69.18%	0.88%	X	X	0.020	0.411***	-0.125	-0.002	-0.749
											(0.024)	(0.035)	(0.241)	(0.014)	(11.245)
Mali	1998	IPUMS-I	49,792	0.51%	841	39.60%	67.52%	2.44%	X		-0.003	0.292***	-0.068	0.007**	0.759
											(0.005)	(0.007)	(0.048)	(0.004)	(0.727)
Mali	2001	DHS	4,067	0.04%	892	65.37%	66.39%	0.50%	X	X	-0.017	0.363***	-0.367	-0.004	4.435
											(0.024)	(0.065)	(0.312)	(0.015)	(18.555)
Mali	2006	DHS	4,623	0.05%	984	63.72%	67.01%	0.87%	X	X	-0.032	0.336***	-0.265	0.007	-1.028
											(0.024)	(0.057)	(0.257)	(0.014)	(3.669)
Mali	2009	IPUMS-I	75,084	0.77%	1036	39.67%	69.27%	2.64%	X	X	0.003	0.294***	0.065*	0.008***	-0.419
											(0.005)	(0.005)	(0.038)	(0.003)	(0.470)
Mali	2012	DHS	3,843	0.04%	1059	45.50%	72.10%	0.86%	X	X	-0.065***	0.331***	-0.545**	0.026*	-0.966
											(0.024)	(0.043)	(0.232)	(0.014)	(0.826)
Mongolia	2000	IPUMS-I	14,378	0.15%	1055	79.34%	40.85%	0.62%	X	X	-0.015**	0.544***	-0.063	0.042***	0.122
											(0.007)	(0.024)	(0.083)	(0.007)	(0.162)
Morocco	1982	IPUMS-I	53,186	0.55%	2261	11.71%	71.54%	1.68%	X		-0.067***	0.324***	-0.073**	0.005	-0.645
											(0.004)	(0.007)	(0.031)	(0.003)	(0.727)
Mozambique	1997	IPUMS-I	82,358	0.85%	1311	69.66%	56.63%	1.52%	X		-0.012***	0.413***	-0.105***	0.003	1.183
											(0.004)	(0.007)	(0.033)	(0.003)	(1.517)
Mozambique	1997	DHS	2,320	0.02%	1311	64.49%	54.93%	0.72%	X	X	0.005	0.383***	-0.228	-0.028	-0.063

											(0.042)	(0.081)	(0.355)	(0.036)	(1.293)
Mozambique	2003	DHS	3,453	0.04%	1849	78.92%	61.75%	1.19%	X	X	0.006	0.410***	0.039	0.008	-0.489
											(0.021)	(0.048)	(0.185)	(0.018)	(2.518)
Mozambique	2007	IPUMS-I	121,872	1.26%	2284	71.73%	63.23%	1.72%	X		0.041***	0.348***	-0.068**	0.004	0.149
											(0.003)	(0.005)	(0.029)	(0.002)	(0.725)
Nepal	1996	DHS	3,299	0.03%	928	79.25%	62.35%	0.26%	X	X	0.017	0.495***	-0.113	-0.006	2.601
											(0.019)	(0.064)	(0.353)	(0.015)	(6.605)
Nepal	2001	DHS	3,511	0.04%	997	84.39%	60.44%	0.33%	X	X	0.013	0.287***	-0.842	0.030**	0.329
											(0.016)	(0.052)	(0.563)	(0.015)	(0.466)
Nepal	2006	DHS	3,251	0.03%	1079	72.41%	51.37%	0.44%	X	X	0.014	0.242**	-1.252	0.010	0.519
											(0.022)	(0.107)	(1.468)	(0.018)	(2.088)
Nicaragua	1995	IPUMS-I	27,148	0.28%	1332	34.89%	63.84%	1.97%	X	X	-0.128***	0.361***	-0.007	0.020***	0.358
											(0.007)	(0.010)	(0.057)	(0.005)	(0.305)
Nicaragua	1998	DHS	3,733	0.04%	1445	39.38%	59.85%	0.61%	X	X	-0.104***	0.381***	0.061	0.015	0.519
											(0.020)	(0.040)	(0.271)	(0.016)	(1.404)
Nicaragua	2001	DHS	3,278	0.03%	1576	41.19%	56.90%	0.72%	X	X	-0.177***	0.412***	0.529**	0.050***	-0.228
											(0.022)	(0.040)	(0.246)	(0.018)	(0.396)
Nicaragua	2005	IPUMS-I	29,130	0.30%	1644	33.91%	51.65%	1.54%	X	X	-0.117***	0.469***	0.039	0.026***	-0.114
											(0.006)	(0.010)	(0.049)	(0.005)	(0.212)
Niger	1992	DHS	2,049	0.02%	511	45.11%	64.81%	0.49%	X	X	-0.060**	0.574***	-0.200	0.012	2.434
											(0.030)	(0.066)	(0.296)	(0.020)	(4.475)
Niger	1998	DHS	2,304	0.02%	455	54.60%	65.48%	0.61%	X	X	-0.055*	0.362***	-0.252	-0.037**	-0.834
											(0.028)	(0.080)	(0.418)	(0.017)	(0.694)
Niger	2006	DHS	3,095	0.03%	491	39.36%	67.79%	0.58%	X	X	-0.051*	0.330***	0.236	-0.020	-0.076
											(0.027)	(0.093)	(0.393)	(0.016)	(1.028)
Niger	2012	DHS	4,520	0.05%	519	23.72%	74.65%	0.88%	X	X	-0.071***	0.193***	0.056	0.015	-2.204
											(0.019)	(0.050)	(0.427)	(0.012)	(1.976)
Nigeria	1990	DHS	2,644	0.03%	1057	70.77%	66.70%	0.69%	X	X	0.037	0.460***	0.497***	0.048**	-0.324
											(0.028)	(0.063)	(0.130)	(0.022)	(0.495)
Nigeria	2003	DHS	1,813	0.02%	1350	66.28%	65.05%	0.84%	X	X	-0.005	0.480***	-0.821***	-0.005	-0.085
											(0.036)	(0.079)	(0.293)	(0.023)	(5.258)
Nigeria	2006	IPUMS-I	4,789	0.05%	1595	46.38%	59.52%	1.83%	X		-0.044**	0.458***	-0.136	0.017	-1.771
											(0.018)	(0.026)	(0.135)	(0.015)	(1.795)
Nigeria	2007	IPUMS-I	4,248	0.04%	1664	51.63%	63.10%	1.91%	X		-0.057**	0.441***	-0.151	0.010	-4.711
											(0.024)	(0.025)	(0.155)	(0.021)	(9.716)
Nigeria	2008	IPUMS-I	5,971	0.06%	1723	56.77%	65.62%	2.22%	X		-0.011	0.417***	-0.163	0.013	-3.923
											(0.018)	(0.029)	(0.145)	(0.014)	(4.459)
Nigeria	2008	DHS	9,291	0.10%	1723	68.66%	65.01%	1.09%	X	X	-0.028**	0.345***	0.028	0.002	-1.073
											(0.013)	(0.026)	(0.136)	(0.009)	(5.887)
Nigeria	2009	IPUMS-I	3,151	0.03%	1790	47.03%	65.56%	1.44%	X		-0.024	0.352***	-0.355	0.016	-0.253
											(0.025)	(0.042)	(0.254)	(0.020)	(1.416)
Nigeria	2010	IPUMS-I	4,028	0.04%	1876	59.05%	61.82%	1.67%	X	X	-0.052***	0.369***	0.079	-0.007	6.533
											(0.020)	(0.031)	(0.171)	(0.016)	(16.018)
Nigeria	2013	DHS	10,596	0.11%	1876	71.30%	67.72%	0.84%	X	X	0.004	0.379***	-0.086	-0.002	5.079
											(0.013)	(0.030)	(0.150)	(0.009)	(18.584)
Norway	1801	NAPP	25,820	0.27%	801	2.13%	56.23%	0.54%			-0.019***	0.443***	0.064	0.004	-0.497
											(0.002)	(0.022)	(0.042)	(0.005)	(0.804)
Norway	1865	NAPP	53,059	0.55%	1269	1.19%	60.23%	0.60%			-0.011***	0.396***	0.001	0.003	0.318
											(0.001)	(0.015)	(0.016)	(0.004)	(0.490)
Norway	1875	NAPP	17,956	0.18%	1520	3.42%	58.76%	0.68%			-0.039***	0.238***	-0.157**	-0.002	2.372
											(0.014)	(0.083)	(0.074)	(0.021)	(23.956)
Norway	1900	NAPP	68,771	0.71%	1880	11.55%	62.84%	0.71%			-0.041***	0.361***	-0.012	0.004	1.405
											(0.003)	(0.011)	(0.040)	(0.003)	(1.265)
Norway	1910	NAPP	75,194	0.77%	2210	9.16%	64.81%	0.90%		X	-0.037***	0.347***	-0.009	0.009***	0.257
											(0.003)	(0.010)	(0.032)	(0.003)	(0.257)
Pakistan	1973	IPUMS-I	76,747	0.79%	957	5.06%	68.01%	1.25%	X		-0.009***	0.340***	-0.016	0.005	-0.278
											(0.002)	(0.008)	(0.022)	(0.003)	(0.446)
Pakistan	1990	DHS	2,757	0.03%	1601	16.53%	76.22%	1.08%	X	X	-0.041	0.241***	-0.401**	0.005	-3.680
											(0.025)	(0.047)	(0.191)	(0.018)	(14.931)
Pakistan	2006	DHS	3,698	0.04%	2266	24.95%	70.26%	0.74%	X	X	0.022	0.277***	-0.831***	0.020	0.116
											(0.021)	(0.041)	(0.183)	(0.015)	(0.844)
Pakistan	2012	DHS	5,043	0.05%	2494	27.19%	66.66%	0.78%	X	X	-0.021	0.290***	0.163	0.034**	0.649
											(0.023)	(0.061)	(0.389)	(0.015)	(0.629)
Panama	1960	IPUMS-I	2,780	0.03%	2484	18.06%	71.33%	1.26%	X		-0.149***	0.315***	-0.063	0.005	2.901
											(0.020)	(0.033)	(0.202)	(0.016)	(9.949)
Paraguay	1962	IPUMS-I	4,420	0.05%	1638	20.11%	71.67%	1.27%	X		-0.133***	0.313***	0.113	0.010	2.751
											(0.016)	(0.024)	(0.180)	(0.012)	(3.755)
Paraguay	1972	IPUMS-I	11,299	0.12%	1990	16.00%	69.20%	0.90%	X		-0.120***	0.302***	-0.065	0.010	0.317
											(0.009)	(0.021)	(0.114)	(0.008)	(0.745)
Philippines	1990	IPUMS-I	347,726	3.58%	2120	30.33%	64.46%	1.31%	X		-0.072***	0.342***	0.056***	0.026***	-0.087
											(0.002)	(0.003)	(0.021)	(0.001)	(0.060)
Philippines	1993	DHS	3,732	0.04%	2162	37.46%	63.02%	0.51%	X	X	-0.103***	0.314***	-0.344	0.028**	0.057
											(0.019)	(0.058)	(0.391)	(0.014)	(0.581)
Philippines	1998	DHS	3,290	0.03%	2290	41.50%	62.18%	0.82%	X	X	-0.091***	0.384***	-0.076	0.028*	-0.707
											(0.022)	(0.056)	(0.289)	(0.017)	(0.762)
Philippines	2003	DHS	3,001	0.03%	2486	42.70%	55.77%	0.65%	X	X	-0.068***	0.475***	0.020	0.030*	-0.540
											(0.022)	(0.061)	(0.239)	(0.016)	(0.662)
Rwanda	1991	IPUMS-I	42,005	0.43%	800	97.33%	67.97%	1.72%		X	0.002	0.276***	-0.047*	0.002	-0.049
											(0.002)	(0.010)	(0.026)	(0.004)	(0.791)
Rwanda	1992	DHS	1,710	0.02%	770	97.99%	64.15%	0.61%	X	X	-0.004	0.285***	-0.045	0.011	0.275
											(0.007)	(0.084)	(0.119)	(0.019)	(0.688)
Rwanda	2000	DHS	2,294	0.02%	743	87.59%	57.34%	0.32%	X	X	0.004	0.407***	-0.589	-0.002	-5.563
											(0.018)	(0.077)	(0.370)	(0.018)	(57.765)
Rwanda	2002	IPUMS-I	41,817	0.43%	794	92.61%	56.63%	1.43%	X	X	-0.011***	0.434***	-0.083***	0.005	-1.176
											(0.003)	(0.010)	(0.029)	(0.004)	(1.174)
Rwanda	2005	DHS	2,668	0.03%	884	71.89%	60.91%	0.69%	X	X	0.009	0.331***	-0.013	-0.012	2.386

											(0.023)	(0.049)	(0.354)	(0.017)	(3.618)
Saotome	2008	DHS	813	0.01%	1484	57.08%	58.60%	2.42%	X	X	0.033	0.432***	0.473**	0.022	-2.658
											(0.052)	(0.078)	(0.235)	(0.038)	(4.827)
Senegal	1988	IPUMS-I	39,875	0.41%	1267	22.70%	67.83%	2.04%	X		-0.020***	0.295***	-0.017	0.005	0.136
											(0.005)	(0.008)	(0.050)	(0.004)	(0.782)
Senegal	1992	DHS	1,814	0.02%	1229	46.64%	68.14%	0.66%	X	X	-0.027	0.275***	-0.208	-0.001	-7.040
											(0.029)	(0.084)	(0.516)	(0.019)	(104.188)
Senegal	1997	DHS	2,320	0.02%	1245	61.92%	63.39%	0.53%	X	X	0.062**	0.322***	-0.282	0.017	-0.460
											(0.030)	(0.069)	(0.571)	(0.019)	(1.430)
Senegal	2002	IPUMS-I	41,222	0.42%	1359	29.16%	65.24%	2.67%	X		-0.009	0.352***	0.012	0.001	-1.602
											(0.005)	(0.007)	(0.039)	(0.004)	(15.525)
Senegal	2005	DHS	3,522	0.04%	1424	37.17%	61.36%	1.00%	X	X	-0.041*	0.436***	0.096	0.018	-1.792
											(0.024)	(0.039)	(0.214)	(0.017)	(2.036)
Senegal	2010	DHS	4,103	0.04%	1507	38.82%	63.53%	1.28%	X	X	-0.040	0.400***	0.299	0.005	1.691
											(0.024)	(0.041)	(0.231)	(0.017)	(7.736)
Senegal	2014	DHS	2,320	0.02%	1507	47.17%	61.10%	1.27%	X	X	-0.048	0.371***	-0.037	-0.047**	-0.260
											(0.034)	(0.058)	(0.300)	(0.022)	(0.576)
Sierra Leone	2004	IPUMS-I	18,744	0.19%	587	70.26%	58.47%	3.93%	X		-0.002	0.436***	0.052	0.005	-1.015
											(0.007)	(0.008)	(0.039)	(0.007)	(1.862)
Sierra Leone	2008	DHS	1,973	0.02%	686	80.81%	52.94%	1.17%	X	X	0.000	0.517***	0.096	0.035	0.652
											(0.022)	(0.051)	(0.140)	(0.023)	(0.712)
Sweden	1880	NAPP	139,113	1.43%	1503	1.90%	56.79%	0.66%			-0.026***	0.407***	0.003	0.009***	0.113
											(0.001)	(0.008)	(0.012)	(0.002)	(0.086)
Sweden	1890	NAPP	152,922	1.58%	1647	3.32%	59.56%	0.61%			-0.039***	0.399***	0.007	0.010***	-0.000
											(0.001)	(0.008)	(0.015)	(0.002)	(0.089)
Sweden	1900	NAPP	149,091	1.54%	2087	2.78%	59.29%	0.63%			-0.028***	0.413***	0.013	0.012***	0.007
											(0.001)	(0.008)	(0.014)	(0.002)	(0.069)
Tajikistan	2012	DHS	2,389	0.02%	1661	24.45%	56.13%	0.61%	X	X	-0.057**	0.295***	0.629	0.027	0.513
											(0.025)	(0.083)	(0.596)	(0.018)	(0.813)
Tanzania	1988	IPUMS-I	112,710	1.16%	540	88.66%	63.44%	2.46%	X		0.006**	0.355***	0.036**	-0.007**	-0.293
											(0.002)	(0.005)	(0.018)	(0.003)	(0.335)
Tanzania	1991	DHS	2,468	0.03%	536	73.19%	61.87%	0.84%	X	X	-0.016	0.458***	-0.437	0.030	1.053
											(0.027)	(0.052)	(0.292)	(0.020)	(1.050)
Tanzania	1996	DHS	2,249	0.02%	525	57.82%	61.91%	0.76%	X	X	0.032	0.397***	0.140	0.002	-5.818
											(0.026)	(0.065)	(0.308)	(0.019)	(55.551)
Tanzania	1999	DHS	1,069	0.01%	546	78.20%	59.00%	1.80%	X	X	0.020	0.499***	0.218	0.010	4.259
											(0.044)	(0.066)	(0.138)	(0.034)	(14.490)
Tanzania	2002	IPUMS-I	191,556	1.97%	591	78.48%	60.83%	2.37%	X		0.033***	0.385***	-0.025	0.001	1.121
											(0.003)	(0.005)	(0.022)	(0.002)	(5.130)
Tanzania	2004	DHS	2,914	0.03%	637	86.06%	59.81%	1.64%	X	X	0.057***	0.286***	0.033	0.006	-0.397
											(0.019)	(0.056)	(0.239)	(0.018)	(2.800)
Tanzania	2010	DHS	2,708	0.03%	804	86.72%	60.98%	0.82%	X	X	0.032*	0.476***	0.043	0.016	-0.234
											(0.018)	(0.051)	(0.153)	(0.020)	(1.065)
Tanzania	2012	IPUMS-I	225,907	2.33%	804	76.60%	60.96%	2.40%	X		0.034***	0.394***	-0.050***	0.000	6.126
											(0.002)	(0.004)	(0.017)	(0.002)	(85.181)
Togo	1998	DHS	2,461	0.03%	661	87.07%	62.35%	2.04%	X	X	0.020	0.435***	-0.064	-0.005	-2.786
											(0.018)	(0.044)	(0.143)	(0.018)	(11.040)
USA	1860	IPUMS-USA	14,364	0.15%	2219	5.01%	63.71%	0.66%			-0.011**	0.321***	0.128	0.023***	0.101
											(0.004)	(0.023)	(0.092)	(0.007)	(0.162)
USA	1870	IPUMS-USA	18,167	0.19%	2497	5.20%	61.49%	0.81%			-0.023***	0.360***	0.114*	0.008	-0.441
											(0.004)	(0.020)	(0.066)	(0.006)	(0.539)
Uganda	1991	IPUMS-I	84,404	0.87%	584	72.66%	66.06%	1.81%	X		0.005	0.328***	-0.024	-0.001	0.430
											(0.004)	(0.007)	(0.040)	(0.003)	(4.401)
Uganda	1995	DHS	2,144	0.02%	654	65.03%	67.39%	0.75%	X	X	0.044	0.329***	-0.218	-0.020	1.404
											(0.030)	(0.062)	(0.390)	(0.019)	(1.692)
Uganda	2000	DHS	2,236	0.02%	780	79.79%	69.26%	0.55%	X	X	0.012	0.389***	-0.079	-0.009	-1.633
											(0.026)	(0.055)	(0.319)	(0.018)	(4.208)
Uganda	2002	IPUMS-I	136,380	1.40%	835	58.16%	69.79%	2.55%	X	X	-0.017***	0.277***	0.023	-0.000	-11.084
											(0.003)	(0.004)	(0.030)	(0.002)	(147.754)
Uganda	2006	DHS	2,685	0.03%	989	88.95%	68.53%	0.97%	X	X	0.025	0.301***	-0.117	0.041**	-0.051
											(0.017)	(0.059)	(0.264)	(0.016)	(0.317)
Uganda	2011	DHS	2,593	0.03%	1158	75.80%	66.13%	1.13%	X	X	-0.037	0.348***	-0.127	0.010	2.776
											(0.025)	(0.048)	(0.266)	(0.018)	(5.307)
Vietnam	1989	IPUMS-I	166,529	1.72%	1009	87.94%	55.40%	1.06%	X	X	-0.008***	0.407***	-0.068**	0.029***	-0.105
											(0.002)	(0.008)	(0.027)	(0.003)	(0.067)
Vietnam	1997	DHS	1,910	0.02%	1560	92.01%	43.07%	0.57%	X	X	0.005	0.624***	-0.125	0.087***	-0.069
											(0.015)	(0.078)	(0.208)	(0.022)	(0.153)
Vietnam	1999	IPUMS-I	133,016	1.37%	1739	85.28%	37.65%	0.61%	X	X	0.003	0.619***	-0.080***	0.068***	-0.029
											(0.003)	(0.010)	(0.029)	(0.003)	(0.036)
Vietnam	2002	DHS	1,634	0.02%	2039	93.13%	28.52%	0.22%	X	X	-0.032**	0.611***	0.070***	0.129***	-0.164
											(0.015)	(0.086)	(0.024)	(0.023)	(0.106)
Yemen	1991	DHS	1,505	0.02%	2380	12.09%	78.55%	0.87%	X	X	-0.031	0.200***	-0.275	0.005	1.355
											(0.024)	(0.057)	(0.302)	(0.019)	(6.358)
Zambia	1990	IPUMS-I	33,408	0.34%	772	27.77%	69.37%	2.25%	X		-0.043***	0.298***	0.007	-0.004	-0.432
											(0.006)	(0.009)	(0.056)	(0.004)	(1.189)
Zambia	1992	DHS	1,963	0.02%	730	59.08%	65.70%	0.70%	X	X	-0.088***	0.292***	-0.239	0.004	4.527
											(0.029)	(0.068)	(0.418)	(0.018)	(20.179)
Zambia	1996	DHS	2,302	0.02%	635	53.16%	62.90%	0.78%	X	X	-0.093***	0.532***	-0.052	-0.015	0.323
											(0.026)	(0.046)	(0.227)	(0.018)	(1.553)
Zambia	2000	IPUMS-I	49,762	0.51%	613	48.83%	64.16%	2.85%	X		-0.028***	0.346***	-0.014	0.008**	0.519
											(0.005)	(0.006)	(0.039)	(0.004)	(0.623)
Zambia	2001	DHS	2,288	0.02%	616	60.77%	62.13%	0.58%	X	X	-0.053**	0.453***	-0.004	-0.000	26.575
											(0.027)	(0.066)	(0.312)	(0.018)	(1277.218)
Zambia	2007	DHS	2,267	0.02%	716	52.59%	64.48%	1.41%	X	X	-0.047*	0.398***	0.061	0.000	11.750
											(0.028)	(0.051)	(0.249)	(0.018)	(769.580)
Zambia	2010	IPUMS-I	78,308	0.81%	795	52.81%	66.79%	1.78%	X		0.017***	0.321***	0.068	-0.002	1.337

											(0.004)	(0.006)	(0.042)	(0.003)	(2.727)
Zambia	2013	DHS	5,091	0.05%	795	56.97%	65.01%	0.99%	X	X	0.042**	0.430***	-0.010	-0.004	2.695
											(0.021)	(0.045)	(0.185)	(0.013)	(9.598)
Zimbabwe	1994	DHS	1,467	0.02%	1341	59.77%	59.07%	1.13%	X	X	-0.058*	0.418***	-0.548	0.036	0.902
											(0.034)	(0.074)	(0.387)	(0.022)	(0.968)
Zimbabwe	1999	DHS	1,240	0.01%	1311	57.10%	47.50%	0.31%	X	X	-0.011	0.562***	0.260	0.025	1.708
											(0.038)	(0.141)	(0.355)	(0.025)	(2.130)
Zimbabwe	2005	DHS	2,135	0.02%	872	37.41%	44.59%	1.00%	X	X	-0.119***	0.627***	0.027	-0.004	-9.176
											(0.027)	(0.063)	(0.185)	(0.022)	(49.492)
Zimbabwe	2010	DHS	2,246	0.02%	750	38.13%	40.60%	1.00%	X	X	-0.119***	0.604***	-0.262	0.008	-1.456
											(0.026)	(0.060)	(0.165)	(0.019)	(4.130)

	Year (num. samples)	Source	N	Percent of bin	Mean real GDP/C	In labor force	3 or more children	2nd child is twin	Educa-tion?	Birth Quarter?	OLS	FS Twin IV	2S Twin IV	FS Same gender IV	2S Same gender IV
	2,500 to 5,000 real GDP/Capita bin														
Pooled	105		13,148,608			23.59%	55.63%	0.99%			-0.044***	0.431***	-0.003	0.021***	-0.002
											(0.007)	(0.030)	(0.010)	(0.005)	(0.016)
Albania	2008	DHS	1,223	0.01%	4916	27.38%	34.49%	0.39%	X	X	-0.182***	0.695***	0.269	0.146***	-0.018
											(0.034)	(0.065)	(0.261)	(0.029)	(0.209)
Armenia	2000	DHS	1,500	0.01%	4912	29.97%	32.82%	0.44%	X	X	-0.006	0.740***	-0.126	0.084***	-0.469
											(0.028)	(0.040)	(0.197)	(0.023)	(0.317)
Bolivia	1976	IPUMS-I	25,165	0.19%	2571	17.46%	61.92%	0.58%	X		-0.076***	0.370***	-0.049	0.014***	-0.008
											(0.006)	(0.021)	(0.082)	(0.005)	(0.337)
Bolivia	1998	DHS	2,850	0.02%	2510	52.76%	58.13%	0.42%	X	X	-0.153***	0.383***	0.109	0.037**	-0.197
											(0.024)	(0.063)	(0.408)	(0.019)	(0.562)
Bolivia	2001	IPUMS-I	38,755	0.29%	2566	41.94%	56.08%	0.87%	X		-0.097***	0.451***	-0.022	0.013***	0.050
											(0.006)	(0.012)	(0.059)	(0.005)	(0.392)
Bolivia	2003	DHS	4,441	0.03%	2611	60.28%	56.54%	0.28%	X	X	-0.066***	0.358***	0.254	0.020	-0.123
											(0.020)	(0.046)	(0.282)	(0.017)	(0.895)
Bolivia	2008	DHS	3,943	0.03%	2920	64.77%	52.31%	0.48%	X	X	-0.056***	0.392***	-0.082	0.058***	0.061
											(0.021)	(0.044)	(0.353)	(0.017)	(0.320)
Botswana	1991	IPUMS-I	5,484	0.04%	3258	47.79%	60.38%	3.37%	X		-0.107***	0.433***	0.106	-0.020*	0.336
											(0.015)	(0.017)	(0.087)	(0.012)	(0.724)
Botswana	2001	IPUMS-I	6,152	0.05%	4157	53.30%	49.38%	3.66%	X		-0.138***	0.507***	0.000	0.001	-7.320
											(0.014)	(0.015)	(0.067)	(0.012)	(139.580)
Brazil	1970	IPUMS-I	255,612	1.94%	3124	11.44%	68.47%	1.93%	X		-0.059***	0.305***	-0.061***	0.021***	-0.067
											(0.002)	(0.003)	(0.014)	(0.002)	(0.062)
Brazil	1980	IPUMS-I	312,368	2.38%	4777	21.49%	59.03%	1.96%	X		-0.080***	0.372***	-0.063***	0.025***	-0.008
											(0.002)	(0.003)	(0.014)	(0.002)	(0.060)
Canada	1911	NAPP	13,428	0.10%	4079	3.25%	62.11%	0.72%		X	-0.006*	0.361***	-0.063**	0.008	0.111
											(0.004)	(0.022)	(0.029)	(0.007)	(0.381)
Colombia	1973	IPUMS-I	97,406	0.74%	3442	14.26%	69.97%	1.47%	X		-0.095***	0.300***	0.050	0.017***	0.109
											(0.003)	(0.006)	(0.032)	(0.003)	(0.133)
Colombia	1985	IPUMS-I	144,601	1.10%	4366	33.52%	53.72%	1.93%	X		-0.084***	0.420***	-0.010	0.035***	-0.030
											(0.003)	(0.004)	(0.022)	(0.002)	(0.072)
Colombia	1990	DHS	1,922	0.01%	4817	35.66%	50.12%	0.88%	X	X	-0.106***	0.497***	0.502	0.037	-0.188
											(0.030)	(0.066)	(0.338)	(0.028)	(0.745)
Costa Rica	1973	IPUMS-I	9,714	0.07%	4202	12.93%	69.26%	0.80%	X		-0.103***	0.323***	0.025	-0.004	1.337
											(0.009)	(0.025)	(0.118)	(0.008)	(3.429)
Costa Rica	1984	IPUMS-I	15,379	0.12%	4413	18.45%	53.64%	1.22%	X		-0.086***	0.484***	0.070	0.043***	0.026
											(0.007)	(0.017)	(0.062)	(0.007)	(0.144)
Cuba	2002	IPUMS-I	36,099	0.27%	2583	35.50%	17.01%	1.00%	X		-0.096***	0.829***	0.006	0.033***	-0.002
											(0.006)	(0.005)	(0.030)	(0.004)	(0.149)
Dominican Republic	2002	IPUMS-I	42,518	0.32%	3803	66.60%	53.00%	2.70%	X	X	-0.038***	0.445***	-0.031	0.031***	-0.271*
											(0.005)	(0.006)	(0.032)	(0.004)	(0.152)
Dominican Republic	1991	DHS	1,762	0.01%	2602	40.66%	58.80%	1.21%	X	X	-0.035	0.493***	-0.519**	0.051*	-0.517
											(0.037)	(0.064)	(0.221)	(0.029)	(0.686)
Dominican Republic	1996	DHS	2,107	0.02%	3120	37.87%	55.60%	0.93%	X	X	-0.078***	0.453***	-0.057	0.082***	-0.175
											(0.027)	(0.072)	(0.327)	(0.023)	(0.300)
Dominican Republic	1999	DHS	314	0.00%	3522	46.27%	50.15%	1.15%	X	X	-0.060	0.522***	-0.338	-0.046	-0.995
											(0.075)	(0.171)	(0.526)	(0.056)	(1.851)
Dominican Republic	2002	DHS	5,718	0.04%	3803	38.52%	51.86%	0.76%	X	X	-0.104***	0.498***	0.172	0.027	-0.472
											(0.019)	(0.032)	(0.180)	(0.017)	(0.684)
Dominican Republic	2007	DHS	5,876	0.04%	4649	42.11%	52.19%	0.90%	X	X	-0.062***	0.469***	-0.120	0.024	0.205
											(0.021)	(0.031)	(0.217)	(0.018)	(0.822)
Ecuador	1974	IPUMS-I	32,604	0.25%	3234	11.15%	68.51%	0.82%	X		-0.070***	0.296***	-0.003	0.011**	0.368
											(0.005)	(0.018)	(0.073)	(0.005)	(0.424)
Ecuador	1982	IPUMS-I	44,110	0.34%	4025	15.79%	63.11%	0.97%	X		-0.101***	0.388***	0.071	0.015***	0.219
											(0.004)	(0.011)	(0.049)	(0.004)	(0.248)
Ecuador	1990	IPUMS-I	52,893	0.40%	3941	25.67%	57.01%	0.91%	X		-0.102***	0.425***	-0.019	0.027***	0.113
											(0.004)	(0.011)	(0.046)	(0.004)	(0.145)
Ecuador	2001	IPUMS-I	56,918	0.43%	4081	31.32%	48.77%	1.15%	X		-0.088***	0.540***	-0.051	0.026***	-0.073
											(0.004)	(0.008)	(0.032)	(0.004)	(0.151)
Egypt	1992	DHS	3,869	0.03%	2563	21.42%	69.28%	0.99%	X	X	-0.027	0.278***	-0.034	0.028**	-0.498
											(0.019)	(0.046)	(0.263)	(0.014)	(0.561)
Egypt	1995	DHS	5,599	0.04%	2726	18.47%	65.34%	0.77%	X	X	-0.048***	0.369***	0.397*	0.038***	-0.322
											(0.016)	(0.049)	(0.209)	(0.013)	(0.340)
Egypt	1996	IPUMS-I	372,603	2.83%	2819	14.60%	63.49%	1.41%	X	X	-0.063***	0.361***	-0.019	0.040***	0.007
											(0.002)	(0.004)	(0.013)	(0.001)	(0.028)
Egypt	2000	DHS	5,707	0.04%	3193	14.90%	60.82%	1.04%	X	X	-0.025*	0.428***	0.154	0.032***	0.336
											(0.013)	(0.039)	(0.132)	(0.011)	(0.341)
Egypt	2003	DHS	3,256	0.02%	3409	18.97%	55.98%	0.67%	X	X	-0.032	0.491***	-0.191	0.061***	-0.249
											(0.020)	(0.049)	(0.169)	(0.016)	(0.267)
Egypt	2005	DHS	6,910	0.05%	3599	18.00%	54.96%	1.30%	X	X	0.001	0.543***	-0.043	0.081***	0.126
											(0.013)	(0.034)	(0.083)	(0.011)	(0.133)
Egypt	2006	IPUMS-I	439,867	3.35%	3714	13.60%	52.46%	1.46%	X		-0.022***	0.474***	0.006	0.048***	0.025
											(0.001)	(0.003)	(0.009)	(0.001)	(0.021)
Egypt	2008	DHS	5,814	0.04%	3992	12.65%	52.73%	1.21%	X	X	-0.022*	0.472***	0.020	0.043***	0.037
											(0.012)	(0.037)	(0.093)	(0.012)	(0.218)
Egypt	2014	DHS	8,447	0.06%	4267	13.17%	52.32%	1.22%	X	X	-0.011	0.444***	0.064	0.041***	-0.166
											(0.011)	(0.029)	(0.104)	(0.010)	(0.217)
El Salvador	2007	IPUMS-I	29,636	0.23%	2897	41.52%	46.04%	1.94%	X	X	-0.111***	0.515***	-0.068*	0.023***	0.156
											(0.006)	(0.009)	(0.040)	(0.005)	(0.254)
Gabon	2000	DHS	1,348	0.01%	4174	43.79%	56.77%	1.60%	X	X	-0.017	0.463***	0.014	-0.009	-6.899

											(0.035)	(0.049)	(0.259)	(0.028)	(20.983)
Great Britain	1851	NAPP	11,693	0.09%	2561	30.30%	64.88%	0.51%			-0.066***	0.391***	-0.114	0.015*	0.228
											(0.011)	(0.027)	(0.151)	(0.008)	(0.616)
Great Britain	1881	NAPP	972,869	7.40%	3530	28.03%	68.77%	0.47%			-0.068***	0.325***	0.006	0.005***	0.053
											(0.001)	(0.003)	(0.021)	(0.001)	(0.180)
Great Britain	1911	NAPP	938,191	7.14%	4699	8.88%	58.16%	0.71%			-0.044***	0.432***	-0.026***	0.012***	0.007
											(0.001)	(0.003)	(0.008)	(0.001)	(0.048)
Guatemala	1995	DHS	3,639	0.03%	3559	28.53%	67.62%	0.62%	X	X	-0.174***	0.213***	-0.187	0.000	-80.537
											(0.027)	(0.065)	(0.735)	(0.018)	(31589.125)
Guatemala	1998	DHS	1,787	0.01%	3760	31.58%	66.74%	0.58%	X	X	-0.122***	0.421***	0.397	0.072**	-0.291
											(0.043)	(0.042)	(0.650)	(0.032)	(0.472)
India	2009	IPUMS-I	29,556	0.22%	3159	27.46%	42.50%	0.39%	X		-0.034***	0.613***	0.007	0.045***	-0.187
											(0.011)	(0.030)	(0.114)	(0.011)	(0.231)
Indonesia	1990	IPUMS-I	57,518	0.44%	2543	42.10%	52.80%	0.63%	X	X	-0.075***	0.464***	-0.109*	0.025***	0.121
											(0.005)	(0.012)	(0.056)	(0.004)	(0.173)
Indonesia	1991	DHS	8,118	0.06%	2690	40.47%	52.30%	0.47%	X	X	-0.058***	0.511***	-0.274	0.001	5.638
											(0.017)	(0.051)	(0.167)	(0.014)	(56.949)
Indonesia	1995	IPUMS-I	41,916	0.32%	3256	42.74%	45.15%	0.50%	X	X	-0.064***	0.538***	-0.058	0.028***	-0.078
											(0.007)	(0.016)	(0.073)	(0.005)	(0.208)
Indonesia	2002	DHS	8,192	0.06%	3429	43.83%	35.19%	0.34%	X	X	-0.051**	0.688***	-0.080	0.001	-2.830
											(0.020)	(0.035)	(0.169)	(0.017)	(96.533)
Indonesia	2007	DHS	8,920	0.07%	4161	50.58%	32.06%	0.51%	X	X	-0.052***	0.705***	0.132	0.046***	-0.460
											(0.018)	(0.024)	(0.143)	(0.015)	(0.386)
Indonesia	2010	IPUMS-I	1,055,321	8.03%	4722	55.55%	29.83%	0.73%	X	X	-0.032***	0.705***	-0.037***	0.030***	-0.072**
											(0.001)	(0.002)	(0.008)	(0.001)	(0.032)
Indonesia	2012	DHS	8,276	0.06%	4722	51.89%	26.40%	0.49%	X	X	-0.049***	0.728***	0.112	0.028**	0.250
											(0.018)	(0.022)	(0.145)	(0.013)	(0.594)
Jamaica	1982	IPUMS-I	9,385	0.07%	3167	51.67%	57.46%	2.16%	X		-0.157***	0.439***	-0.013	0.006	-0.713
											(0.011)	(0.015)	(0.081)	(0.009)	(2.037)
Jamaica	1991	IPUMS-I	11,693	0.09%	3731	44.33%	51.22%	2.30%	X	X	-0.150***	0.504***	0.002	0.001	-1.729
											(0.010)	(0.013)	(0.060)	(0.008)	(25.555)
Jamaica	2001	IPUMS-I	9,267	0.07%	3700	52.77%	48.41%	2.06%	X	X	-0.125***	0.513***	0.015	0.011	0.581
											(0.011)	(0.016)	(0.072)	(0.010)	(1.143)
Jordan	1990	DHS	2,767	0.02%	4080	10.34%	80.61%	0.55%	X	X	0.006	0.231***	-0.405**	0.045***	-0.246
											(0.020)	(0.060)	(0.194)	(0.013)	(0.266)
Jordan	1997	DHS	2,490	0.02%	4039	10.75%	72.19%	1.05%	X	X	-0.023	0.261***	0.642*	-0.014	0.851
											(0.018)	(0.052)	(0.346)	(0.016)	(1.309)
Jordan	2002	DHS	2,559	0.02%	4504	7.83%	73.62%	0.98%	X	X	-0.016	0.333***	0.362	0.042**	0.303
											(0.018)	(0.046)	(0.419)	(0.017)	(0.320)
Jordan	2004	IPUMS-I	28,275	0.22%	4799	16.51%	69.81%	1.38%	X	X	-0.062***	0.324***	0.001	0.019***	0.007
											(0.006)	(0.014)	(0.059)	(0.004)	(0.230)
Kyrgyz Republic	2012	DHS	2,070	0.02%	2947	21.72%	49.90%	0.74%	X	X	-0.209***	0.480***	-0.092	0.001	-9.926
											(0.027)	(0.059)	(0.272)	(0.022)	(141.397)
Kyrgyz Republic	2009	IPUMS-I	30,670	0.23%	2976	66.30%	49.59%	0.91%	X	X	-0.029***	0.519***	-0.032	0.052***	-0.088
											(0.006)	(0.015)	(0.055)	(0.005)	(0.102)
Malaysia	1980	IPUMS-I	10,040	0.08%	3619	32.32%	63.84%	1.25%	X		-0.058***	0.381***	-0.007	0.025***	-0.434
											(0.011)	(0.023)	(0.109)	(0.008)	(0.392)
Mexico	1970	IPUMS-I	26,355	0.20%	4331	9.98%	76.47%	1.16%	X		-0.061***	0.252***	0.204**	0.015***	0.089
											(0.005)	(0.011)	(0.082)	(0.005)	(0.244)
Moldova	2005	DHS	1,026	0.01%	3311	50.51%	18.23%	1.12%	X	X	-0.131***	0.799***	-0.232	0.069***	0.041
											(0.043)	(0.027)	(0.172)	(0.025)	(0.468)
Morocco	1992	DHS	1,943	0.01%	2590	18.94%	68.81%	0.57%	X	X	-0.050**	0.230***	-0.031	0.031*	0.470
											(0.024)	(0.083)	(0.491)	(0.018)	(0.645)
Morocco	1994	IPUMS-I	60,890	0.46%	2626	11.45%	66.00%	1.43%	X	X	-0.057***	0.331***	0.034	0.021***	0.129
											(0.004)	(0.009)	(0.034)	(0.003)	(0.123)
Morocco	2003	DHS	2,718	0.02%	3167	11.81%	53.23%	0.58%	X	X	-0.037**	0.542***	0.187	0.046***	-0.176
											(0.017)	(0.067)	(0.199)	(0.017)	(0.284)
Morocco	2004	IPUMS-I	60,390	0.46%	3286	10.40%	53.04%	1.23%	X	X	-0.015***	0.486***	-0.010	0.034***	0.034
											(0.003)	(0.010)	(0.023)	(0.003)	(0.074)
Mozambique	2011	DHS	3,843	0.03%	2613	41.57%	63.00%	1.18%	X	X	-0.044*	0.426***	-0.160	-0.008	-0.053
											(0.023)	(0.034)	(0.163)	(0.016)	(2.293)
Namibia	1992	DHS	988	0.01%	3335	35.62%	54.21%	0.93%	X	X	-0.085**	0.427***	-0.675**	0.026	-0.795
											(0.036)	(0.067)	(0.269)	(0.029)	(1.421)
Namibia	2000	DHS	1,108	0.01%	3652	38.46%	44.15%	1.85%	X	X	-0.210***	0.511***	-0.043	0.011	-3.780
											(0.041)	(0.058)	(0.302)	(0.034)	(11.569)
Namibia	2006	DHS	1,413	0.01%	4277	49.00%	38.30%	1.26%	X	X	-0.135***	0.503***	-0.133	0.047*	-0.677
											(0.036)	(0.055)	(0.258)	(0.027)	(0.737)
Nicaragua	1971	IPUMS-I	10,485	0.08%	2906	17.22%	74.54%	0.77%	X		-0.112***	0.284***	-0.052	0.025***	-0.169
											(0.010)	(0.023)	(0.145)	(0.008)	(0.292)
Panama	1970	IPUMS-I	8,373	0.06%	3828	23.59%	72.14%	1.28%	X		-0.152***	0.298***	0.204	0.008	0.279
											(0.012)	(0.019)	(0.149)	(0.009)	(1.302)
Panama	1980	IPUMS-I	10,736	0.08%	4850	30.39%	62.67%	1.45%	X		-0.145***	0.398***	0.026	0.012	-0.279
											(0.011)	(0.019)	(0.094)	(0.008)	(0.757)
Panama	1990	IPUMS-I	12,549	0.10%	4818	29.91%	55.26%	1.44%	X		-0.153***	0.465***	0.034	0.046***	0.152
											(0.009)	(0.016)	(0.074)	(0.008)	(0.180)
Paraguay	1982	IPUMS-I	15,623	0.12%	3193	15.34%	63.15%	0.97%	X		-0.112***	0.394***	0.100	0.020***	0.209
											(0.007)	(0.018)	(0.080)	(0.007)	(0.310)
Paraguay	1990	DHS	1,519	0.01%	3226	34.25%	60.14%	1.02%	X	X	-0.162***	0.422***	0.145	0.029	-0.682
											(0.030)	(0.074)	(0.305)	(0.023)	(0.959)
Paraguay	1992	IPUMS-I	22,777	0.17%	3274	19.37%	61.58%	0.97%	X		-0.127***	0.398***	-0.010	0.031***	0.181
											(0.006)	(0.015)	(0.065)	(0.006)	(0.177)
Paraguay	2002	IPUMS-I	24,926	0.19%	2997	36.77%	56.31%	1.03%	X		-0.141***	0.447***	0.050	0.025***	-0.089
											(0.007)	(0.014)	(0.068)	(0.006)	(0.242)
Peru	1991	DHS	3,929	0.03%	3196	52.80%	56.91%	0.59%	X	X	-0.043**	0.439***	-0.525*	0.001	-31.330
											(0.020)	(0.053)	(0.269)	(0.015)	(603.874)
Peru	1993	IPUMS-I	113,466	0.86%	3220	24.40%	55.08%	0.92%	X		-0.092***	0.449***	0.014	0.023***	0.037

											(0.003)	(0.007)	(0.030)	(0.003)	(0.111)
Peru	1996	DHS	7,325	0.06%	3531	51.25%	55.41%	0.40%	X	X	-0.094***	0.500***	-0.055	0.031**	-0.384
											(0.017)	(0.039)	(0.216)	(0.013)	(0.482)
Peru	2000	DHS	6,371	0.05%	3766	57.25%	49.41%	0.48%	X	X	-0.053***	0.503***	0.014	0.010	-0.283
											(0.018)	(0.042)	(0.186)	(0.014)	(1.652)
Peru	2007	IPUMS-I	115,601	0.88%	4923	33.95%	41.38%	0.94%	X		-0.095***	0.591***	-0.024	0.026***	0.068
											(0.003)	(0.006)	(0.024)	(0.003)	(0.107)
Peru	2007	DHS	7,867	0.06%	4923	67.08%	44.31%	0.64%	X	X	-0.006	0.607***	0.099	0.026*	0.311
											(0.016)	(0.044)	(0.139)	(0.014)	(0.611)
Philippines	2008	DHS	2,717	0.02%	2863	42.30%	53.96%	0.67%	X	X	-0.096***	0.456***	-0.156	0.011	1.836
											(0.023)	(0.053)	(0.222)	(0.018)	(3.809)
Philippines	2013	DHS	3,014	0.02%	3024	42.14%	48.66%	0.54%	X	X	-0.113***	0.501***	0.262	0.036**	0.064
											(0.020)	(0.065)	(0.243)	(0.017)	(0.520)
Romania	1992	IPUMS-I	100,657	0.77%	3191	74.01%	34.41%	0.89%	X	X	-0.192***	0.655***	-0.066***	0.034***	-0.075
											(0.003)	(0.006)	(0.023)	(0.003)	(0.079)
Romania	2002	IPUMS-I	71,737	0.55%	3456	54.25%	22.54%	0.87%	X	X	-0.168***	0.786***	-0.081***	0.035***	0.083
											(0.004)	(0.005)	(0.025)	(0.003)	(0.107)
Romania	2011	IPUMS-I	46,774	0.36%	4653	57.10%	23.02%	1.33%	X	X	-0.126***	0.784***	-0.039	0.038***	0.230*
											(0.005)	(0.005)	(0.024)	(0.004)	(0.119)
South Africa	1996	IPUMS-I	133,590	1.02%	3700	68.23%	47.38%	2.34%	X		-0.095***	0.515***	-0.015	0.022***	-0.093
											(0.003)	(0.004)	(0.016)	(0.002)	(0.117)
South Africa	2001	IPUMS-I	136,950	1.04%	4005	72.27%	43.16%	2.44%	X		-0.093***	0.559***	0.005	0.019***	0.221*
											(0.003)	(0.004)	(0.014)	(0.002)	(0.133)
South Africa	2007	IPUMS-I	33,071	0.25%	4783	82.26%	39.58%	2.47%	X		-0.063***	0.613***	-0.034	0.016***	0.096
											(0.005)	(0.007)	(0.024)	(0.005)	(0.275)
South Africa	1998	DHS	2,067	0.02%	3812	32.69%	39.99%	0.85%	X	X	-0.134***	0.576***	-0.145	0.012	1.310
											(0.027)	(0.043)	(0.232)	(0.022)	(3.417)
Sudan	2008	IPUMS-I	289,810	2.20%	3021	24.70%	71.98%	1.47%	X		-0.009***	0.288***	0.011	0.018***	0.095
											(0.003)	(0.005)	(0.035)	(0.002)	(0.138)
Swaziland	2006	DHS	851	0.01%	2967	42.78%	50.21%	0.68%	X	X	-0.086**	0.509***	-0.030	-0.011	1.808
											(0.040)	(0.099)	(0.435)	(0.031)	(6.471)
Turkey	1985	IPUMS-I	150,756	1.15%	4578	39.19%	57.39%	1.39%	X		0.103***	0.462***	0.293***	0.051***	-0.045
											(0.003)	(0.005)	(0.024)	(0.002)	(0.050)
USA	1880	US Full Count	2,391,227	18.19%	3032	6.19%	64.10%	0.66%			-0.023***	0.343***	0.039***	0.009***	0.042
											(0.000)	(0.002)	(0.006)	(0.001)	(0.035)
USA	1900	US Full Count	3,139,566	23.88%	4161	6.39%	61.24%	1.07%		X	-0.028***	0.373***	0.014***	0.010***	0.080***
											(0.000)	(0.001)	(0.004)	(0.000)	(0.029)
Ukraine	2007	DHS	755	0.01%	4487	72.01%	12.81%	0.57%	X	X	-0.152***	0.867***	-0.016	0.048*	-0.423
											(0.054)	(0.020)	(0.219)	(0.026)	(0.704)
Uruguay	1963	IPUMS-I	9,974	0.08%	4909	17.38%	44.42%	1.02%	X		-0.062***	0.583***	0.055	0.039***	0.152
											(0.008)	(0.017)	(0.067)	(0.009)	(0.200)
Uzbekistan	1996	DHS	1,275	0.01%	3223	46.16%	55.92%	0.78%	X	X	-0.147***	0.391***	0.059	0.082***	0.295
											(0.036)	(0.053)	(0.504)	(0.025)	(0.388)
Vietnam	2009	IPUMS-I	745,767	5.67%	3063	87.23%	20.74%	0.61%	X	X	-0.013***	0.792***	-0.065***	0.084***	-0.041***
											(0.001)	(0.002)	(0.010)	(0.001)	(0.012)
Yemen	2013	DHS	6,699	0.05%	3165	9.39%	69.87%	0.63%		X	-0.018*	0.207***	0.021	0.024**	-0.590
											(0.010)	(0.058)	(0.265)	(0.012)	(0.447)

	5,000 to 7,500 real GDP/Capita bin														
	Year (num. samples)	Source	N	Percent of bin	Mean real GDP/C	In labor force	3 or more children	2nd child is twin	Educa-tion?	Birth Quarter?	OLS	FS Twin IV	2S Twin IV	FS Same gender IV	2S Same gender IV
Pooled	55		17,092,548			15.57%	53.35%	0.91%			-0.047***	0.482***	0.005	0.019***	0.004
											(0.010)	(0.017)	(0.009)	(0.003)	(0.013)
Argentina	1970	IPUMS-I	19,209	0.11%	7206	16.41%	45.69%	1.46%	X		-0.047***	0.538***	-0.013	0.040***	-0.161
											(0.006)	(0.011)	(0.041)	(0.007)	(0.133)
Argentina	1991	IPUMS-I	205,654	1.20%	7173	40.36%	51.84%	1.17%	X		-0.102***	0.482***	-0.006	0.031***	-0.189**
											(0.003)	(0.005)	(0.024)	(0.002)	(0.079)
Armenia	2001	IPUMS-I	17,771	0.10%	5412	71.52%	32.50%	0.81%	X	X	-0.031***	0.678***	-0.036	0.112***	-0.122**
											(0.008)	(0.016)	(0.054)	(0.006)	(0.060)
Azerbaijan	2006	DHS	1,658	0.01%	5773	12.81%	30.58%	0.55%	X	X	-0.022	0.589***	0.181	0.179***	0.204*
											(0.023)	(0.067)	(0.264)	(0.026)	(0.119)
Belarus	1999	IPUMS-I	30,957	0.18%	6097	83.30%	13.11%	0.67%	X		-0.092***	0.866***	-0.011	0.029***	-0.096
											(0.007)	(0.005)	(0.028)	(0.004)	(0.142)
Brazil	1991	IPUMS-I	475,199	2.78%	5007	33.14%	48.87%	1.20%	X		-0.103***	0.492***	-0.052***	0.036***	-0.037
											(0.002)	(0.003)	(0.014)	(0.001)	(0.042)
Brazil	1991	DHS	1,356	0.01%	5007	44.53%	61.08%	0.60%	X	X	-0.000	0.258***	-1.492*	0.048*	-0.077
											(0.041)	(0.060)	(0.830)	(0.029)	(0.729)
Brazil	1996	DHS	2,687	0.02%	5241	46.53%	42.38%	0.77%	X	X	-0.058**	0.598***	0.214	0.009	1.814
											(0.023)	(0.056)	(0.205)	(0.020)	(4.612)
Brazil	2000	IPUMS-I	498,571	2.92%	5400	51.25%	41.49%	1.25%	X		-0.104***	0.573***	-0.052***	0.030***	-0.037
											(0.002)	(0.002)	(0.012)	(0.001)	(0.050)
Brazil	2010	IPUMS-I	392,152	2.29%	6879	58.86%	36.35%	1.44%	X		-0.105***	0.630***	-0.032***	0.028***	-0.015
											(0.002)	(0.003)	(0.012)	(0.002)	(0.065)
Chile	1970	IPUMS-I	41,509	0.24%	5241	12.37%	64.12%	1.04%	X		-0.098***	0.400***	0.049	0.029***	0.070
											(0.004)	(0.010)	(0.043)	(0.004)	(0.112)
Chile	1982	IPUMS-I	55,984	0.33%	5263	18.19%	45.85%	1.19%	X		-0.088***	0.540***	-0.014	0.026***	-0.009
											(0.004)	(0.009)	(0.027)	(0.004)	(0.122)
Chile	1992	IPUMS-I	69,678	0.41%	7416	20.33%	37.75%	1.29%	X		-0.079***	0.628***	-0.062***	0.033***	0.018
											(0.003)	(0.006)	(0.020)	(0.003)	(0.092)
Colombia	1993	IPUMS-I	168,635	0.99%	5144	28.54%	48.49%	1.53%	X		-0.126***	0.537***	0.022	0.031***	0.045
											(0.002)	(0.004)	(0.017)	(0.002)	(0.072)
Colombia	1995	DHS	2,399	0.01%	5359	45.26%	45.42%	0.60%	X	X	-0.101***	0.585***	0.420**	0.032	0.221
											(0.022)	(0.044)	(0.196)	(0.020)	(0.676)
Colombia	2000	DHS	2,317	0.01%	5473	46.76%	41.69%	0.89%	X	X	-0.105***	0.608***	-0.201	0.013	0.936
											(0.023)	(0.051)	(0.185)	(0.021)	(2.281)
Colombia	2005	IPUMS-I	185,928	1.09%	6116	33.71%	42.64%	1.43%	X	X	-0.134***	0.572***	-0.016	0.035***	-0.013
											(0.005)	(0.007)	(0.033)	(0.004)	(0.120)
Colombia	2005	DHS	7,234	0.04%	6116	49.77%	41.38%	0.74%	X	X	-0.097***	0.586***	0.010	0.028**	-0.728
											(0.017)	(0.026)	(0.143)	(0.014)	(0.635)
Colombia	2010	DHS	9,053	0.05%	7063	51.82%	35.62%	0.73%	X	X	-0.102***	0.644***	-0.080	0.025**	-0.407
											(0.015)	(0.023)	(0.114)	(0.012)	(0.563)
Costa Rica	2000	IPUMS-I	20,566	0.12%	6046	24.56%	47.31%	1.19%	X		-0.109***	0.516***	0.074	0.034***	-0.091
											(0.007)	(0.014)	(0.056)	(0.006)	(0.175)
Dominican Republic	2010	IPUMS-I	39,222	0.23%	5379	43.69%	46.09%	1.63%	X	X	-0.087***	0.514***	-0.026	0.034***	-0.044
											(0.005)	(0.008)	(0.038)	(0.005)	(0.144)
Dominican Republic	2013	DHS	1,818	0.01%	5379	50.70%	45.54%	1.29%	X	X	-0.076**	0.529***	-0.161	0.064**	-0.404
											(0.034)	(0.040)	(0.239)	(0.028)	(0.503)
Ecuador	2010	IPUMS-I	70,502	0.41%	5050	44.68%	43.74%	0.93%	X	X	-0.109***	0.558***	-0.057	0.039***	-0.048
											(0.004)	(0.008)	(0.035)	(0.003)	(0.095)
Greece	1971	IPUMS-I	35,148	0.21%	6610	22.29%	23.69%	1.31%	X		0.015***	0.771***	0.019	0.065***	-0.124*
											(0.006)	(0.006)	(0.026)	(0.004)	(0.069)
Hungary	1990	IPUMS-I	22,785	0.13%	6271	64.74%	19.27%	0.96%	X		-0.303***	0.812***	-0.098***	0.046***	-0.333***
											(0.008)	(0.007)	(0.037)	(0.005)	(0.122)
Hungary	2001	IPUMS-I	16,781	0.10%	7090	46.48%	25.85%	1.08%	X		-0.446***	0.738***	-0.109**	0.032***	-0.343
											(0.007)	(0.010)	(0.044)	(0.006)	(0.211)
Iran	2006	IPUMS-I	59,264	0.35%	5694	9.15%	35.12%	0.91%	X	X	-0.030***	0.663***	0.026	0.041***	0.142**
											(0.003)	(0.010)	(0.024)	(0.004)	(0.068)
Iran	2011	IPUMS-I	60,204	0.35%	6456	6.98%	24.91%	1.02%	X	X	-0.027***	0.745***	-0.025**	0.032***	-0.033
											(0.002)	(0.006)	(0.012)	(0.003)	(0.064)
Ireland	1971	IPUMS-I	8,860	0.05%	6426	3.63%	59.18%	1.66%	X		-0.015***	0.366***	-0.028	0.016*	0.473
											(0.005)	(0.022)	(0.037)	(0.009)	(0.380)
Jordan	2007	DHS	4,244	0.02%	5290	10.83%	68.85%	1.49%	X	X	-0.078***	0.432***	-0.052	0.064***	0.436
											(0.026)	(0.048)	(0.141)	(0.019)	(0.282)
Jordan	2009	DHS	3,774	0.02%	5585	11.92%	65.13%	1.81%	X	X	-0.023	0.341***	0.117	0.033	0.701
											(0.023)	(0.046)	(0.170)	(0.021)	(0.663)
Jordan	2012	DHS	4,169	0.02%	5647	12.75%	66.64%	1.60%	X	X	-0.091***	0.277***	-0.153	-0.016	-0.725
											(0.025)	(0.046)	(0.163)	(0.020)	(1.175)
Kazakhstan	1995	DHS	771	0.00%	5157	48.73%	35.18%	1.00%	X	X	-0.251***	0.591***	-0.206	0.022	-0.209
											(0.040)	(0.043)	(0.272)	(0.038)	(1.655)
Kazakhstan	1999	DHS	885	0.01%	5456	34.49%	36.23%	0.39%	X	X	-0.199***	0.544***	0.177	0.104***	0.309
											(0.038)	(0.058)	(0.508)	(0.035)	(0.370)
Malaysia	1991	IPUMS-I	19,157	0.11%	5502	30.39%	62.01%	1.43%	X	X	-0.101***	0.365***	-0.132*	0.020***	-0.655*
											(0.008)	(0.015)	(0.072)	(0.006)	(0.368)
Mexico	1990	IPUMS-I	453,455	2.65%	6067	15.94%	60.53%	1.05%	X		-0.113***	0.402***	0.013	0.029***	-0.059
											(0.001)	(0.004)	(0.013)	(0.001)	(0.038)
Mexico	1995	IPUMS-I	20,788	0.12%	6381	34.88%	54.16%	0.67%	X		-0.112***	0.534***	0.011	0.037***	0.002
											(0.012)	(0.028)	(0.098)	(0.010)	(0.284)
Mexico	2000	IPUMS-I	602,523	3.53%	6993	28.54%	49.86%	0.95%	X		-0.102***	0.503***	0.041***	0.032***	-0.083*
											(0.002)	(0.004)	(0.016)	(0.002)	(0.048)
Panama	2000	IPUMS-I	14,174	0.08%	5597	36.14%	50.36%	1.39%	X		-0.155***	0.493***	0.010	0.026***	0.463

											(0.009)	(0.016)	(0.070)	(0.008)	(0.353)
Panama	2010	IPUMS-I	14,272	0.08%	6675	38.38%	46.99%	1.05%	X		-0.172***	0.530***	-0.081	0.026***	-0.397
											(0.009)	(0.017)	(0.075)	(0.008)	(0.309)
Peru	2009	DHS	4,832	0.03%	5505	60.97%	41.75%	0.51%	X	X	-0.032	0.629***	-0.344*	0.001	23.032
											(0.021)	(0.034)	(0.177)	(0.017)	(435.103)
Peru	2010	DHS	4,564	0.03%	5774	60.29%	42.78%	0.96%	X	X	-0.041**	0.577***	-0.226	0.050***	0.063
											(0.021)	(0.034)	(0.214)	(0.018)	(0.380)
Peru	2011	DHS	4,448	0.03%	5774	62.32%	40.03%	0.52%	X	X	-0.020	0.532***	0.261	0.054***	0.267
											(0.021)	(0.047)	(0.208)	(0.018)	(0.368)
Peru	2012	DHS	4,588	0.03%	5774	56.39%	39.53%	0.43%	X	X	-0.068***	0.544***	-0.157	0.035**	-0.206
											(0.021)	(0.058)	(0.231)	(0.017)	(0.547)
South Africa	2011	IPUMS-I	139,743	0.82%	5080	74.28%	36.08%	2.30%	X	X	-0.066***	0.637***	-0.058***	0.013***	0.099
											(0.003)	(0.003)	(0.013)	(0.002)	(0.176)
Turkey	1990	IPUMS-I	163,770	0.96%	5333	38.55%	51.07%	1.17%	X		0.107***	0.503***	0.200***	0.066***	0.036
											(0.003)	(0.005)	(0.023)	(0.002)	(0.036)
Turkey	1993	DHS	2,349	0.01%	5648	32.95%	47.27%	0.55%	X	X	-0.077***	0.433***	0.140	0.109***	-0.172
											(0.022)	(0.055)	(0.324)	(0.019)	(0.182)
Turkey	1998	DHS	2,093	0.01%	6215	29.33%	42.32%	1.02%	X	X	-0.051**	0.554***	0.045	0.084***	-0.118
											(0.025)	(0.050)	(0.200)	(0.022)	(0.263)
Turkey	2000	IPUMS-I	180,069	1.05%	6358	38.20%	42.58%	1.36%	X		0.073***	0.601***	0.150***	0.070***	-0.013
											(0.002)	(0.004)	(0.017)	(0.002)	(0.033)
Turkey	2003	DHS	2,579	0.02%	6841	22.40%	43.09%	0.72%	X	X	-0.049**	0.674***	0.052	0.094***	0.073
											(0.021)	(0.051)	(0.164)	(0.020)	(0.203)
USA	1910	US Full Count	3,532,739	20.67%	5022	9.96%	58.00%	0.64%			-0.008***	0.421***	0.052***	0.012***	0.037
											(0.000)	(0.001)	(0.005)	(0.000)	(0.027)
USA	1920	US Full Count	4,516,473	26.42%	5595	7.76%	56.59%	0.90%			-0.034***	0.439***	0.006**	0.013***	0.022
											(0.000)	(0.001)	(0.003)	(0.000)	(0.019)
USA	1930	US Full Count	4,850,513	28.38%	5948	8.62%	53.34%	0.85%			-0.047***	0.479***	-0.002	0.018***	0.037**
											(0.000)	(0.001)	(0.003)	(0.000)	(0.015)
Uruguay	1975	IPUMS-I	10,546	0.06%	5368	24.20%	43.23%	1.09%	X		-0.082***	0.572***	-0.150**	0.050***	0.246
											(0.009)	(0.017)	(0.060)	(0.009)	(0.176)
Uruguay	1985	IPUMS-I	11,929	0.07%	5926	36.06%	42.37%	1.05%	X		-0.119***	0.583***	-0.025	0.041***	-0.278
											(0.009)	(0.017)	(0.073)	(0.008)	(0.215)

7,500 to 10,000 real GDP/Capita bin

	Year (num. samples)	Source	N	Percent of bin	Mean real GDP/C	In labor force	3 or more children	2nd child is twin	Educa- tion?	Birth Quarter?	OLS	FS Twin IV	2S Twin IV	FS Same gender IV	2S Same gender IV
Pooled	22		6,909,510			18.35%	46.06%	0.99%			-0.078***	0.543***	-0.025***	0.025***	0.042**
											(0.011)	(0.010)	(0.006)	(0.003)	(0.021)
Argentina	1980	IPUMS-I	135,408	1.96%	7826	20.41%	47.70%	1.38%	X		-0.079***	0.530***	-0.050**	0.043***	0.103
											(0.003)	(0.006)	(0.021)	(0.003)	(0.068)
Argentina	2001	IPUMS-I	150,620	2.18%	8049	49.12%	50.04%	1.22%	X		-0.116***	0.509***	-0.055**	0.023***	-0.133
											(0.003)	(0.005)	(0.023)	(0.002)	(0.110)
Armenia	2005	DHS	1,315	0.02%	8617	21.54%	25.53%	0.89%	X	X	-0.061*	0.851***	-0.204***	0.117***	0.120
											(0.035)	(0.072)	(0.058)	(0.028)	(0.245)
Costa Rica	2011	IPUMS-I	17,905	0.26%	7997	34.86%	34.15%	0.99%	X		-0.096***	0.656***	0.009	0.033***	0.056
											(0.008)	(0.014)	(0.055)	(0.007)	(0.212)
France	1962	IPUMS-I	92,331	1.34%	8073	20.32%	49.31%	2.68%	X		-0.124***	0.519***	-0.103***	0.026***	-0.173*
											(0.003)	(0.004)	(0.014)	(0.003)	(0.100)
Greece	1981	IPUMS-I	45,467	0.66%	8897	21.31%	23.95%	1.19%	X	X	-0.024***	0.761***	-0.011	0.063***	-0.046
											(0.005)	(0.005)	(0.023)	(0.004)	(0.060)
Hungary	2011	IPUMS-I	9,789	0.14%	8353	47.65%	28.69%	1.09%	X		-0.397***	0.699***	-0.189***	0.022***	-0.171
											(0.010)	(0.017)	(0.059)	(0.008)	(0.414)
Ireland	1981	IPUMS-I	13,484	0.20%	8641	8.93%	53.24%	1.28%	X		-0.070***	0.456***	0.031	0.040***	0.126
											(0.006)	(0.017)	(0.051)	(0.008)	(0.128)
Ireland	1986	IPUMS-I	12,809	0.19%	9597	16.74%	50.61%	1.12%			-0.100***	0.481***	-0.039	0.058***	-0.105
											(0.007)	(0.020)	(0.062)	(0.008)	(0.112)
Malaysia	2000	IPUMS-I	20,415	0.30%	7759	34.08%	57.88%	1.66%	X		-0.080***	0.462***	0.208***	0.028***	-0.680***
											(0.008)	(0.014)	(0.056)	(0.006)	(0.264)
Mexico	2010	IPUMS-I	644,670	9.33%	7716	33.66%	43.39%	0.94%	X		-0.111***	0.582***	-0.004	0.030***	0.082
											(0.003)	(0.006)	(0.020)	(0.002)	(0.089)
Mexico	2015	IPUMS-I	584,788	8.46%	7716	32.78%	40.69%	1.01%	X		-0.109***	0.596***	-0.019	0.033***	-0.030
											(0.003)	(0.005)	(0.018)	(0.002)	(0.069)
Poland	2002	IPUMS-I	115,456	1.67%	7683	76.94%	27.24%	1.00%	X	X	-0.110***	0.729***	-0.057***	0.028***	-0.067
											(0.003)	(0.004)	(0.018)	(0.003)	(0.086)
Portugal	1981	IPUMS-I	19,031	0.28%	7979	46.28%	28.97%	1.02%	X		-0.141***	0.703***	-0.045	0.043***	0.252
											(0.008)	(0.011)	(0.051)	(0.006)	(0.174)
Puerto Rico	1980	IPUMS-PR	8,246	0.12%	7918	35.07%	51.75%	1.84%	X	X	-0.167***	0.464***	-0.062	0.048***	-0.191
											(0.011)	(0.018)	(0.082)	(0.010)	(0.216)
USA	1940	US Full Count	4,621,433	66.89%	7942	10.61%	47.15%	0.86%	X		-0.064***	0.539***	-0.016***	0.021***	0.071***
											(0.000)	(0.001)	(0.003)	(0.000)	(0.014)
USA	1950	IPUMS-USA	103,494	1.50%	9643	14.02%	43.10%	1.02%	X		-0.079***	0.588***	-0.042***	0.024***	0.117
											(0.002)	(0.006)	(0.015)	(0.003)	(0.093)
Uruguay	1996	IPUMS-I	11,642	0.17%	8086	54.76%	39.92%	1.22%	X		-0.116***	0.584***	-0.019	0.029***	-0.195
											(0.010)	(0.017)	(0.071)	(0.008)	(0.311)
Uruguay	2006	IPUMS-I	9,121	0.13%	9084	62.78%	40.98%	1.24%	X		-0.148***	0.563***	-0.076	0.027**	-0.306
											(0.013)	(0.028)	(0.100)	(0.011)	(0.459)
Venezuela	1981	IPUMS-I	80,451	1.16%	9827	26.08%	60.94%	2.36%	X		-0.134***	0.380***	-0.012	0.029***	0.062
											(0.004)	(0.005)	(0.026)	(0.003)	(0.106)
Venezuela	1990	IPUMS-I	98,117	1.42%	8785	32.08%	56.01%	2.35%	X		-0.152***	0.427***	-0.075***	0.030***	-0.157
											(0.004)	(0.005)	(0.026)	(0.003)	(0.113)
Venezuela	2001	IPUMS-I	113,518	1.64%	8138	33.54%	49.48%	1.45%	X		-0.132***	0.518***	-0.043*	0.035***	0.064
											(0.003)	(0.005)	(0.022)	(0.003)	(0.081)

	Year (num. samples)	Source	N	Percent of bin	Mean real GDP/C	In labor force	3 or more children	2nd child is twin	Educa-tion?	Birth Quarter?	OLS	FS Twin IV	2S Twin IV	FS Same gender IV	2S Same gender IV
	10,000 to 15,000 real GDP/Capita bin														
Pooled	20		1,084,881			31.37%	44.47%	1.41%			-0.126***	0.525***	-0.063***	0.035***	-0.067***
											(0.015)	(0.048)	(0.017)	(0.002)	(0.021)
Armenia	2010	DHS	1,178	0.11%	10215	22.47%	19.86%	0.73%	X	X	-0.062	0.804***	0.076	0.128***	-0.235
											(0.038)	(0.040)	(0.210)	(0.025)	(0.230)
Armenia	2011	IPUMS-I	15,059	1.39%	10215	47.35%	22.72%	0.93%	X	X	-0.013	0.787***	-0.112**	0.107***	-0.088
											(0.010)	(0.013)	(0.052)	(0.006)	(0.076)
Austria	1971	IPUMS-I	30,982	2.86%	10195	34.31%	40.88%	1.04%	X		-0.076***	0.593***	-0.056	0.026***	-0.060
											(0.006)	(0.010)	(0.044)	(0.005)	(0.210)
Austria	1981	IPUMS-I	27,991	2.58%	13779	43.65%	29.86%	1.00%	X		-0.102***	0.697***	-0.144***	0.042***	-0.253*
											(0.007)	(0.008)	(0.041)	(0.005)	(0.140)
Belarus	2009	IPUMS-I	22,000	2.03%	12992	78.71%	14.59%	0.88%	X		-0.138***	0.854***	-0.036	0.021***	-0.074
											(0.008)	(0.005)	(0.034)	(0.005)	(0.257)
Chile	2002	IPUMS-I	56,760	5.23%	10777	31.44%	31.07%	0.94%	X		-0.081***	0.688***	-0.044	0.026***	-0.187
											(0.004)	(0.007)	(0.028)	(0.004)	(0.149)
France	1968	IPUMS-I	95,250	8.78%	10432	24.54%	46.56%	1.05%	X		-0.153***	0.539***	-0.084***	0.033***	-0.104
											(0.003)	(0.006)	(0.024)	(0.003)	(0.082)
France	1975	IPUMS-I	103,331	9.52%	13254	36.94%	38.92%	1.13%	X		-0.249***	0.607***	-0.172***	0.026***	0.088
											(0.003)	(0.006)	(0.021)	(0.003)	(0.120)
Greece	1991	IPUMS-I	40,657	3.75%	10062	37.03%	21.80%	1.22%	X	X	-0.080***	0.781***	-0.054**	0.059***	-0.035
											(0.006)	(0.005)	(0.027)	(0.004)	(0.081)
Greece	2001	IPUMS-I	28,882	2.66%	12660	51.60%	20.43%	1.13%	X		-0.070***	0.801***	-0.086**	0.042***	0.038
											(0.007)	(0.006)	(0.034)	(0.005)	(0.139)
Ireland	1991	IPUMS-I	10,937	1.01%	11843	31.29%	45.71%	1.24%	X		-0.145***	0.550***	-0.096	0.060***	-0.279*
											(0.010)	(0.021)	(0.068)	(0.008)	(0.146)
Portugal	1991	IPUMS-I	15,987	1.47%	10872	63.32%	22.76%	1.15%	X		-0.184***	0.771***	-0.046	0.021***	0.120
											(0.010)	(0.009)	(0.047)	(0.006)	(0.375)
Portugal	2001	IPUMS-I	11,704	1.08%	13831	74.49%	16.77%	1.13%	X		-0.144***	0.866***	-0.061	0.026***	-0.559*
											(0.012)	(0.010)	(0.045)	(0.007)	(0.330)
Portugal	2011	IPUMS-I	8,445	0.78%	14279	80.69%	17.18%	1.35%	X		-0.164***	0.851***	-0.017	0.025***	-0.225
											(0.013)	(0.011)	(0.042)	(0.008)	(0.331)
Puerto Rico	1990	IPUMS-PR	8,442	0.78%	10477	41.70%	47.01%	1.42%	X		-0.148***	0.509***	-0.096	0.055***	0.011
											(0.012)	(0.018)	(0.089)	(0.011)	(0.204)
Puerto Rico	2000	IPUMS-PR	7,809	0.72%	13881	43.14%	40.70%	1.41%	X		-0.106***	0.561***	-0.194**	0.042***	-0.458
											(0.013)	(0.020)	(0.084)	(0.011)	(0.283)
Spain	1991	IPUMS-I	59,957	5.53%	12030	40.02%	23.21%	1.07%	X		-0.112***	0.768***	-0.095***	0.045***	-0.051
											(0.005)	(0.006)	(0.024)	(0.003)	(0.088)
USA	1960	IPUMS-USA	470,378	43.36%	11380	22.85%	55.09%	1.70%	X	X	-0.117***	0.452***	-0.033***	0.035***	-0.084**
											(0.001)	(0.002)	(0.010)	(0.001)	(0.034)
Uruguay	2011	IPUMS-I	10,012	0.92%	11526	65.74%	36.49%	0.88%	X	X	-0.142***	0.628***	-0.015	0.026***	-0.478
											(0.011)	(0.020)	(0.080)	(0.009)	(0.380)
Venezuela	1971	IPUMS-I	59,120	5.45%	10429	15.96%	70.48%	2.28%	X		-0.083***	0.289***	-0.043	0.017***	0.416**
											(0.004)	(0.006)	(0.034)	(0.003)	(0.207)

15,000 to 20,000 real GDP/Capita bin

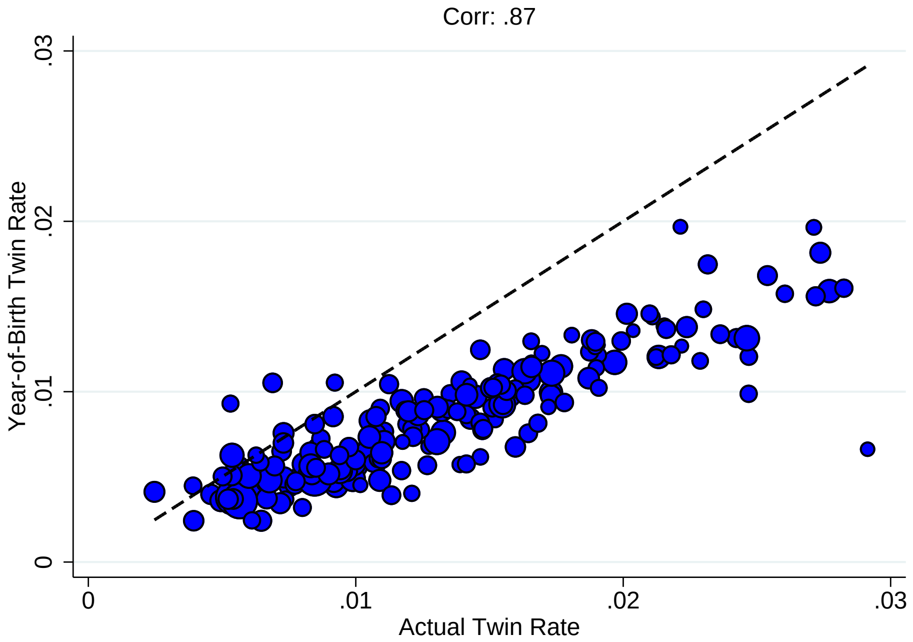
	Year (num. samples)	Source	N	Percent of bin	Mean real GDP/C	In labor force	3 or more children	2nd child is twin	Educa-tion?	Birth Quarter?	OLS	FS Twin IV	2S Twin IV	FS Same gender IV	2S Same gender IV
Pooled	12		1,107,326			49.12%	37.03%	1.26%			-0.203***	0.627***	-0.082***	0.046***	-0.145***
											(0.028)	(0.037)	(0.021)	(0.004)	(0.017)
Austria	1991	IPUMS-I	28,036	2.53%	16956	51.39%	24.72%	0.93%	X		-0.117***	0.763***	-0.136***	0.036***	-0.232
											(0.007)	(0.008)	(0.040)	(0.005)	(0.167)
France	1982	IPUMS-I	117,660	10.63%	15076	52.24%	33.51%	1.08%	X		-0.339***	0.663***	-0.212***	0.041***	-0.243***
											(0.003)	(0.005)	(0.020)	(0.003)	(0.068)
France	1990	IPUMS-I	91,261	8.24%	17309	64.14%	34.07%	1.04%	X		-0.358***	0.656***	-0.207***	0.042***	-0.160**
											(0.003)	(0.006)	(0.025)	(0.003)	(0.072)
France	1999	IPUMS-I	86,473	7.81%	19690	68.14%	29.60%	1.24%	X		-0.279***	0.706***	-0.061***	0.039***	-0.203***
											(0.004)	(0.006)	(0.020)	(0.003)	(0.076)
Great Britain	1991	IPUMS-I	20,003	1.81%	16403	46.22%	32.15%	1.11%			-0.221***	0.705***	-0.160***	0.079***	-0.232***
											(0.008)	(0.012)	(0.045)	(0.006)	(0.086)
Ireland	1996	IPUMS-I	9,165	0.83%	15683	43.06%	39.77%	1.16%	X		-0.172***	0.634***	-0.066	0.064***	-0.217
											(0.011)	(0.019)	(0.076)	(0.009)	(0.156)
Puerto Rico	2010	IPUMS-PR	4,397	0.40%	15074	57.14%	36.00%	1.39%	X	X	-0.159***	0.635***	-0.150	0.064***	-0.070
											(0.018)	(0.029)	(0.106)	(0.014)	(0.243)
Spain	2001	IPUMS-I	34,927	3.15%	15874	51.25%	16.22%	2.31%	X	X	-0.066***	0.882***	-0.025	0.034***	-0.072
											(0.007)	(0.003)	(0.020)	(0.004)	(0.156)
Switzerland	1970	IPUMS-I	11,998	1.08%	16668	21.80%	35.64%	0.81%	X		-0.083***	0.655***	-0.075	0.019**	-0.230
											(0.008)	(0.016)	(0.058)	(0.008)	(0.403)
Switzerland	1980	IPUMS-I	11,241	1.02%	18315	28.42%	23.09%	0.70%	X		-0.079***	0.789***	-0.167***	0.042***	-0.339*
											(0.010)	(0.011)	(0.048)	(0.008)	(0.202)
USA	1970	IPUMS-USA	186,891	16.88%	15334	33.17%	52.60%	1.40%	X	X	-0.137***	0.463***	-0.007	0.037***	-0.101*
											(0.002)	(0.004)	(0.020)	(0.002)	(0.057)
USA	1980	IPUMS-USA	505,274	45.63%	18487	49.31%	36.51%	1.27%	X	X	-0.177***	0.621***	-0.076***	0.053***	-0.127***
											(0.001)	(0.002)	(0.010)	(0.001)	(0.026)

20,000 to 35,000 real GDP/Capita bin

	Year (num. samples)	Source	N	Percent of bin	Mean real GDP/C	In labor force	3 or more children	2nd child is twin	Educa-tion?	Birth Quarter?	OLS	FS Twin IV	2S Twin IV	FS Same gender IV	2S Same gender IV
Pooled	12		2,397,575			67.77%	33.18%	1.45%			-0.191***	0.668***	-0.086***	0.044***	-0.140***
											(0.024)	(0.016)	(0.008)	(0.003)	(0.015)
Austria	2001	IPUMS-I	24,022	1.00%	20997	72.72%	23.55%	1.00%	X		-0.127***	0.782***	-0.153***	0.041***	-0.200
											(0.007)	(0.008)	(0.041)	(0.005)	(0.140)
Canada	2011	IPUMS-I	19,894	0.83%	24941	69.05%	29.17%	2.13%	X		-0.152***	0.686***	-0.169***	0.045***	-0.124
											(0.009)	(0.008)	(0.039)	(0.007)	(0.157)
France	2006	IPUMS-I	510,203	21.28%	21540	73.34%	28.81%	1.43%	X		-0.263***	0.707***	-0.100***	0.037***	-0.210***
											(0.002)	(0.002)	(0.008)	(0.001)	(0.034)
France	2011	IPUMS-I	485,266	20.24%	21477	76.23%	29.28%	1.46%	X		-0.248***	0.702***	-0.105***	0.038***	-0.156***
											(0.002)	(0.003)	(0.008)	(0.001)	(0.032)
Ireland	2002	IPUMS-I	7,664	0.32%	22315	45.76%	35.43%	1.55%	X		-0.180***	0.663***	-0.159**	0.037***	-0.097
											(0.013)	(0.018)	(0.067)	(0.010)	(0.300)
Ireland	2006	IPUMS-I	8,025	0.33%	24076	55.64%	32.77%	1.37%	X		-0.182***	0.681***	0.035	0.047***	-0.128
											(0.013)	(0.018)	(0.070)	(0.010)	(0.231)
Ireland	2011	IPUMS-I	10,654	0.44%	22013	61.96%	33.96%	1.40%	X		-0.176***	0.680***	-0.188***	0.048***	0.172
											(0.011)	(0.013)	(0.059)	(0.009)	(0.200)
Switzerland	1990	IPUMS-I	10,612	0.44%	20699	38.74%	26.71%	1.05%	X		-0.116***	0.751***	-0.022	0.043***	-0.274
											(0.011)	(0.012)	(0.058)	(0.008)	(0.213)
Switzerland	2000	IPUMS-I	8,685	0.36%	22122	61.04%	26.09%	1.01%	X		-0.152***	0.762***	-0.165**	0.043***	0.143
											(0.012)	(0.016)	(0.069)	(0.009)	(0.244)
USA	1990	IPUMS-USA	505,189	21.07%	22901	60.60%	35.68%	1.28%	X		-0.166***	0.647***	-0.084***	0.051***	-0.134***
											(0.002)	(0.002)	(0.011)	(0.001)	(0.030)
USA	2000	IPUMS-USA	438,854	18.30%	28100	62.82%	36.49%	1.58%	X		-0.136***	0.638***	-0.073***	0.049***	-0.102***
											(0.002)	(0.002)	(0.010)	(0.002)	(0.033)
USA	2010	IPUMS-USA	368,507	15.37%	30491	66.16%	38.14%	1.57%	X	X	-0.141***	0.622***	-0.049***	0.048***	-0.125***
											(0.002)	(0.003)	(0.012)	(0.002)	(0.040)

Notes: This table displays summary statistics, OLS results, twin IV results, and same sex IV results for each of the surveys included in our analysis. Each sample is restricted according to the rules: all two-child mothers aged 21 to 35 that were at least 15 when they had their first child, their oldest child is younger than 18, they do not live in group quarters, their first child is not a multiple birth, and mother and child have no imputations on age and gender. A twin is defined as the second and third birth being the same age. Regressions control for mother's age, age at first birth, gender of first child (and second child for same gender IV), and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level.

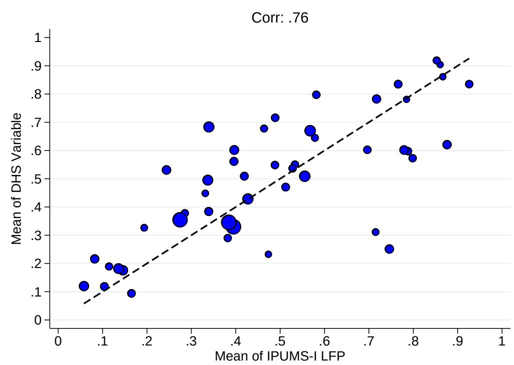
Figure A1: Comparison of twinning rates in DHS



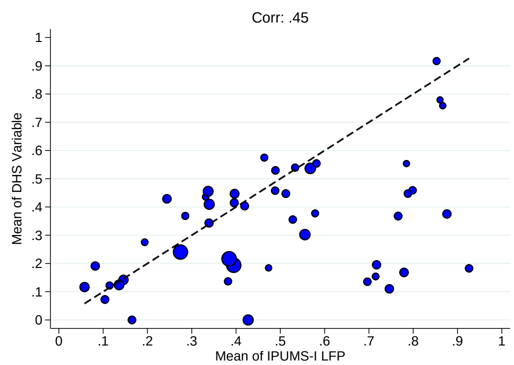
Notes: This figure plots the twin rate when twins are identified using the survey variable that indicates twinning (x-axis) against the twin rate when twins are identified as two children born to the same mother in the same year (y-axis). Each observation is a DHS survey.

Figure A2: Comparison of DHS work measures with IPUMS-International LFP

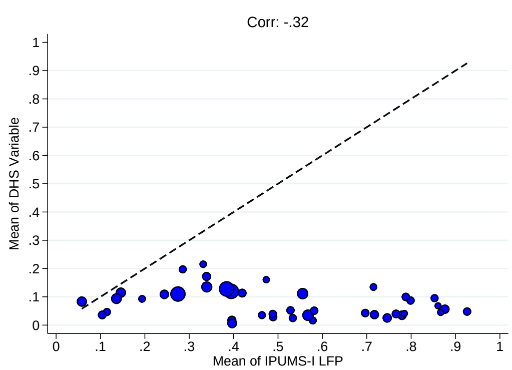
(a) Any current work



(b) Any current work for cash

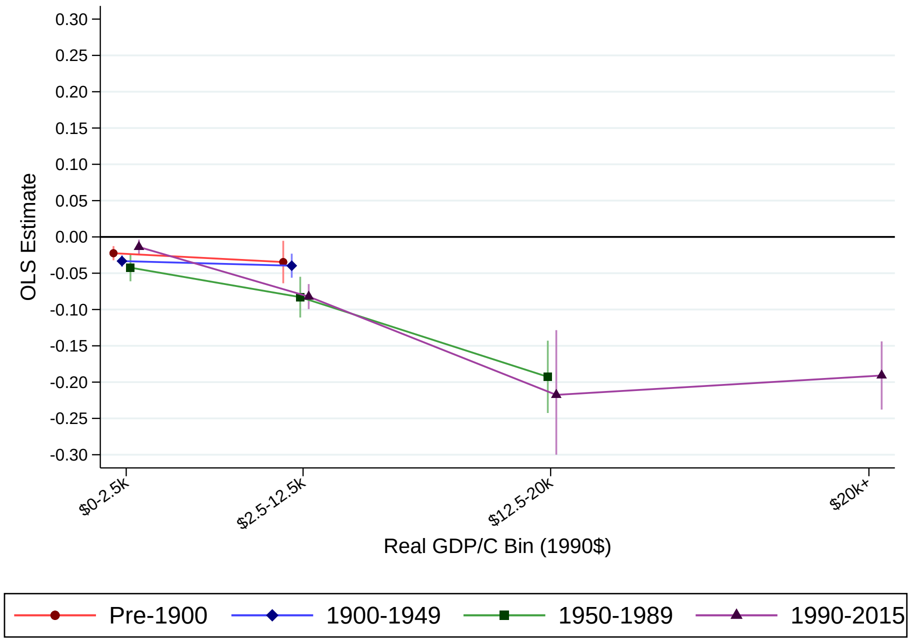


(c) Any current work for cash away from home



Notes: These figures plot a measure of female labor force participation from IPUMS-I against various measures of female LFP taken from DHS. Each observation is a country that has IPUMS-I and DHS data in the same year. Dot sizes correspond to the square root of the number of mothers observed in the IPUMS-I sample.

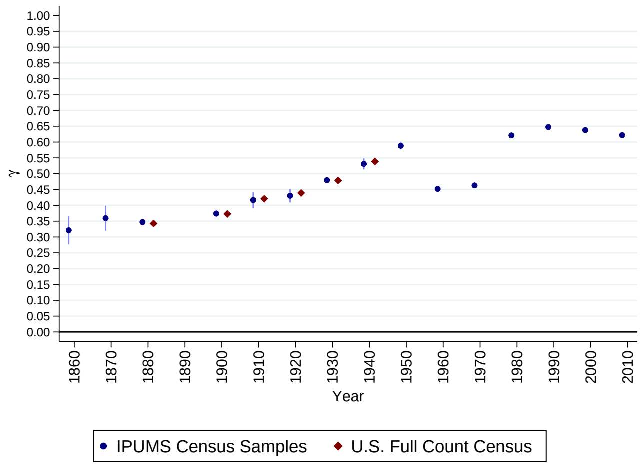
Figure A3: OLS, by time and real GDP/capita bin



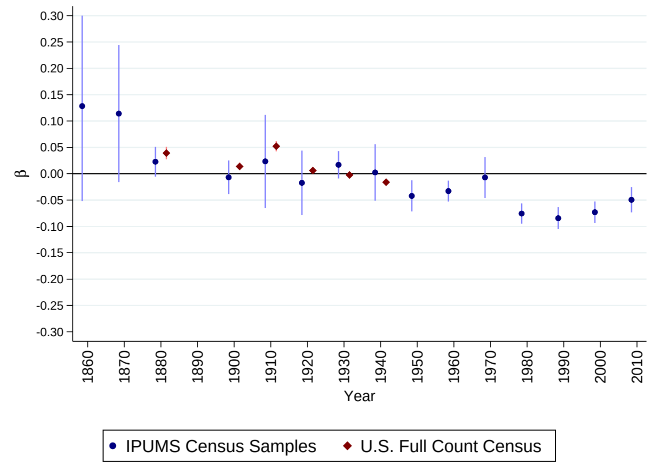
Notes: This figure displays OLS estimates of the relationship between having a third birth and mothers' labor force participation, stratified by time period. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A4: Twin IV, U.S. by time

(a) First Stage: Third Birth



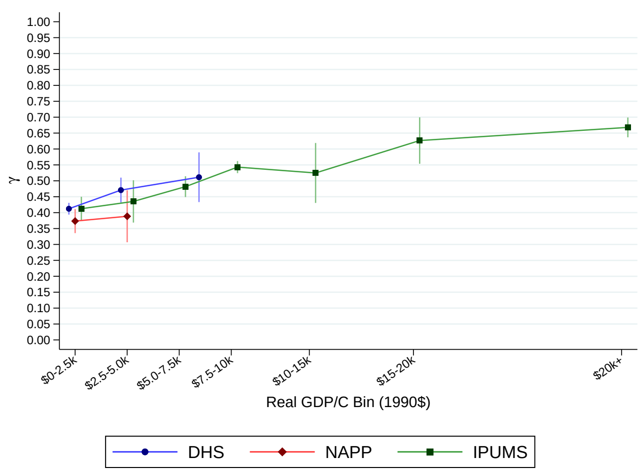
(b) Second Stage: Labor Force Participation



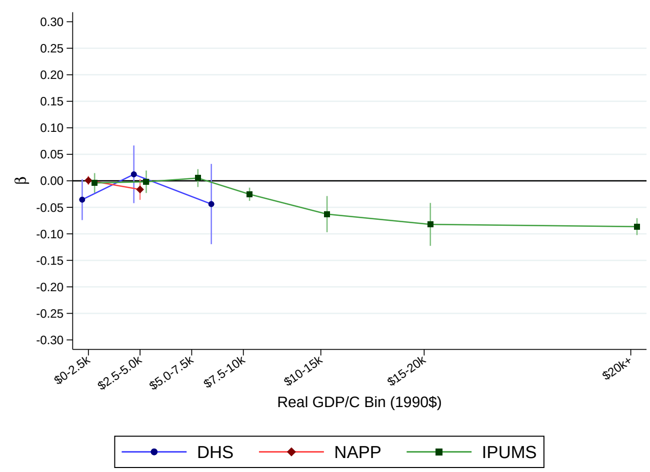
Notes: This figure displays twin IV estimates, binned by census year. It uses the baseline sample of mothers for the US only. Regressions control for mother's age, age at first birth, and gender of first child. Standard errors are robust to heteroskedasticity. 95 percent confidence intervals based on robust standard errors clustered at the year level are displayed but may not always be visible at the scale of the figure.

Figure A5: Twin IV by data source

(a) First Stage: Third Birth



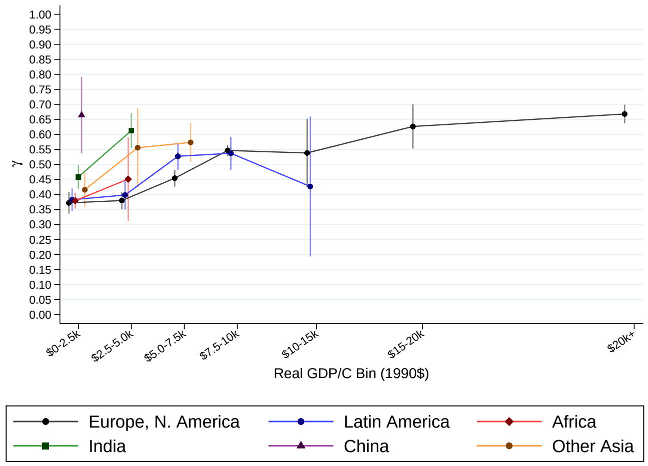
(b) Second Stage: Labor Force Participation



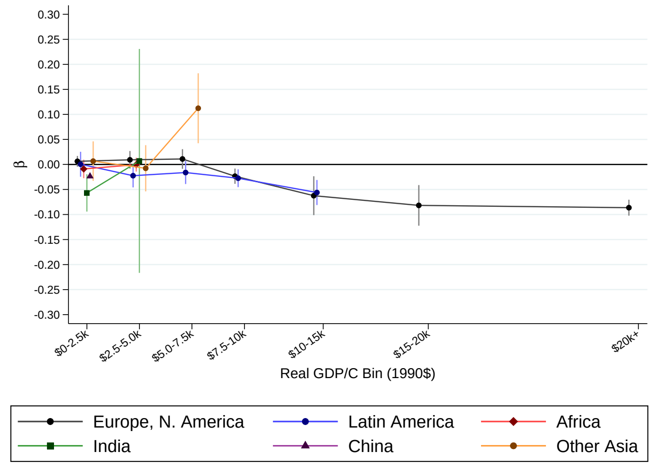
Notes: This figure displays twin IV estimates, binned by data source. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A6: Twins IV by region

(a) First Stage: Third Birth



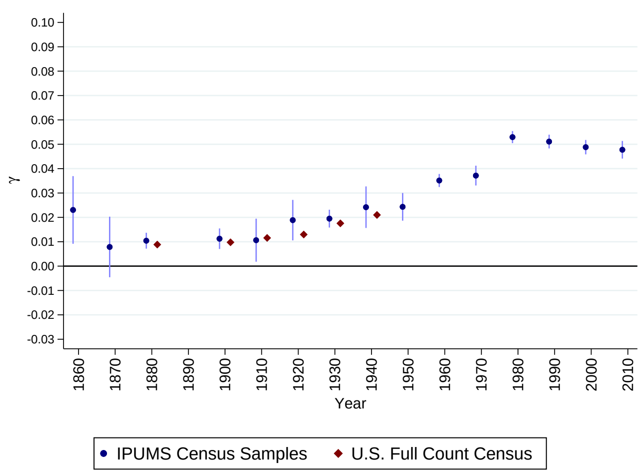
(b) Second Stage: Labor Force Participation



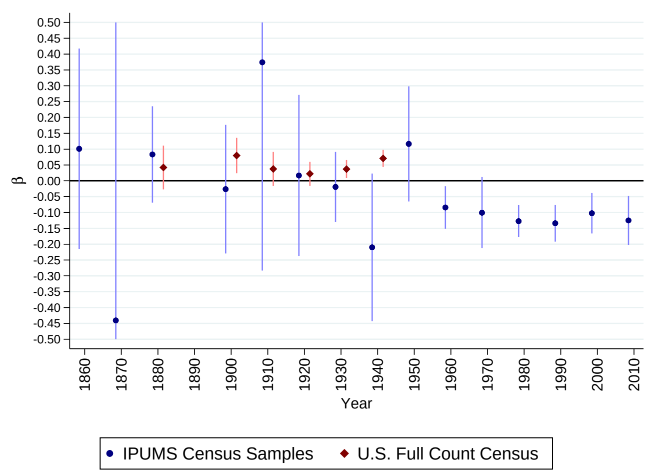
Notes: This figure displays twin IV estimates, binned by world region. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A7: Same gender IV, U.S. by time

(a) First Stage: Third Birth



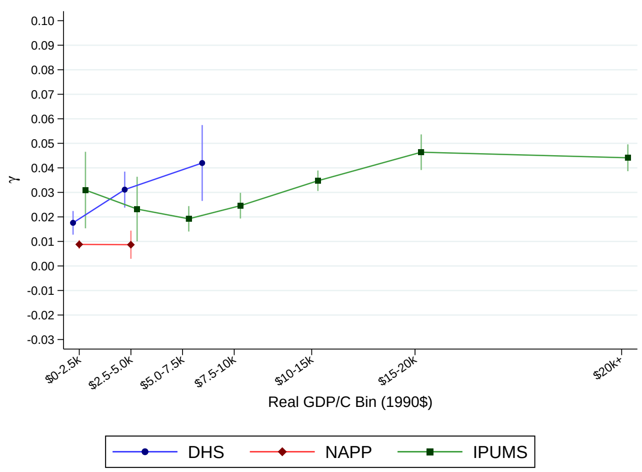
(b) Second Stage: Labor Force Participation



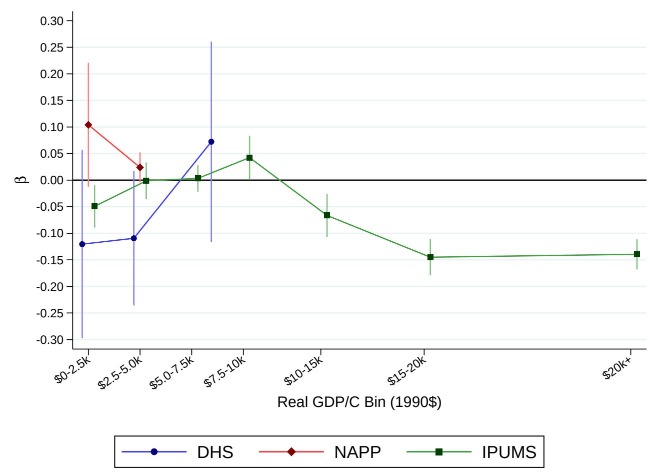
Notes: This figure displays same sex IV estimates, binned by census year. It uses the baseline sample of mothers for the US only. Regressions control for mother's age, age at first birth, and gender of first child. Standard errors are robust to heteroskedasticity. 95 percent confidence intervals based on robust standard errors clustered at the year level are displayed but may not always be visible at the scale of the figure.

Figure A8: Same Sex IV by data source

(a) First Stage: Third Birth



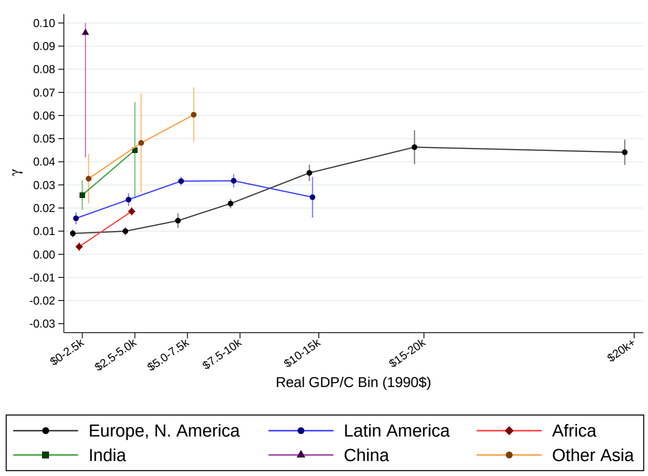
(b) Second Stage: Labor Force Participation



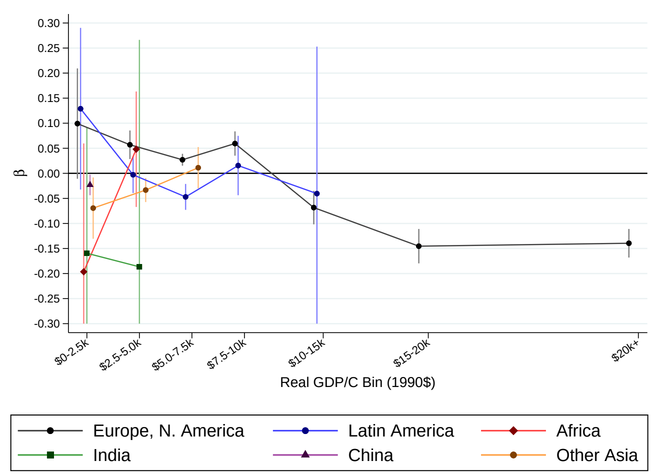
Notes: This figure displays same sex IV estimates, stratified by data source. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A9: Same Sex IV by region

(a) First Stage: Third Birth



(b) Second Stage: Labor Force Participation

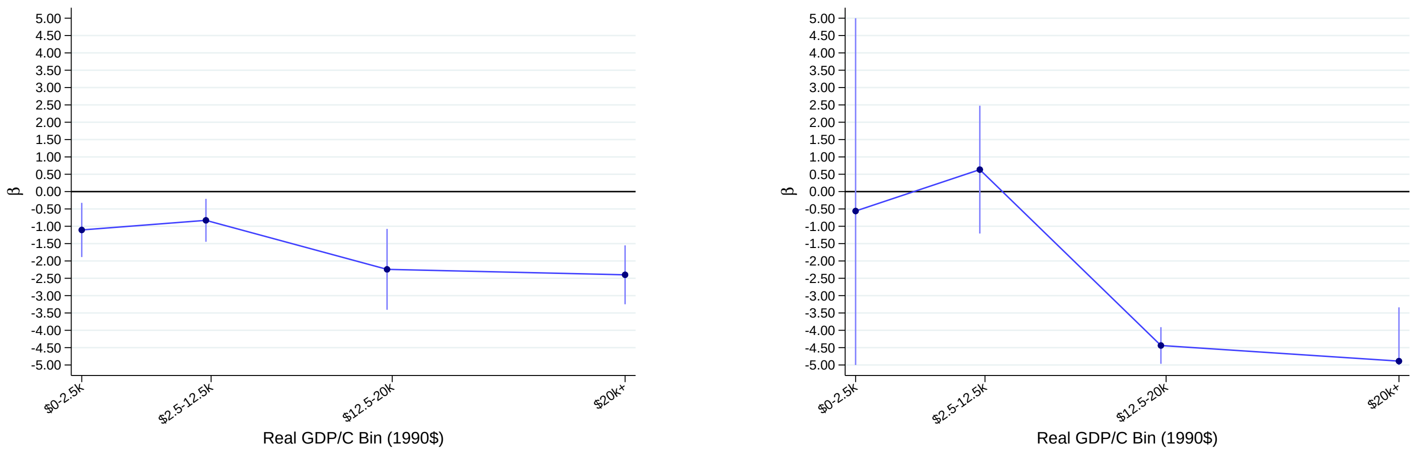


Notes: This figure displays same sex IV estimates, stratified by world region. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A10: Twin and Same Sex IV estimates of hours

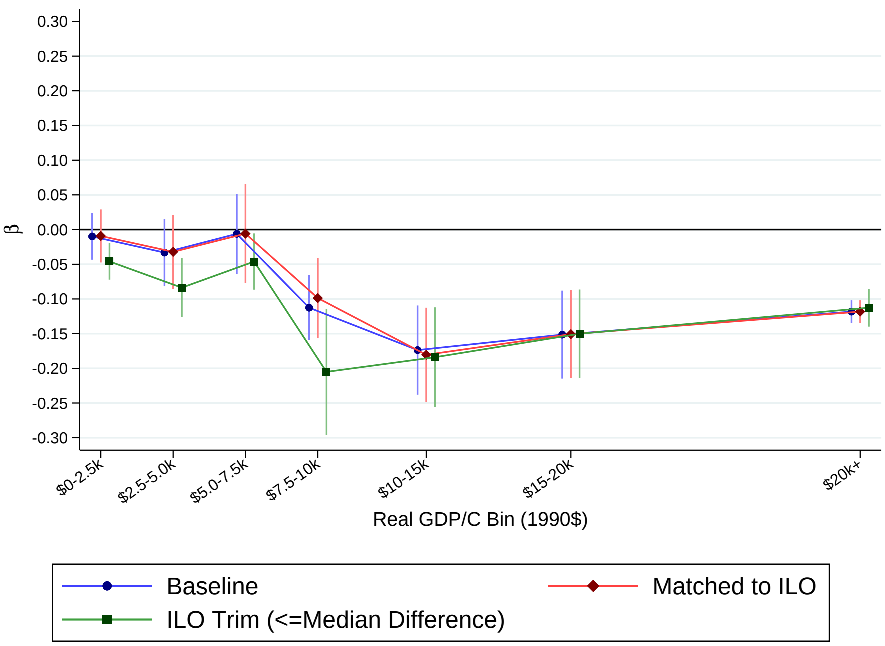
(a) Twin IV

(b) Same Sex IVl



Notes: This figure displays twin (panel A) and same sex (panel B) IV estimates using hours worked not conditional on working as the outcome variable. Surveys that do not report information on hours worked are excluded. When hours are reported in ranges, we take the median point of the range. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

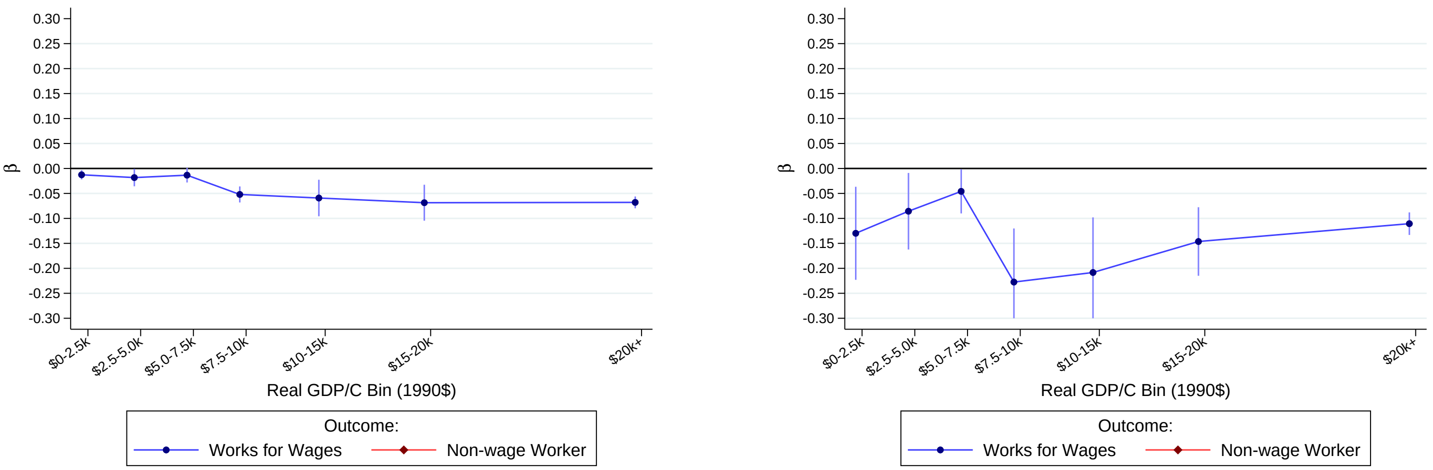
Figure A11: Twin IV estimates, rescaled by the complier-control outcome mean, ILO comparison



Notes: This figure rescales the baseline twin IV estimates by the complier-control mean of mothers' labor force status. We compare the baseline result to the results using only surveys that match ILO measures of female LFP. We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A12: Twin IV estimates by class of worker, median ILO trim

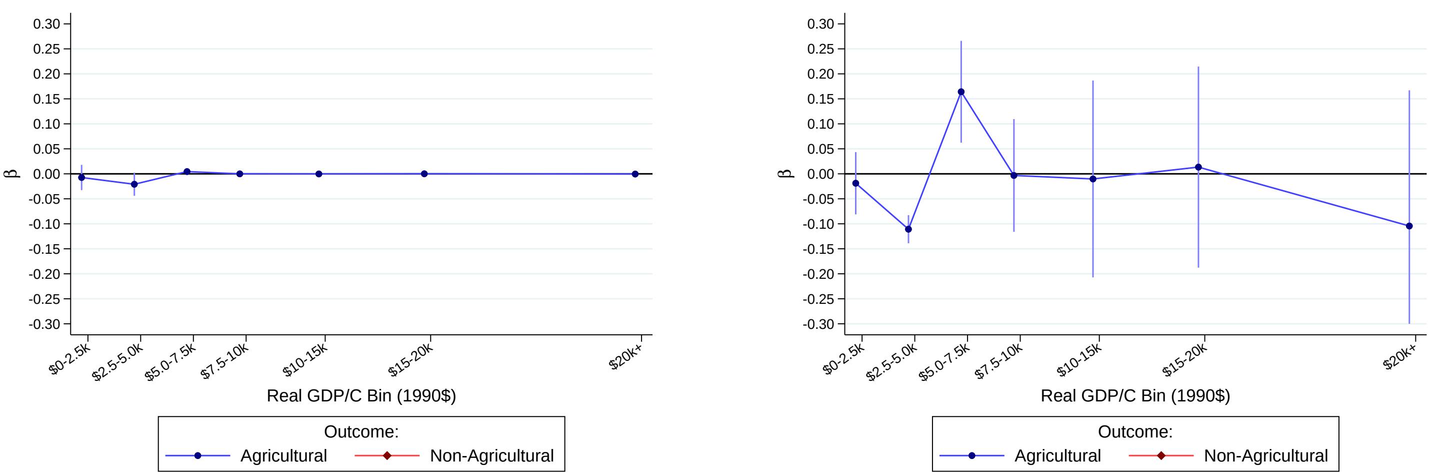
(a) Unscaled | (b) Scaled by complier-control mean



Notes: This figure displays twin IV estimates, unscaled (panel A) and scaled by the complier-control mean (panel B). The outcome for the blue line (circles) is an indicator of whether the mother works for wages. The outcome for the red line (triangles) an indicator of whether a mother works but not for wages. The sample is restricted to the set of mothers with nonmissing data on wage work and held constant across panels. We further restrict to surveys that match ILO measures of female LFP (diff<= 4.8). We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A13: Twin IV estimates by agricultural occupation, median ILO trim

(a) Unscaled | (b) Scaled by complier-control mean

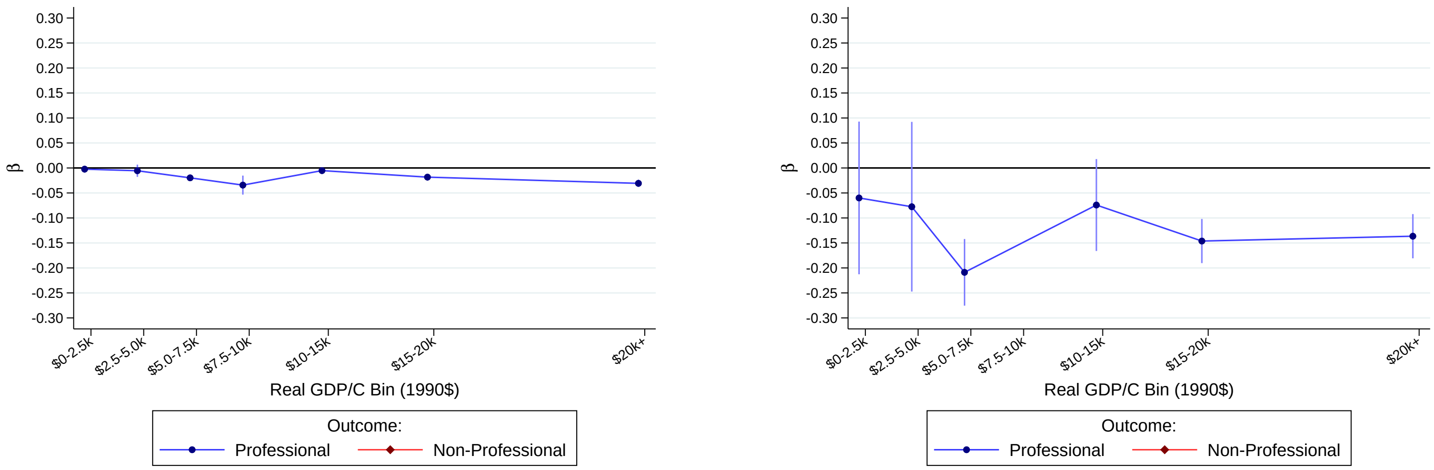


Notes: This figure displays IV estimates unscaled (panel A) and scaled by the complier-control mean (panel B). The outcome for the blue line (circles) is an indicator of whether the mother works in agriculture (defined as a farm laborer, tenant, manager, or owner). The outcome for the red line (triangles) is an indicator of whether a mother works but not in agriculture. The sample is restricted to the set of mothers with nonmissing data on wage work and held constant across panels. We further restrict to surveys that match ILO measures of female LFP (diff<= 4.8). We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A14: Twin IV estimates by professional occupation, median ILO trim

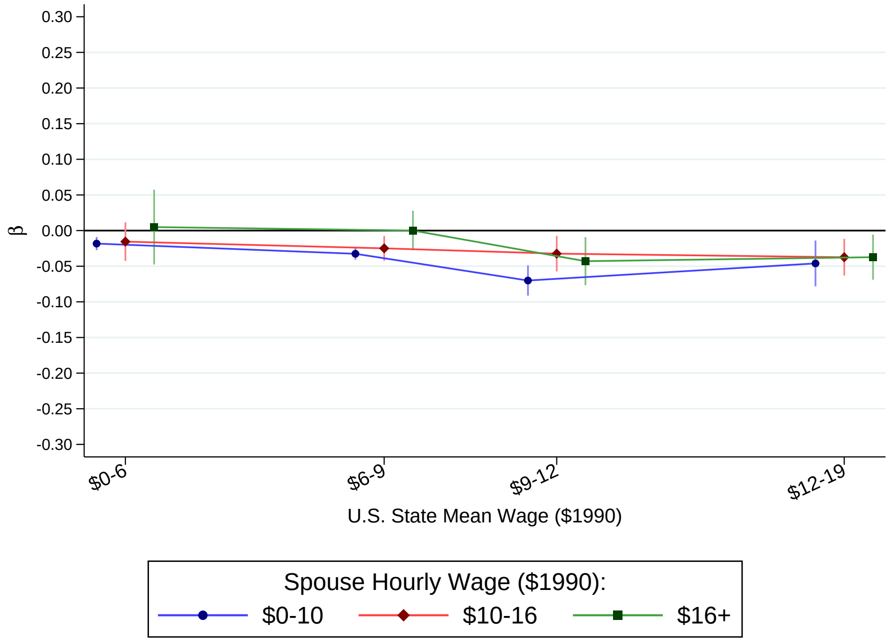
(a) Unscaled

(b) Scaled by complier-control mean



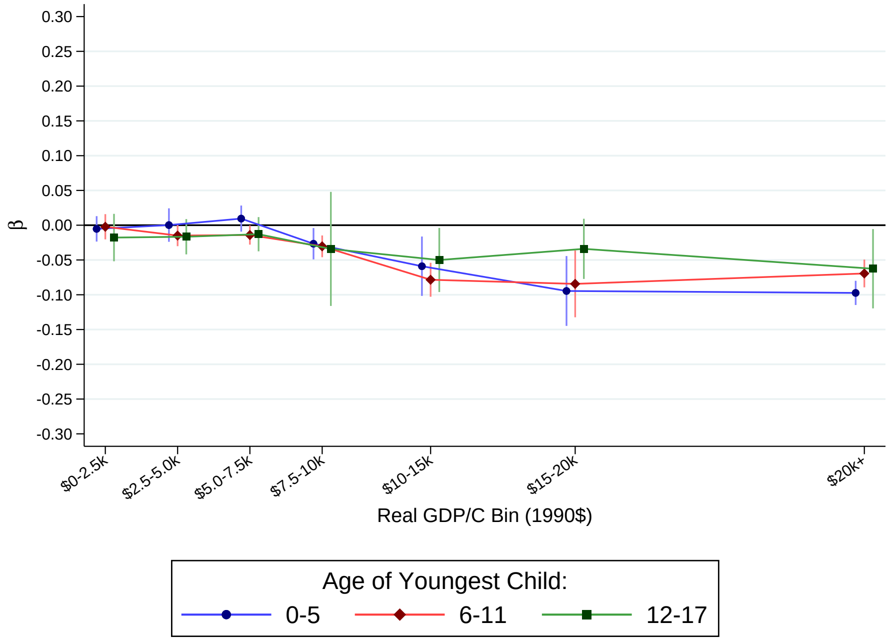
Notes: This figure displays IV estimates unscaled (panel A) and scaled by the complier-control mean (panel B). The outcome for the blue line (circles) is an indicator of whether the mother works in a professional occupation. The outcome for the red line (triangles) is an indicator of whether a mother works but not in a professional occupation The sample is restricted to the set of mothers with nonmissing data on wage work and held constant across panels. We further restrict to surveys that match ILO measures of female LFP (diff<= 4.8). We exclude U.S. samples prior to 1920 since these surveys often exhibit strongly positive labor supply responses. The calculation of the complier-control mean follows the IV methodology of Angrist, Pathak, and Walters (2013). To get standard errors, unscaled coefficients and the complier-control mean are calculated in a seemingly unrelated regression framework and the standard errors of the ratio of the unscaled estimate to the control mean are calculated via the delta method. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A15: Twin IV estimates, by state and husband wage, U.S. 1940-2010



Notes: This figure displays twin IV estimates, binned by state average wage and stratified by husband's wage. It uses the sample of US mothers with husband wage information available. Regressions control for mother's age, age at first birth, gender of first child, and year fixed effects. Year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the year level. 95 percent confidence intervals based on robust standard errors clustered at the year level are displayed but may not always be visible at the scale of the figure.

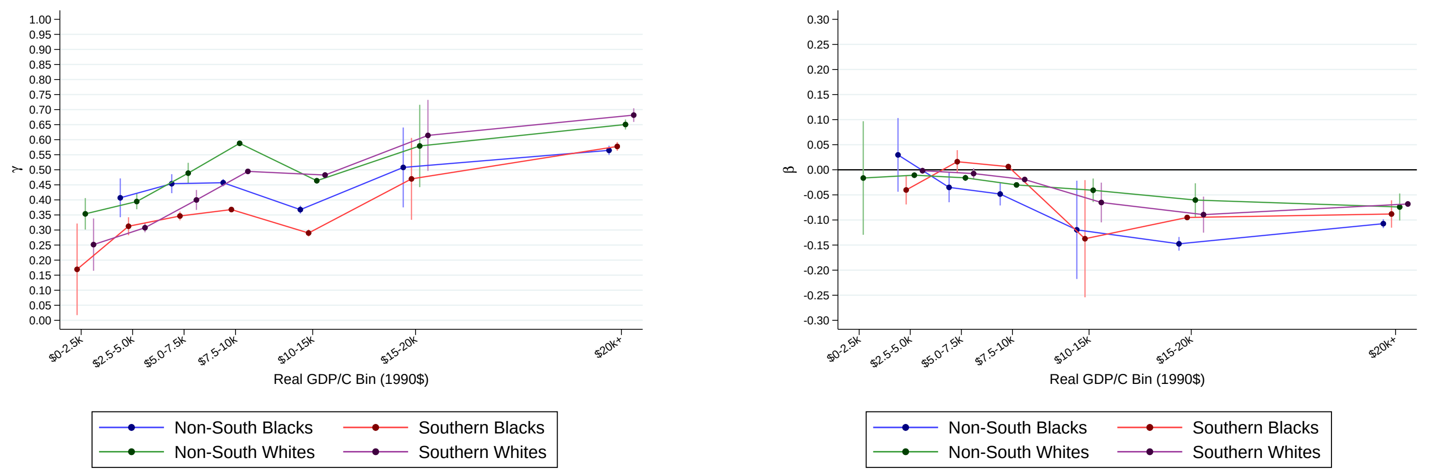
Figure A16: Twin IV estimates by the age of the youngest child



Notes: This figure displays twin IV estimates, stratified by age of the youngest child at observation. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

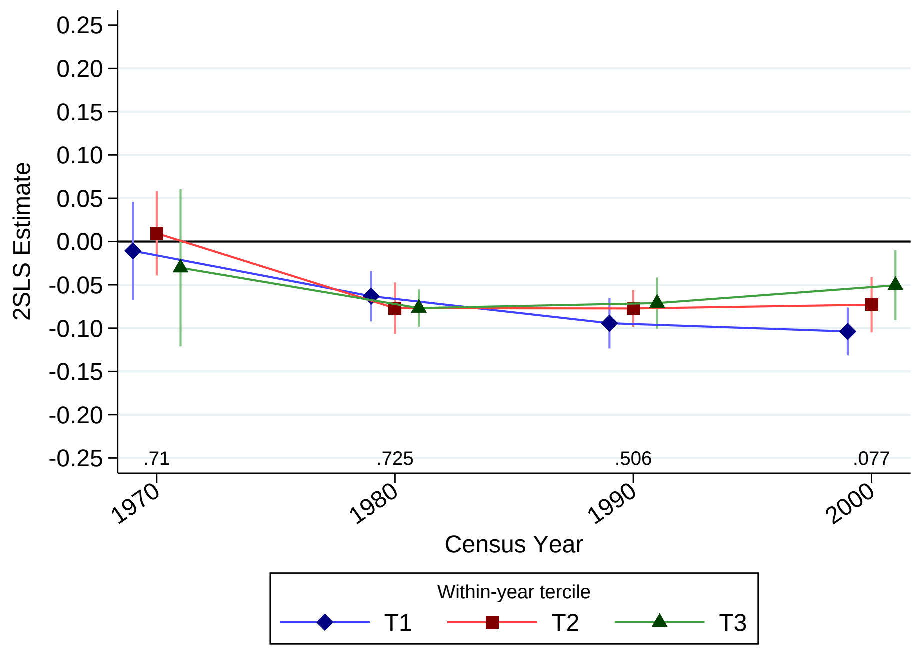
Figure A17: Twin IV estimates, US by race and region

(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates, stratified by race and US region. It uses the baseline sample of mothers for the US only. Regressions control for mother's age, age at first birth, gender of first child, and year fixed effects. Year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the year level. 95 percent confidence intervals based on robust standard errors clustered at the year level are displayed but may not always be visible at the scale of the figure.

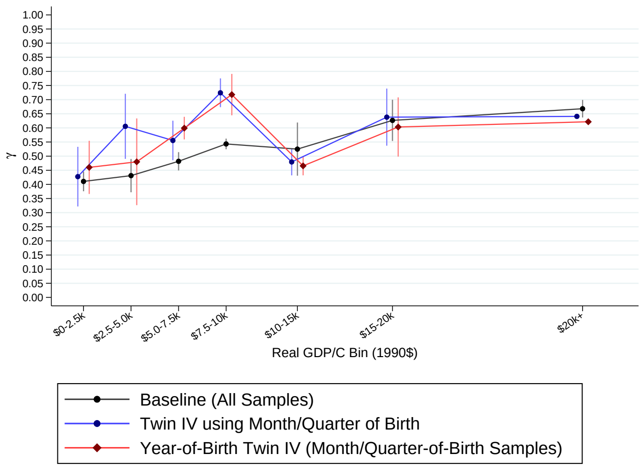
Figure A18: Twin IV estimates by attitudes about female labor supply



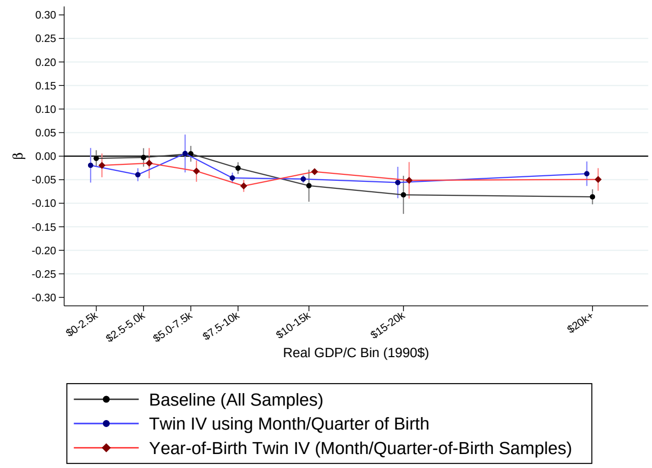
Notes: This figure displays twin IV estimates, binned by year and stratified by state tercile of attitudes about the acceptability of female labor force participation in the GSS. It uses the baseline sample of mothers for the US only. The numbers above the x-axis give the p-value of an F-test that the three point estimates are jointly equal in that year. Regressions control for mother's age, age at first birth, gender of first child, and year fixed effects. Year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the year level are displayed but may not always be visible at the scale of the figure.

Figure A19: Twin IV estimates by definition of twin

(a) First Stage: Third Birth



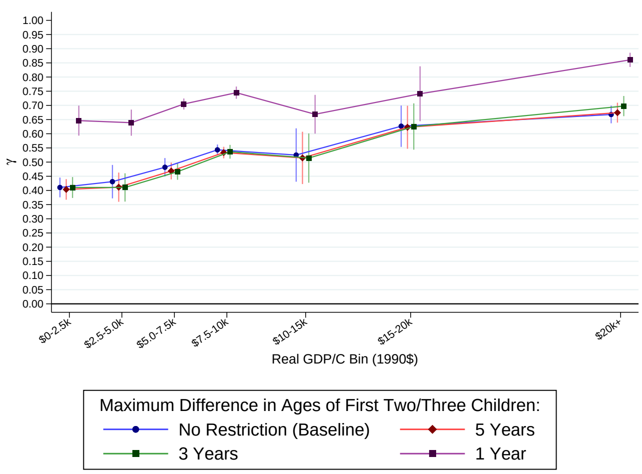
(b) Second Stage: Labor Force Participation



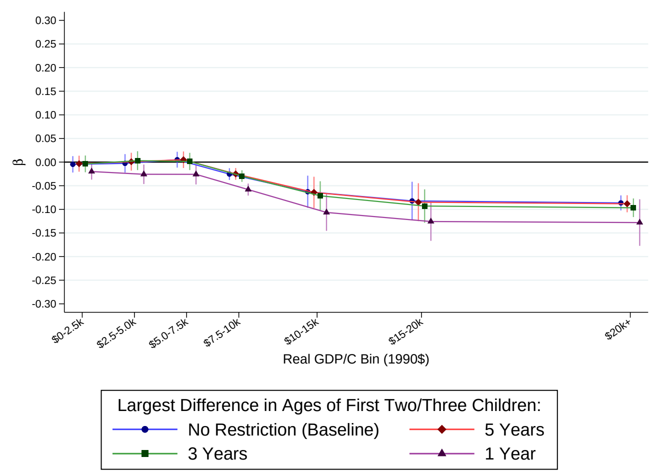
Notes: This figure displays twin IV estimates using alternative definitions of twins. The black baseline defines a twin as two children born to the same mother in the same calendar year. The blue line (circles) defines twins as being born in the same month or quarter, depending on the census or survey. See Appendix Table A1 for censuses and surveys where month/quarter or birth is available. The red line (triangles) uses the baseline definition of twins but the sample of censuses/surveys with month/quarter of birth. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A20: Twin IV by spacing of births

(a) First Stage: Third Birth



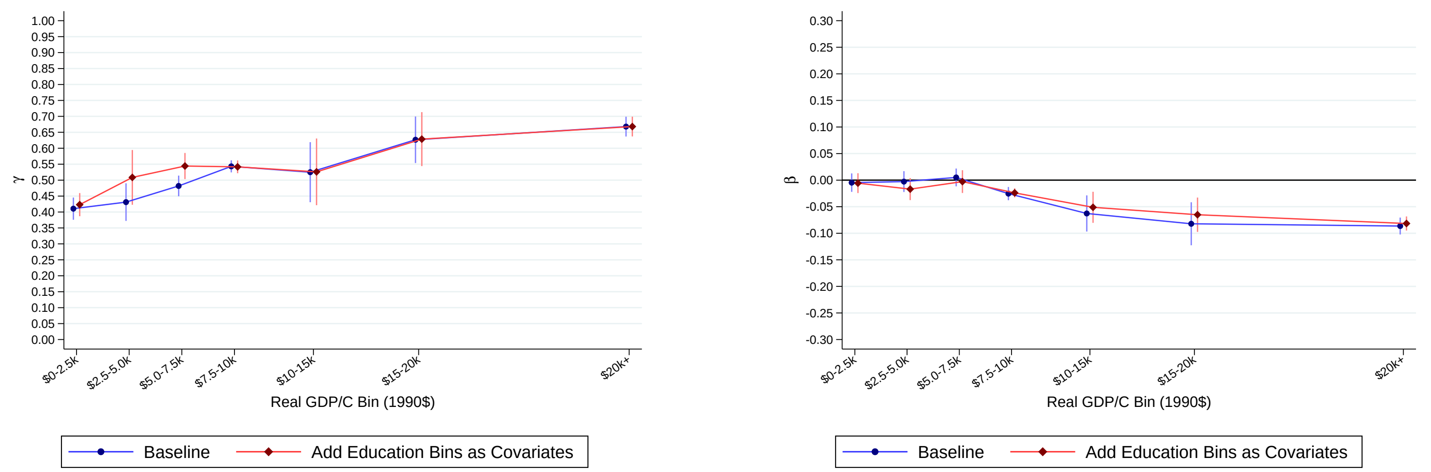
(b) Second Stage: Labor Force Participation



Notes: This figure compares baseline twin IV estimates to estimates that exclude mothers with gaps greater than 5, 3, and 1 year(s) between the births of their first two/three children. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A21: Robustness to education, twin IV

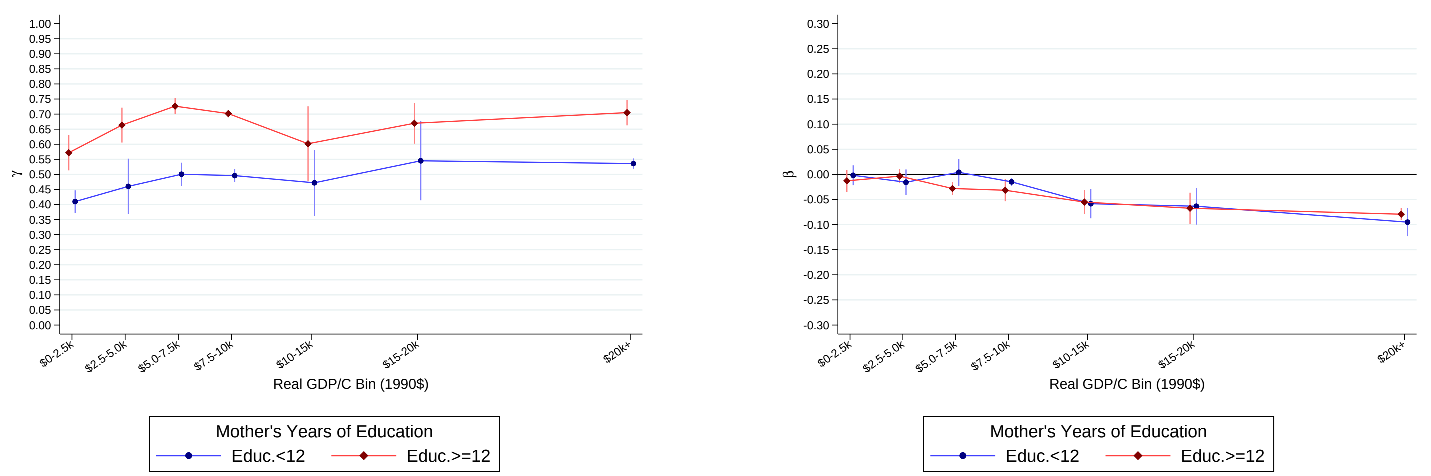
(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



Notes: This figure compares baseline twin IV estimates to estimates that include mother's education as a covariate. The sample is restricted to the set of mothers who report education. Regressions also control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A22: Twin IV by mother's education

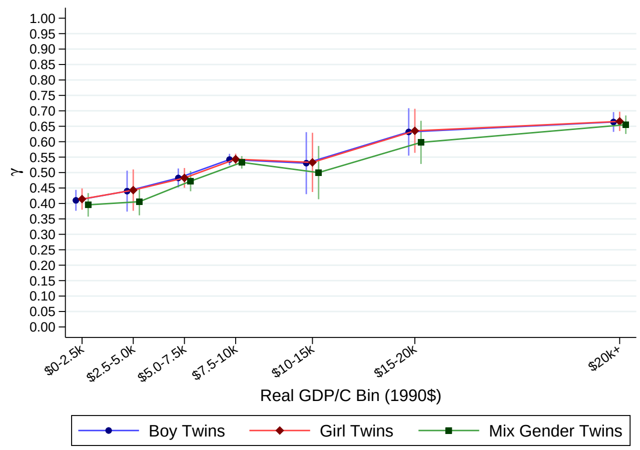
(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



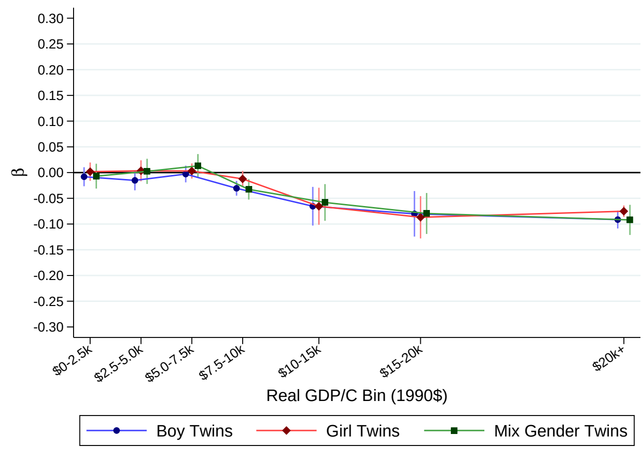
Notes: This figure displays twin IV estimates, stratified by mother's education. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure..

Figure A23: Twin IV estimates by gender of twins

(a) First Stage: Third Birth



(b) Second Stage: Labor Force Participation

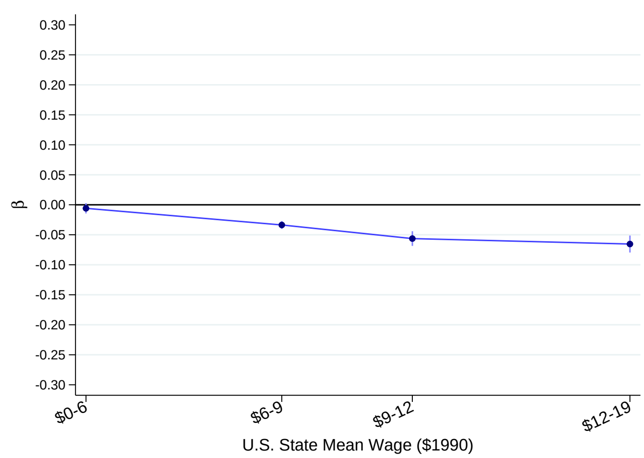


Notes: This figure displays twin IV estimates, stratified by age mix of the twins. Regressions control for mother's age, age at first birth, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A24: Twin IV estimates by U.S. state mean hourly wage, 1940-2010

(a) First Stage: Third Birth

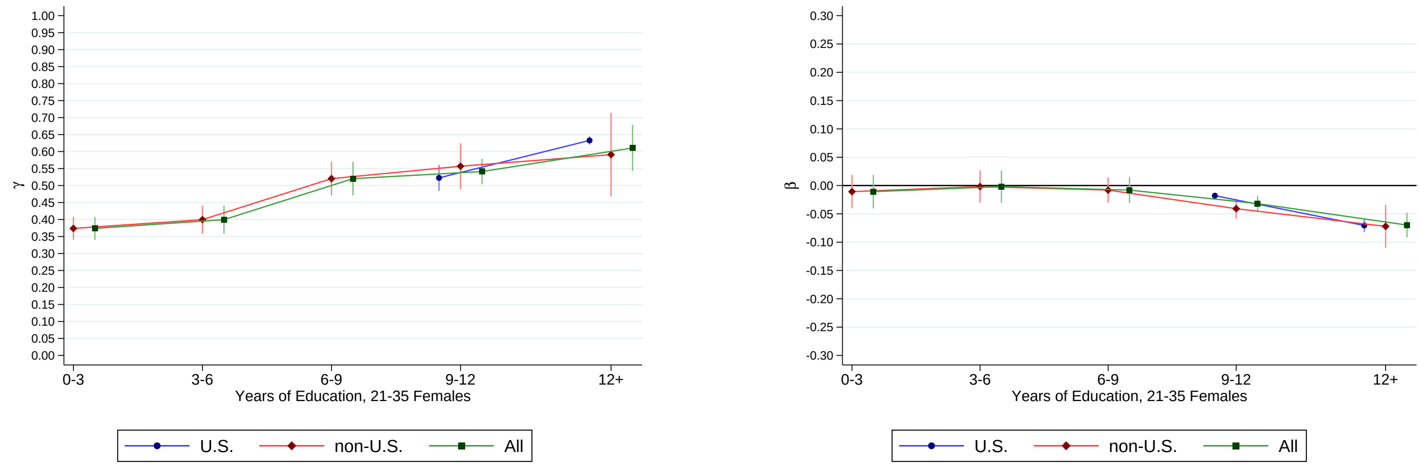
(b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates, binned by state average wage. It uses the sample of US mothers with husband wage information available. Regressions control for mother's age, age at first birth, gender of first child, and year fixed effects. Year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the year level. 95 percent confidence intervals based on robust standard errors clustered at the year level are displayed but may not always be visible at the scale of the figure.

Figure A25: Twin IV estimates by female education

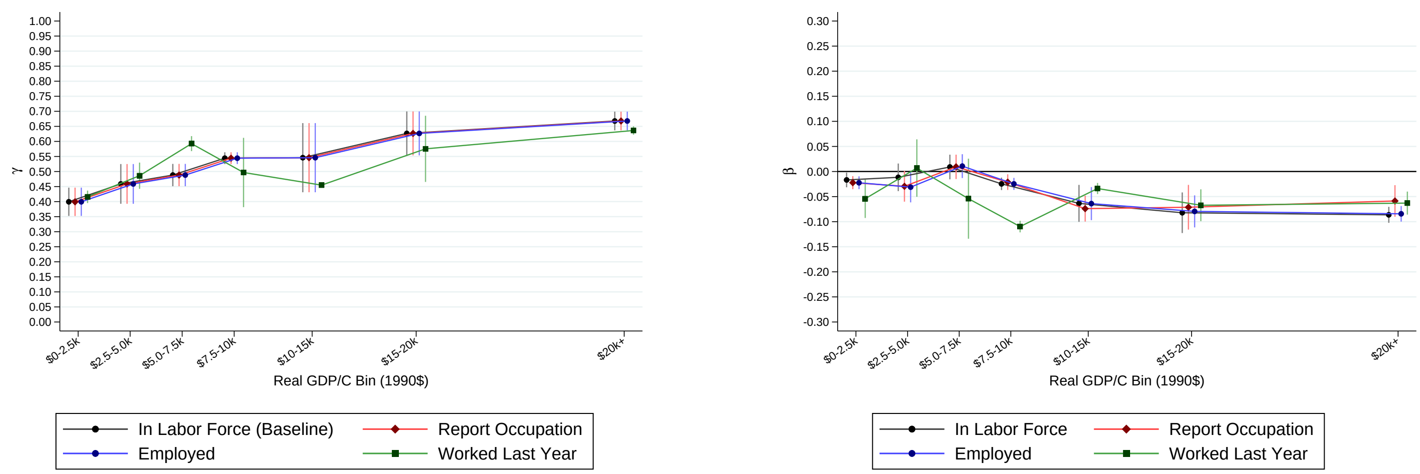
(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates, binned by average years of education of women aged 21-35 within each survey. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A26: Twin IV estimates using alternative labor supply measures

(a) First Stage: Third Birth (b) Second Stage: Labor Force Participation

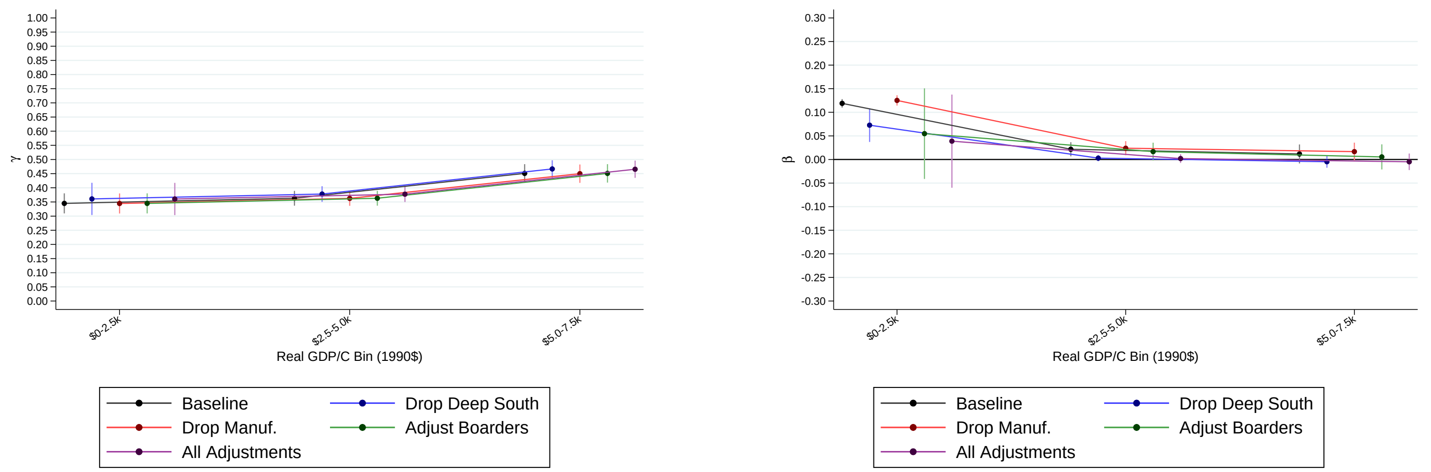


Notes: This figure displays twin IV estimates using alternative definitions of employment. The black, blue, and red lines use whether a mother is in the labor force (baseline), employed, and report any occupation, respectively. The sample is constant across all three indicators. The green (squares) line uses whether a mother worked in the previous year, and these results are based on a smaller sample of surveys/censuses. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A27: Twin IV estimates adjusted for mismeasured occupations, 1860-1930 U.S.

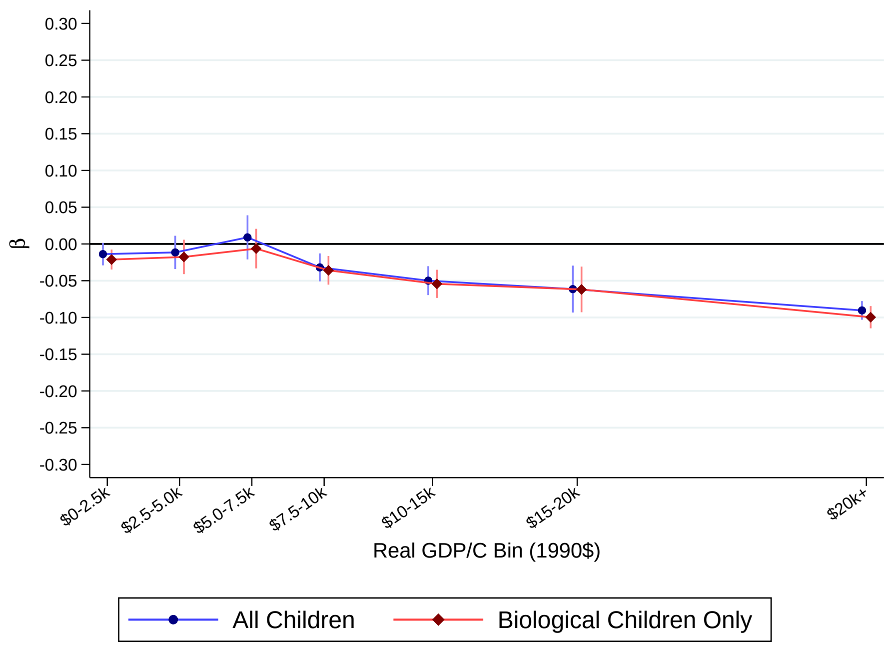
(a) First Stage: Third Birth

(b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates for the US that assess the sensitivity of our results to accounting for a variety of possible mismeasurement issues in pre-1940 U.S. occupational status, as identified in Goldin (1990). The black line is our baseline. The blue line drops the deep South (Alabama, Florida, Georgia, Louisiana, Mississippi, South Carolina, and Texas). The red line drops mothers who list their industry as manufacturing. The green line indicates a mother as working if there is at least one boarder in her household. The purple line makes all of these adjustments simultaneously. Only the first three real GDP/capita bins are impacted, so we do not show the others. Regressions control for mother's age, age at first birth, gender of first child, and year fixed effects. Year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the year level. 95 percent confidence intervals based on robust standard errors clustered at the year level are displayed but may not always be visible at the scale of the figure.

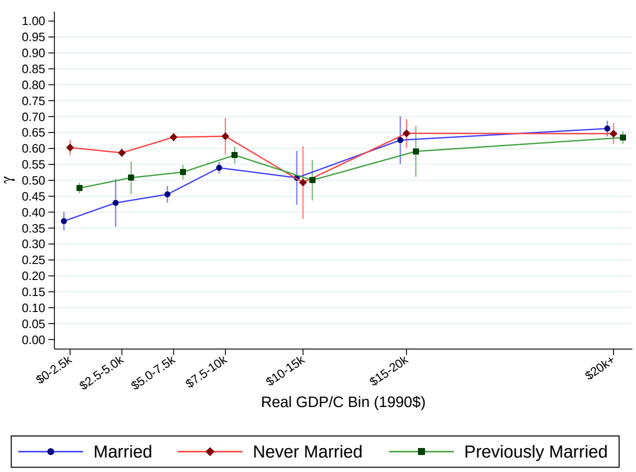
Figure A28: Twin IV estimates, robustness to non-biological children



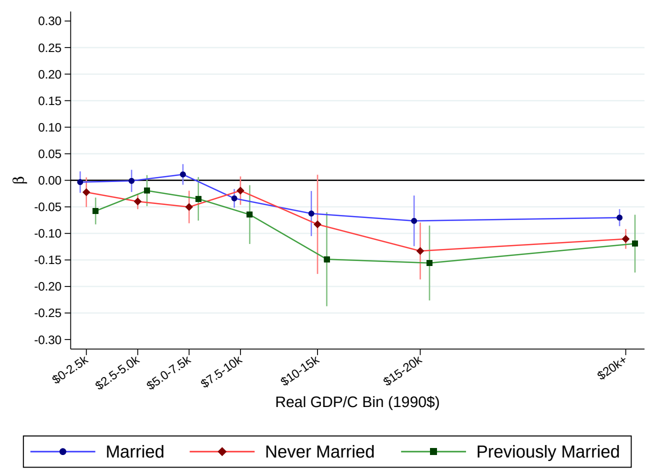
Notes: This figure compares baseline twin IV estimates to estimates that restrict to biological children only by requiring that the number of children in the household equal the number of children ever born to the mother. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.

Figure A29: Twin IV by marital status

(a) First Stage: Third Birth



(b) Second Stage: Labor Force Participation



Notes: This figure displays twin IV estimates, stratified by mother's marital status. Regressions control for mother's age, age at first birth, gender of first child, and country-year fixed effects. Country-year weights are normalized to the number of mothers in a survey. Robust standard errors are clustered at the country-year level. 95 percent confidence intervals based on robust standard errors clustered at the country-year level are displayed but may not always be visible at the scale of the figure.