Intergenerational Dynamics and the Fertility Transition

Tom S. Vogl*

Princeton University, BREAD, and NBER

January 2016

Abstract

Fertility change is distinct from other forms of social and economic change because it directly alters the size and composition of the next generation. This paper studies how changes in population composition over the fertility transition feed back into the pace of aggregate fertility decline. Theory predicts that changes in the relationship between human capital and fertility first weaken and then strengthen similarities in fertility choice between mothers and daughters, a process that first hastens and then slows aggregate fertility decline. Consistent with these predictions, microdata from 40 developing countries over the second half of the 20^{th} century show that intergenerational fertility associations strengthen late in the fertility transition, due to the alignment across generations of the relationship between human capital and fertility. As fertility approaches the replacement level, the strengthening of these associations reweights the population to raise aggregate fertility rates, pushing back against aggregate fertility decline.

---

*E-mail: tvogl@princeton.edu. I am grateful for comments from seminar and conference participants at Bocconi University, Hebrew University, Lund University, Northeastern University, Stockholm University, Princeton University, UC Berkeley, UC Louvain, University of Namur, the AEA/ASSA meetings, the NBER Children's Group Meeting, and the Warwick Summer Workshop in Economic Growth.

1 Introduction

Research in macroeconomics has shown increasing interest in whether and how heterogeneity across households—in income, wealth, or skill, for example—affects aggregate outcomes (Guvenen 2011; Perri et al. 2015). In the study of economic development and growth, classic papers on this topic (Banerjee and Newman 1993; Galor and Zeira 1993) focus on how capital market imperfections interact with the distribution of wealth to affect development in the long run. More recent work seeks to model the interplay between these forces and demographic change (Galor and Moav 2002; de la Croix and Doepke 2003). Because the escape from the Malthusian trap commonly coincides with the demographic transition (Galor 2012), an understanding of this economic-demographic interplay may be crucial for theories of current and historical economic growth. In this context, heterogeneity becomes especially interesting because it may feed back into the dynamics of population composition, with implications for the pace of aggregate change. Families with high relative fertility in the one generation comprise a larger share of the next.

This paper studies how skill heterogeneity combines with fertility and child investment decisions to influence the pace of aggregate fertility decline across generations. The core insight is that the fertility transition occurs unevenly, changing the relative fertility levels and population shares of different skill groups. As documented by Bengtsson and Dribe (2014), Clark and Cummins (2015), and Vogl (2016), the history of differential fertility spans two regimes, Malthusian and modern. In the past, when Malthusian population dynamics prevailed, higher-skill parents maintained higher levels of both fertility and child investment. But around the time of the demographic transition, their fertility fell below that of the lower-skilled, even as they continued to invest more in their children. Existing theoretical research (Galor and Moav 2002; Galor and Michalopoulous 2012) argues that these changes in differential fertility alter the skill composition of the population, with important consequences for economic growth and other forms of aggregate progress.[1] Building on that literature, this paper theoretically explores the mechanism's implications for the path of aggregate fertility and then quantifies these implications using micro-data from 40 developing countries. The main theoretical result is that forces commonly associated with the fertility transition—such as increases in the return to human capital or decreases in the cost of schooling—induce compositional

---

[1] See also Clark (2007). These theories have much in common with the eugenic arguments of Francis Galton, Karl Pearson, and Ronald Fisher (Kevles 1985).

shifts that initially promote fertility decline and then restrain it. This result squares well with evidence from the demography literature (Bongaarts 2006, 2008) that following periods of sustained fertility decline, many populations "stall" at fertility levels well above the replacement level.

A key determinant of these compositional shifts is the reduced-form association between a mother's fertility and her daughter's fertility, which provides a useful estimand for the empirical work. The theoretical framework predicts that within the Malthusian regime or the modern regime, this association is positive, but for a generation in the transition between regimes, the association turns negative. Higher-skill parents bear more children and invest more in their human capital, but when their children grow up, their greater human capital decreases their own fertility. Drawing on data from 82 Demographic and Health Surveys, the paper thus estimates mother-daughter associations in fertility for 40 developing countries in birth cohorts spanning the second half of the twentieth century. The results suggest that women of these cohorts predominantly gave birth in the modern regime, with fertility declining in parental skill. However, while more recent cohorts were also born in the modern regime, older cohorts are part of the transitional generation born in the Malthusian regime, with fertility rising in parental skill. For women in these older cohorts, a larger childhood family size is positively associated with education but unrelated to negatively related with relative fertility in adulthood.

Due to this strengthening of the intergenerational fertility association, the population share of high-fertility parents rises as a population approaches low fertility, pushing back against aggregate fertility decline. To quantify this phenomenon, the paper non-parametrically estimates the difference between the average number of children among mothers and the average number of grandchildren per daughter among grandmothers. This gap, which I term the *composition effect*, captures the extent to which heterogeneity in fertility across women who grew up in families of different size affects average fertility.[2] Estimating this effect at the cohort level, I find that the composition effect becomes significantly stronger as cohort fertility declines; well into the fertility transition, heterogeneity in fertility across mothers in one generation raises the mean fertility of their daughters by as much as 10 percent, and by 4 percent on average. To link these cohort-level associations to the evolution of total fertility rates (TFR), the paper also carries out a complementary analysis of period

---

[2] The composition effect bears similarities to the stable population theory concept of population momentum (Keyfitz 1971). Population momentum stems from the gradual evolution of the population age structure following a change in age-specific fertility rates. Analogously, one can view the composition effect of differential fertility as reflecting the gradual evolution of the population shares of different lineages following a change in their relative fertility rates.

fertility, aggregating over the cross-sectional age distribution in a given year. Here again, the composition effect on TFR becomes significantly stronger as TFR declines. In populations with TFRs less than 3, differential fertility raises TFR by as much as 6 percent, and by 3 percent on average.

These findings demonstrate a regularity in the role of heterogeneity in driving aggregate fertility dynamics. In this sense, they relate to the economics literature on how micro-level demographic phenomena like differential fertility and assortative mating aggregate up to the population or the economy (Kremer 1993; Lam 1986; Fernandez and Rogerson 2001; de la Croix and Doepke 2003). That literature focuses largely on aggregating skill or income across households, whereas this paper shifts attention toward aggregating fertility behavior, yielding results that may help resolve the puzzle of fertility "stalls."[3] The paper also expands on classic demographic research relating the average family sizes of women and children. Preston (1976) points out that because children from larger families are over-represented in the population, the average family size of children (i.e., sibship size) is generally larger than the average family size of women (i.e., fertility). In a hypothesis recently confirmed by Lam and Marteleto (2014), Preston argues that this difference shifts slowly during the demographic transition, so that average sibship size falls more slowly than average fertility. Just as in this paper, differential fertility reweights the population to slow the pace of aggregate demographic change as experienced by the next generation. Both lines of research highlight the value of considering how the cross-section interacts with the aggregate time series of important behaviors and outcomes.

Additionally, the results add to a large body of research, spanning the social sciences, on intergenerational associations in a range of outcomes, behaviors, and traits. Especially relevant is the recent economics literature that studies how intergenerational associations vary across space and over time. Much of this literature focuses on intergenerational earnings mobility, investigating its variation both across countries (Björklund and Jantti 1997; Solon 1999; Corak 2013) and across subnational areas within a country (Chetty et al. 2014). International comparisons are also available for intergenerational associations in of other outcomes, including educational attainment (Hertz et al. 2007; Chevalier et al. 2009) and health (Bhalotra and Rawlings 2013). In the field of demography, interest in the association of mothers' and daughters' fertility dates back over a century, to one of

---

[3] Indeed, my data contain three of the six countries that Bongaarts (2003) singles out for having pronounced "stalls"—the Dominican Republic, Kenya, and Peru—and all three exhibit recent upward swings in composition effects on cohort fertility. Composition effects on TFR are also significantly positive in Dominican Republic and Peru, but not in Kenya.

the earliest applications of the correlation coefficient (Pearson et al. 1899).[4] Summarizing and extending this literature, Murphy (1999, 2012) finds that the association strengthened recently in rich countries and that it is weaker in poor than in rich countries. This paper further extends Murphy's analysis to shed light on changes over time within many countries, linking it to a broader theory of aggregate demographic change.[5]

More broadly, the paper contributes to the economics literature on the demographic transition and its relation to economic development (Galor 2011; de la Croix 2013). In both theory and empirics, this work has focused mostly on aggregate demographic and economic variables, with a small subset of the theoretical literature exploring the role of within-population heterogeneity. While research in development economics has given more attention to within-population heterogeneity (Schultz 1997), its primary goal has been to understand the determinants of fertility choice in poor countries, not to shed light on how heterogeneity affects the pace of aggregate change. By analyzing heterogeneity and its consequences for many countries over half a century, this paper seeks to improve the empirical basis for theories linking population dynamics with aggregate change.

2 Heterogeneity and Aggregation in a Quality-Quantity Model

To shed light on how intergenerational shifts in population composition feed back into the process of fertility decline, this section studies a model in the tradition of the theoretical literature on demography and long-run economic growth. First, it derives how associations in the fertility behavior of parents and children aggregate up to the population level, irrespective of the specific behavioral model underlying these associations. Next, it analyzes a model in which parents of heterogeneous skill choose the quality and quantity of their children, using the earlier aggregation results to explore the evolving role of heterogeneity over the fertility transition. Finally, it broadens the discussion to consider whether the empirical implications change if factors outside the model also contribute to parent-child fertility associations.

---

[4] Seminal contributions include Huestis and Maxwell (1932), Duncan et al. (1965), Ben-Porath (1975), Wise and Condie (1975), Anderton et al. (1987), Kahn and Anderson (1992), and Axinn et al. (1994).

[5] Two recent contributions to biodemography (Murphy and Wang 2003; Kolk et al. 2014) study how intergenerational fertility associations affect the time series of aggregate fertility rates, but only in the context of highly stylized microsimulation models.

2.1 Preliminaries: Intergenerational Associations and Average Fertility

Let n_{it} denote the fertility of parent i from generation t, with mean μ_t and variance σ_t^2. Let $n_{i,t+1}$ denote the fertility (in adulthood) of each of the parent's children, such that if the parent has more than one child, then the $i, t+1$ subscript represents more than one member of generation $t+1$. Because the results below considers only expectations and conditional expectations, one can assume that siblings have the same fertility without loss of generality, so this slight abuse of notation is inconsequential.[6] Given the notation, parent i from generation t has $n_{it} n_{i,t+1}$ grandchildren.

The paper is interested in how the reweighting of the population due to differential fertility in generation t affects average fertility in generation $t+1$. To quantify this phenomenon:

Definition 1. The *composition effect* of differential fertility in generation t on average fertility in generation $t+1$, denoted Δ_{t+1}, equals the average fertility of generation $t+1$ minus the average number of grandchildren per child in generation t.

Using the notation above, one can express the composition effect as:

$$\begin{aligned}
\Delta_{t+1} &= E_{t+1}\left[n_{i,t+1}\right] - E_t\left[E\left(n_{i,t+1}|n_{it}\right)\right] \\
&= E_t\left[\left(\frac{n_{it}}{\mu_t}\right) E\left(n_{i,t+1}|n_{it}\right)\right] - E_t\Big[E\left(n_{i,t+1}|n_{it}\right)\Big] \\
&= E_t\left[\left(\frac{n_{it}}{\mu_t} - 1\right) E\left(n_{i,t+1}|n_{it}\right)\right]
\end{aligned} \tag{1}$$

where $E_{t+1}[\cdot]$ is an expectation evaluated under the distribution of $n_{i,t+1}$; $E_t[\cdot]$ is an expectation evaluated under the distribution of n_t; and $E\left(n_{i,t+1}|n_{it}\right)$ is the conditional expectation of a child's fertility given her parent's fertility, also known as her childhood family size or sibship size. In the first term of the second line, the distribution of n_{it} is reweighted by the factor $\frac{n_{it}}{\mu_t}$ to give more weight to members of generation $t+1$ from larger sibships, reflecting the changing composition of the population from t to $t+1$. This factor gives more weight to members of generation $t+1$ from larger sibships. If the conditional expectation $E\left(n_{i,t+1}|n_{it}\right)$ is an increasing function, then the added weight given to larger families raises average fertility, so that $\Delta_{t+1} > 0$. In this case, fertility heterogeneity in generation t raises average fertility in generation $t+1$.

Both the model and the empirical work use a linear projection to summarize the conditional expectation function $E\left(n_{i,t+1}|n_{it}\right)$.

---

[6] In fact, the model below exhibits no intra-family heterogeneity: siblings are endowed with the same human capital and optimally choose the same fertility. But the results in this subsection are general and do not only apply to the model.

Definition 2. The *intergenerational fertility association*, denoted β_{t+1}, is the coefficient from the linear projection of $n_{i,t+1}$ on n_{it}.

This association has received much attention in the literature on the intergenerational transmission of behaviors, outcomes, and traits. Here, it is particularly convenient because it can simplify the formula for the composition effect. Because expression (3) represents the composition effect for a conditional expectation of any form, computation may prove complicated. With a linear conditional expectation, however, the expression reduces.

Lemma 1. *If $E\left(n_{i,t+1}|n_{it}\right)$ is linear, then $\Delta_{t+1} = \beta_{t+1}\frac{\sigma_t^2}{\mu_t}$.*

The Theory Appendix contains all proofs. This result will prove useful for both the theory and the empirical work. Δ_{t+1} increases in β_{t+1} and σ_t^2, while decreasing in μ_t. In other words, the composition effect of differential fertility is more positive when the intergenerational association is stronger or when the variance of childhood family size is larger relative to its mean. The role of the ratio σ_t^2/μ_t in linking the cross-section with aggregate outcomes is not unique to the composition effect. Exactly the same ratio appears in Preston's (1976) formula for the difference between the average family size of children and the average family size of women. Both there and here, it implies that the aggregate implications differential fertility are larger when fertility is more dispersed relative to its mean.

Lemma 1 also provides a helpful decomposition of the composition effect in the linear case, which the empirical work will implement. One component measures similarity in the reproductive behavior of parents and children, while the other measures the relative spread of the fertility distribution among parents. Because the second component is always positive, the composition effect is negative—promoting fertility decline—only when the intergenerational association is negative.

2.2 Setup

The model setup closely follows the theoretical literature on differential fertility and growth.[7] Parents maximize a log-linear utility function over their own consumption c_{it}, the number of children n_{it}, and human capital per child $h_{i,t+1}$:

$$U\left(c_{it}, n_{it}, h_{i,t+1}\right) = \alpha \log\left(c_{it}\right) + (1-\alpha)\left(\log\left(n_{it}h_{i,t+1}\right)\right) \tag{2}$$

---

[7] See Galor and Moav (2002), Hazan and Berdugo (2002), de la Croix and Doepke (2003, 2005, 2008), and Moav (2005).

$\alpha \in (0,1)$ indexes the weight the parents place on their own consumption relative to the combined quantity and quality of children.

To produce child quality, or human capital, parents have access to a human capital production function that is isoelastic with respect to total inputs x_{it}: $h_{i,t+1} = x_{it}^{\eta}$, with $\eta \in (0,1)$. Each children is born with an input endowment $\theta > 0$ (reflecting basic human capacity or free public school), which parents can augment by purchasing additional education e_{it} at price p, so that $x_{it} = \theta + e_{it}$. One can view p as the inverse of the return to educational expenditure, so that an increase in this return is reflected in falling p. The presence of the input endowment implies that low-skill parents will choose a corner solution in which fertility rises with their human capital (de la Croix 2013; Vogl 2016), and also that the elasticity of human capital with respect to education is positive, which allows fertility to decline with parental human capital in the interior solution (Jones et al. 2010). Parents themselves are endowed with human capital $h_{it} > 0$, drawn from an distribution F_t.

The aggregate production function is $Y_t = AH_t$, where $A > 0$ is the level of technology and H_t is the aggregate quantity of human capital in generation t. Thus, the wage rate per unit of human capital is A. Notably, the aggregate and human capital productions function have one input each and are fixed over time. A fuller version of the model could incorporate human capital spillovers, scale effects, and additional inputs, but this section seeks to shed light on the compositional mechanism and its empirical implications, not to provide a full theory of the growth process.

Apart from education spending, each child costs $\tau \in (0,1)$ units of time and $\kappa > 0$ goods. These assumptions result in the following budget constraint:

$$c_{it} + \kappa n_{it} + p e_{it} n_{it} \leq A h_{it} \left(1 - \tau n_{it}\right) \tag{3}$$

Parents spend their full income Ah_{it} on their own consumption (c_{it}), the quantity costs of children (κn_{it} and $\tau A h_{it} n_t$), and the quality costs of children ($p e_{it} n_{it}$).

2.3 Results

The first order conditions imply that parents set education at:

$$e_{it} = \max\left\{o, \frac{1}{1-\eta}\left(\frac{\eta}{p}\left(Ah_{it}\tau + \kappa\right) - \theta\right)\right\} \tag{4}$$

As a result, education starts at a corner solution of zero for low-skill parents and then rises linearly with parental skill in the interior solution. Parents spend on education if and only if their human capital exceeds $\widetilde{h} \equiv \frac{1}{\tau A}\left(\frac{p\theta}{\eta} - \kappa\right)$. This threshold also plays an important role in fertility determination, as revealed in the solution:

$$n_{it} = \begin{cases} \frac{(1-\alpha)Ah_{it}}{\kappa+\tau Ah_{it}} & \text{if } h_{it} \leq \widetilde{h} \\ \frac{(1-\alpha)(1-\eta)Ah_{it}}{\kappa-p\theta+\tau Ah_{it}} & \text{if } h_{it} > \widetilde{h} \end{cases} \tag{5}$$

From inspection of equation (5), one can see that an additional assumption is necessary for unambiguous comparative statics with respect to parental skill.

Assumption 1. *The value of the input endowment is strictly larger than the goods cost:* $p\theta > \kappa$*.*

To obtain its value, the input endowment θ is scaled by the price of education p. This assumption guarantees that $\widetilde{h} > 0$, so parents with very low skill choose a corner solution, and also that the substitution effect of higher parental skill dominates its income effect in the interior solution.

Lemma 2. *Fertility* n_{it} *increases with parental human capital* h_{it} *over the interval* $(0, \widetilde{h}]$ *and decreases with* h_{it} *over the interval* $(\widetilde{h}, \infty)$*.*

Under Assumption 1, the model thus predicts a hump-shaped relationship between parental skill and fertility. Below $\widetilde{h}$, parents choose a corner solution for human capital expenditure, so the income effect of higher skill dominates the substitution effect, and fertility rises with skill. Above $\widetilde{h}$, the substitution effect dominates, so fertility declines with skill.[8] In data from 48 developing countries, Vogl (2016) documents exactly this hump-shaped pattern, such that the least-skilled exhibit lower fertility in the early stages of development and higher fertility in the later stages.

In addition to the hump-shaped fertility profile, the corner solution also generates an s-shaped relationship between the human capital of parents and children. Substitution of equation (4) into the human capital production function leads to a first-order, non-linear dynamical system $h_{i,t+1} = \phi(h_{it})$, where:

$$\phi\left(h_{it}\right) \equiv \begin{cases} \theta^{\eta} & \text{if } h_{it} \leq \widetilde{h} \\ \left\{\left(\frac{\eta}{1-\eta}\right)\left(\frac{\kappa+\tau Ah_{it}}{p} - \theta\right)\right\}^{\eta} & \text{if } h_{it} > \widetilde{h} \end{cases} \tag{6}$$

---

[8] If $p\theta \leq \kappa$, the declining portion of the hump disappears, so that fertility everywhere increases with parental skill.

Equation (6) implies several useful properties.

Lemma 3. *The dynamical system $\phi(\cdot)$ is continuous, constant over the interval $(o, \widetilde{h}]$, and increasing and concave over the interval $(\widetilde{h}, \infty)$, with $\lim_{h_{it} \to \infty} \phi'(h_{it}) = 0$.*

These properties are key to identifying the steady states of the dynamical system—with constant levels of human capital (and therefore fertility) across generations—which help shed light on the model's implications for intergenerational associations and composition effects over the fertility transition. In particular, $\phi(\cdot)$ may have multiple steady states, generating heterogeneity that is sufficiently tractable for the study of β_t and Δ_t.

Proposition 1. *The dynamical system $\phi(\cdot)$ is characterized by at least one and at most three steady state levels of human capital and fertility, of which at least one and at most two are stable.*

The dynamical system is guaranteed to have a steady state and may have up to three, two of which have positive human capital expenditure and one of which has none. Of these three steady states, those with the lowest and highest levels of human capital are stable, while the intermediate one (which exists only if one of the others also exist) is unstable. For an environment with three steady states, Figure 1 graphs the dynamical system together with the relationship between human capital and fertility. Steady states ss_1, ss_2, and ss_3 occur when the dynamical system intersects the 45-degree line. ss_1 and ss_3 are stable because the dynamical system crosses the 45-degree line from above, whereas ss_2 is unstable because it crosses from below. In this case, the fertility level associated with ss_1 is higher than that associated with ss_3, but other parameter values could give rise to the opposite result. The remainder of the section focuses on the two stable steady states of an environment with three steady states.

Definition 3. A *fertility regime* is a distribution of fertility in which all lineages are in stable steady states. A fertility regime is *non-degenerate* if fertility is heterogeneous within it.

The history of differential fertility over the course of economic development suggests two types of fertility regimes: one in which the higher-skill steady state has higher fertility, and one in which it has lower. I refer to the first type of regime as "Malthusian" and the second as "modern." In Figure 1, a population distributed across ss_1 and ss_3 would be in a modern fertility regime.

In this model, two forces are particularly natural for generating transitions between Malthusian and modern regimes: rising productivity and a rising return to human capital expenditure, as represented by a decrease in the price of education, p. Rising productivity pushes up the wage return to human capital, which increases fertility in the low-skill steady state, while decreasing fertility in the high-skill steady state. Falling p decreases the fertility of the high skilled, while leaving the low-skill steady state unchanged.

Proposition 2. *Consider a population initially in a non-degenerate fertility regime that experiences a one-time, permanent increase in A or decrease in p. If the population attains a new non-degenerate fertility regime, fertility in both the transition and the new regime is weakly higher among low human capital parents and strictly lower among high human capital parents.*

Consequently, if the population starts in a Malthusian regime, either force pushes it toward a modern regime. Each mechanism has a different implication for average fertility, however. A decline in p decreases average fertility in the population, while an increase in A has an ambiguous effect. In aggregate data on the demographic transition, average fertility typically first rises and then falls (Dyson and Murphy 1985; Galor 2011), so both mechanisms are consistent with features of the historical fertility transition. Most directly, these mechanisms relate to theories that place rising income and human capital at the center of the transition (Galor 2011). But one can also link them to a theory popular among demographers, which involves changes in social norms regarding the role of children (Caldwell 1981). In particular, one can interpret new social norms favoring the education of children as as an increase in the social return to education expenditure, which is equivalent to a decline in p.

So far, the model has omitted child mortality, which classic demographic transition theory views as a key driver of fertility change. Because child mortality occurs before substantial human capital investment, the model can capture it by dividing the quantity costs of children by the survival rate. If m is the child mortality rate, then the quantity costs become $\frac{\kappa}{1-m}$ and $\frac{\tau}{1-m}$. In this setup, a decrease in child mortality raises (surviving) fertility in both steady states, with has ambiguous effects on fertility differentials. The prediction that net fertility rises with survival is common in fertility models in which a share m of children die deterministically, although it is not unique to them. In particular, Doepke (2004) finds that it also holds in appropriately parameterized models with stochastic child mortality and sequential fertility choice. Thus, in the standard economic models of the demographic transition, child mortality does play a major role in net fertility change. As such, the remainder of this section focuses on the changes in fertility regimes that result from rising productivity or returns to educational expenditure.

We are now equipped to consider how the intergenerational fertility association and the composition effect evolve over the transition from a Malthusian regime to a modern regime.

Proposition 3. *In any non-degenerate fertility regime and any transition between non-degenerate regimes,* β_t *and* Δ_t *are strictly positive, except during a transition from a Malthusian regime to a modern regime, during which they are weakly negative for a single generation.*

Because parents' human capital is positively associated with children's human capital, fertility is also positively associated in all pairs of generations except one: the generations between which skill differentials in fertility flip. In fact, because human capital and fertility are constant in a steady state, $\beta_t = 1$ within a fertility regime. This implication is quantitatively unrealistic but would be weakened by the introduction of stochastic shocks to human capital or income. As such, the model's predictions should be viewed primarily as qualitative.

From the perspective of population dynamics, the most important implication of Proposition 3 is that the composition effect Δ_t is negative for the first generation to exhibit a reversal of differential fertility. If this reversal occurs around the same time as the onset of aggregate fertility decline, then this negative composition effect initially speeds the decline. In subsequent generations, however, the composition effect turns positive again, pushing back against aggregate fertility decline. These results reflect how the reversal of differential fertility alters the population shares of higher- and lower-fertility parents.

2.4 Empirical Implications

The most general empirical prediction is that when the fertility differential between high- and low-skill parents flips, the intergenerational fertility association and the composition effect turn negative for a generation. However, the model highlights only one source of the intergenerational association, human capital persistence. Although the literature on demography and long-run growth (Galor 2011) suggests that this source is likely to be important, other sources are possible.[9] Chief among

---

[9] Galor and Moav (2002) propose a model with genetic or cultural transmission in the taste for child quantity or quality, which has similar properties to the model here.

them are the genetic heritability of fecundity and the cultural determinants of fertility preferences or behavior.[10] The behavioral genetics literature documents some genetic heritability in markers of fecundity—including menstrual regularity and the ages at menarche and menopause (van Akker et al. 1987; Treloar and Martin 1990)—which may lead to the heritability of fertility. Meanwhile, a large literature in the social sciences emphasizes the role of culture in generating intergenerational transmission of various traits and behaviors. Sociologists (e.g., Duncan et al. 1965) have long emphasized this explanation for intergenerational fertility associations, positing that growing up in a large family shapes the preference for a large family of one's own.

To combine the model's prediction with these alternative sources of the intergenerational fertility association, one can express the association as:

$$\frac{\partial E\left[n_{t+1}|n_t\right]}{\partial n_t} = \frac{\partial E\left[n_{t+1}|n_t, h_{t+1}\right]}{\partial n_t} + \frac{\partial E\left[n_{t+1}|n_t, h_{t+1}\right]}{\partial h_t} \times \frac{\partial E\left[h_{t+1}|n_t\right]}{\partial n_t} \tag{7}$$

If the conditional expectation is linear, then $\beta_t = \frac{\partial E[n_{t+1}|n_t]}{\partial n_t}$. This equation merely expands the overall fertility association, $\frac{\partial E[n_{t+1}|n_t]}{\partial n_t}$, into sub-components using the chain rule. These subcomponents include any intergenerational transmission of fertility net of the human capital mechanism, $\frac{\partial E[n_{t+1}|n_t,h_{t+1}]}{\partial n_t}$; the relationship between skill and fertility, $\frac{\partial E[n_{t+1}|n_t,h_{t+1}]}{\partial h_t}$; and the relationship between childhood family size and skill, $\frac{\partial E[h_{t+1}|n_t]}{\partial n_t}$. The intergenerational transmission sub-component captures persistence in the determinants of fertility that are not correlated with education, perhaps genetics or culture. The model focuses on the second and third sub-components, which change in a staggered fashion during the demographic transition. In a Malthusian regime, higher-skill parents bear more children, so $\frac{\partial E[n_{t+1}|n_t,h_{t+1}]}{\partial n_t} > 0$, and women from larger sibships obtain more education, so $\frac{\partial E[h_{t+1}|n_t]}{\partial n_t} > 0$. In a modern regime, both patterns are flipped, so $\frac{\partial E[n_{t+1}|n_t,h_{t+1}]}{\partial n_t} < 0$ and $\frac{\partial E[h_{t+1}|n_t]}{\partial n_t} < 0$. In the transition between these two regimes, a single generation experiences long and auxiliary parameters of opposite sign: $\frac{\partial E[n_{t+1}|n_t,h_{t+1}]}{\partial n_t} < 0$ and $\frac{\partial E[h_{t+1}|n_t]}{\partial n_t} > 0$. If the 'pure' intergenerational transmission component is small or varies little over time, then the intergenerational association $\frac{\partial E[n_{t+1}|n_t]}{\partial n_t}$ follows the trajectory predicted by the model: large within the Malthusian and modern regimes, small during the transition between them. However, in the presence of 'pure' intergenerational transmission, $\frac{\partial E[n_{t+1}|n_t]}{\partial n_t}$ may not turn negative in the transition.

---

[10] Intergenerational associations in other socioeconomic outcomes, such as income or health, may also play a role. Because these outcomes are closely linked to human capital, they are broadly related to the model.

Given Lemma 1, equation (7) has similar implications for the composition effect. In particular, the composition effect will be large and positive within the Malthusian and modern regimes, and will become small (and possibly negative) during the transition between them. If it indeed turns negative during the transition, then it will initially speed the process of fertility decline, as predicted by the model. But if 'pure' transmission keeps it from turning negative, then the composition effect will never make fertility decline faster than it would in the absence of population reweighting. Nevertheless, in either case, the growth of the composition effect after the transition will subsequently slow the process of fertility decline.

I will argue that in the postwar era, developing countries moved from from the transition generation to the modern regime. As a result, the results will show that the intergenerational association and the composition effect grew as the relationship between education and sibship size flipped from positive to negative. In other words, all cohorts in the Demographic and Health Surveys gave birth in an era when the relationship between human capital and fertility, $\frac{\partial E[n_{t+1}|n_t,h_{t+1}]}{\partial h_t}$, was negative. Earlier cohorts were born in an era when the association of sibship size with human capital, $\frac{\partial E[h_{t+1}|n_t]}{\partial n_t}$, was positive, while later cohorts were born in an era when this association was negative. Because this reversal of differential fertility coincided with fertility decline (Vogl 2016), one can expect the intergenerational association and the composition effect to grow as aggregate fertility falls.

3 Data on the Fertility of Mothers and Daughters

To link mothers' and daughters' fertilities, I draw on data from the Demographic and Health Surveys, a collection of nationally-representative samples of women of childbearing age (generally 15-49). Two survey modules are key for the analysis. The first, the fertility history module, asks women to list all of their own children ever born, with several details like date of birth and survival status. The second, the sibling history module, asks women to list all of their siblings ever born to their mothers, with similar information. Combined, these two lists allow one to retrospectively track fertility behavior over two generations in a lineage.

Although the design of the DHS allows for comparisons across surveys and countries, questionnaires and sampling methods occasionally differ. Surveys must meet two criteria for inclusion in the study sample. First, they must sample all women, not only ever-married women, to avoid sample selection on an outcome closely related to fertility.[11] Second, they must include information on the survival of both offspring and siblings, to allow analysis of both ever-born and surviving fertility. Based on these criteria, the paper uses 82 surveys in 40 African, Asian, Caribbean, Latin American, and Pacific countries (see Appendix Table 1 for a list of countries and survey years).[12] Together, these surveys provide data on sibship size and fertility for 966,498 women aged 15-49.

As its main unit of analysis, the analysis focuses on birth cohorts, estimating intergenerational associations and composition effects among women born in the same country and the same 5-year interval. Because this strategy generates parameter estimates for many cells per country, it allows a rich analysis of changes in intergenerational associations and composition effects within a country over time. Notably, this approach abstracts from the model, which considers generations, or all women born to the same group of mothers. DHS data do not allow estimation of generational statistics. Instead, each cohort statistic will summarize fertility for women from a hypothetical generation of families whose size is distributed according to the cross-sectional distribution of sibship size in the birth cohort.

Analyses of both the intergenerational association and the composition effect involve two steps: first, estimating parameters for each country-cohort cell, and second, documenting patterns across (rather than within) cells. In the first step, I use sampling weights provided by the DHS, but I rescale the weights to maximize efficiency with multiple surveys per country. Specifically, if a country-cohort cell includes data from surveys $k = 1, \cdots, K$, each with N_k observations, then I rescale the survey weights from survey k to sum to $\frac{N_k}{\sum_k N_k}$. This approach weights individual surveys in proportion their contribution to the overall country-cohort sample. In the second step, analyses are unweighted to ensure representativeness for all cells. For precision, however, I drop cells with fewer than 250 observations, or the 5th percentile of the cell size distribution for the main analysis.

In an extension to the cohort analysis, I quantify the implications of the main results for the total fertility rate (TFR), a period measure of fertility that reflects the expected number of children for a woman who experiences current age-specific fertility rates throughout her lifetime. Here, the unit of analysis is a country-year, so each survey generates a single estimate. Because the period analysis does not pool surveys before the first step, no weight rescaling is necessary.

---

[11] The one exception to this rule is the 1996 Nepal DHS, which surveyed only ever-married women. Because 98 percent of Nepalese women over 30 in that year were ever-married, I include data from that survey on women over 30.

[12] Two additional surveys, the 1989 Bolivia DHS and the 1999 Nigeria DHS, meet the sample inclusion criteria but are omitted due to irregularities in the sibling history data.

4 Intergenerational Fertility Associations

This section estimates intergenerational fertility associations at the country-cohort level and studies how they evolve over time. For completeness, it reports results for both ever-born fertility and surviving fertility. If one were purely interested in intergenerational associations in fertility *behavior* (the number of pregnancies), then ever-born fertility would be more relevant. If one were instead interested in intergenerational associations in fertility *outcomes* (the effective number of offspring), then surviving fertility would be more relevant. But for those interested in intergenerational associations in the *demand for children*, both measures of fertility may be relevant, depending on parents' ability to target the number of surviving offspring subject to mortality risk. From the model's perspective (and from an evolutionary perspective), only surviving fertility is relevant. Aside from these substantive issues, surviving fertility is also attractive because it is less subject to recall bias. Respondents may forget their deceased children, but more importantly, they may forget their deceased siblings, some of whom may have died before the respondents were born.

4.1 Methods

I estimate intergenerational associations in a linear model relating a woman's fertility outcome at a specific age to her sibship size. For woman i from country c and 5-year birth cohort t:

$$y^a_{ict} = \alpha^a_{ct} + \beta^a_{ct} s_{ict} + u^a_{ict} \tag{8}$$

where y^a_{ict} is cumulative fertility at age a, and s_{ict} is sibship size. β^a_{ct} is country-cohort ct's intergenerational fertility association at age a: for short, the cohort fertility association. As discussed in Section 2, β^a_{ct} captures *all* mechanisms linking a mother's fertility with her daughter's. It is not a causal effect, and its sources are not necessarily limited to the mechanism specified in the model.

The choice of a poses an important methodological tradeoff. The DHS interviews most respondents midway through their childbearing years, so the analysis must consider cumulative fertility before the end of reproductive age. An earlier a allows for a larger sample and coverage of later birth cohorts, while a later a covers more of the reproductive lifespan. To deal with this trade-off concerning the age of measurement, I follow the literature on intergenerational income associations, where similar issues arise (Solon and Haider 2006).[13] In their research on intergenerational income mobility in the United States, Chetty et al. (2014) use the age at which the intergenerational association becomes stable for the rest of the lifecycle, which they estimate to be 30. If fertility gaps are concentrated early in the reproductive lifespan, then one can take a similar approach to analyzing intergenerational fertility associations. Along these lines, I choose the earliest age a for which β_{ct}^{a} approximates the completed fertility transmission coefficient. To find this age, I draw on the full fertility histories of women at least 45 years old to estimate β_{ct}^{a} at ages 20, 25, 30, 35, 40, and 45. I then regress the cohort's age-45 association on its association at an earlier age.[14] Figure 1 plots the results by age, indicating that intergenerational associations at earlier ages are strongly informative about the association for completed fertility (as measured at age 45). None of the estimated slopes are significantly different from 1, and all are significantly different from 0. However, the estimated slope for surviving fertility is substantially below 1 for the association at age 20. By age 25, the slope is extremely close to 1 for ever-born and surviving fertility. Given these results, the analysis focuses on cumulative fertility at age 25.

Motivated by equation (7), one can decompose the intergenerational fertility association into its driving forces: (1) the association of sibship size and fertility, conditional on human capital; (2) the association of human capital and fertility, conditional on sibship size; and (3) the association of sibship size and human capital. Two equations summarize these three partial derivatives. The first expresses fertility outcomes as a function of sibship size and education:

$$y_{ict}^{a} = \mathrm{A}_{ct}^{a} + \mathrm{B}_{ct}^{a} s_{ict} + \Gamma_{ct}^{a} edu_{ict} + U_{ict}^{a} \tag{9}$$

while the second expresses education as a function of sibship size:

$$edu_{ict} = \psi_{ct} + \varphi_{ct} s_{ict} + v_{ict} \tag{10}$$

In the language Goldberger (1991) uses to characterize omitted variables bias, equation (8) is a

---

[13] In estimating the parent-child association in income, researchers are often forced to use data from early in the child's career. Since age-earnings profiles are steeper for individuals with high lifetime earnings, early measurement of the child's income can bias estimates of the parent-child association downward. This lifecycle bias is analogous to the problem that arises here, except that the dependent variable is a flow (income) rather than a stock (cumulative fertility).

[14] A limitation of this approach is that it cannot account for any changes in the age profile of fertility that may have occurred in recent cohorts.

"short" regression, while equations (9) and (10) are the corresponding "long" and "auxiliary" regressions. B^a_{ct} is the intergenerational transmission term discussed in Section 2.4, accounting for factors outside the theoretical framework. Γ^a_{ct} is the education-fertility association, representing differential fertility in the current generation, while φ_{ct} is the sibsize-fertility association, representing differential fertility in the last generation. By equation (7), the cohort coefficients are related by the identity $\beta^a_{ct} = \mathrm{B}^a_{ct} + \varphi_{ct}\Gamma^a_{ct}$, which confirms that the fertility association is smaller when the sibsize-education and education-fertility associations are of opposite sign.

4.2 Results

Figure 3 plots kernel densities of the estimated β^a_{ct} and B^a_{ct} across country-cohort cells, revealing considerable dispersion. The dashed blue curves correspond to β^a_{ct} from the "short" regression, while the solid red curves correspond to B^a_{ct} from the "long" regression. Three patterns emerge for both ever-born and surviving fertility. First, the central tendencies of β^a_{ct} and B^a_{ct} are close to zero, and the supports of the distributions include both positive and negative values. Second, the dispersion of B^a_{ct}—from the "long" regression, which controls for education—is smaller than that of β^a_{ct} from the "short" regression. This second finding implies that heterogeneity in the role of education explains part of the observed heterogeneity in intergenerational fertility association. Third, and perhaps more interestingly, the distribution of B^a_{ct} has shorter right *and* left tails than the distribution of β^a_{ct}, suggesting that the education amplifies intergenerational fertility associations when they are both positive and negative.

One interpretation of Figure 3 is that intergenerational fertility associations are small everywhere, so that the distributions observed in Figure 2 are the result of sampling variability. However, the distributions of the associated *t*-statistics, summarized in Table 2, do not fit such an interpretation. At the 5 percent significance level, the ever-born fertility association is significantly positive in 28 percent of cells and significantly negative in 1 percent; the surviving fertility association is significantly positive in 15 percent of cells and significantly negative in 9 percent. As a result, the distribution of the surviving fertility associations contains too many positive values *and* too many negative values to be explained by randomness.

The fact that the β^a_{ct} distribution has more mass in both tails than the B^a_{ct} distribution means that education can amplify intergenerational fertility associations in both directions. This result could stem from two sources: (1) heterogeneity in the association of education with fertility (Γ_{ct}) and (2) heterogeneity in the association of sibship size with education (φ_{ct}). Although both theory and existing research suggest that both associations flip from positive to negative over the fertility transition, these reversals are one generation apart. The education-fertility association flips first, inducing a flip in the next generation's sibsize-education association. For insight into which of these sources is at play, Table 2 also reports the distribution of t-statistics for estimates of Γ_{ct} and φ_{ct}. The results suggest that the reversal of the education-fertility association had already occurred before the sample cohorts began childbearing, with 89 percent of Γ_{ct} estimates significantly negative for ever-born fertility and 79 percent for surviving fertility. In contrast, the distribution of the sibsize-education association has significant mass in both tails, suggesting that some cohorts were born when Malthusian fertility differences prevailed, so that children with higher skill parents had more siblings and obtained more education. For ever-born fertility, 38 percent of φ_{ct} estimates are significantly positive and 26 percent negative. For surviving fertility, 51 percent are positive and 15 percent negative.

The distribution of estimates from equations (8)-(10) suggests that most women in the sample bore children during the modern era, but some were born during the Malthusian era. One can shed more light on this issue by considering how the slope parameters change on average across successive cohorts. To characterize this evolution, I regress each estimated parameter on country fixed effects and cohort fixed effects, plotting the cohort effects in Figure 4. I set the omitted category to the the only cohort for which data are available for all countries, 1965-9, and then add the mean parameter value to all the cohort effects, so the figure provides information about average levels and changes. The horizontal line reflects the 1965-9 mean; the point estimates and 95 percent confidence intervals are for differences relative to this line.

Starting from the right-hand side of Figure 4, one can see clear evidence of the transition from Malthusian to modern regimes. For both ever-born and surviving sibship size, the average sibsize-education association is positive for the 1965-9 cohort, significantly more positive for earlier cohorts, and significantly more negative for later cohorts. Moving forward one generation, the average education-fertility association is negative for the 1965-9 cohort. But similar to the sibsize-education association, it is significantly more positive for earlier cohorts, and significantly more negative for later cohorts. Meanwhile, the transmission coefficient averages near zero with little change.

The intergenerational association equals the transmission coefficient plus the product of the two education-related associations. Since the transmission coefficient is constant while the the education-related associations become more negative across successive cohorts, intergenerational fertility associations should become more positive. Indeed, the left-most panels of Figure 4 indicate that intergenerational associations for both ever-born and surviving fertility have become significantly more positive since the 1965-9 cohort. Importantly, the magnitudes are not large: the intergenerational association averages 0.02-0.03 for the 1980-4 cohort, implying that a woman with 10 siblings can expect to have $\frac{1}{10}$ more children by age 25 than a woman with no siblings. Nevertheless, the data show a clear strengthening of the association, consistent with the theory.

While Figure 4 provides evidence of average changes across cohorts, one can glean further evidence of the link between intergenerational fertility associations and skill differentials in fertility by directly studying the relationship between the estimated associations. Figure 5 carries out such an analysis. On the left-hand side, the figure draws scatterplots and local linear regressions linking the intergenerational fertility association to the education-sibsize association. Both the scatterplot and non-parametric fit point negative relationships for both ever-born and surviving fertility, with the scatterplot also suggesting a role for regional variation. In particular, Latin American and Caribbean (LAC) cohorts have more negative education-sibsize associations and more positive intergenerational fertility associations. In the starkest cases (recent cohorts in the Andes), a women with 10 siblings can expect to have 5 years less education and 1 more child than a woman with no siblings. At the other end of the spectrum (older cohorts in Africa), the woman with 10 siblings would expect to have 3 years less education and $\frac{1}{3}$ fewer children than her counterpart without siblings. The intercepts of the local linear regressions are positive, reflecting a transmission coefficient B_{ct}^{a} that is on average greater than zero.

Despite the apparent importance of regional variation, within-country variation produces similar results. The right-hand side of Figure 5 plots residuals after regressing the estimated associations on country and cohort fixed effects. A negative slope obtains again, suggesting that as the education-sibsize assocation turns more negative across successive cohorts within a country, the intergenerational fertility association turns more positive. The LAC region reaches furthest into the southeast and northwest quadrants, implying that the simultaneous drop in the education-sibsize association and rise in the intergenerational fertility assocation were strongest in that region.

To place magnitudes on these negative slopes, Table 2 runs ordinary least squares regressions of the intergenerational fertility association on the education-sibsize association, for both ever-born and surviving fertility, with and without country and cohort fixed effects (columns [1]-[2] and [5]-[6]). All four variants of the regression produce significantly negative slopes, such that a one-unit decrease in the sibsize-education association leads to a $\frac{1}{10}$-unit increase in the intergenerational fertility association. Even more, the R^2 terms in the specifications without fixed effects (columns [1] and [5]) suggest that the sibsize-education association explains fully one-third of the variation in the intergenerational fertility association. Conditional on the fixed effects, this share drops to roughly one-sixth, as indicated by the within-R^2 terms in columns [2] and [6]. But even this share is notable, especially in light of the sampling variability in the association estimates. In unreported results, all four coefficients remain significantly negative if any one of the three regions is omitted from the sample.

The negative slopes in Figure 5 and the univariate regressions of Table 2 square well with Table 1's result that most education-fertility associations in the sample are negative. For most women in the sample, more education is associated with fewer children, so the intergenerational association has the opposite sign of the sibsize-education association. But equation (7) is also directly testable, by adding the education-fertility association and its interaction with the sibsize-education association to the regression. Equation (7) predicts coefficients of 0 on the individual associations and a coefficient of 1 on their interaction. Columns (3)-(4) and (7)-(8) of Table 2 confirm this prediction for both measures of fertility, with and without fixed effects. The coefficients on the individual associations are insignificantly different from 0, while the coefficients on the interactions are insignificantly different from 1.

To what extent do the preceding results reflect changes in the ages at marriage and first birth? The fact that intergenerational fertility associations measured early in the reproductive lifecycle strongly predict associations measured late in the reproductive lifecycle suggests that marriage and the initiation of childbearing may play an important role. Ordinary least squares estimation of equation(8) is inappropriate for answering this question because never-married and childless women have censored outcomes. As such, I estimate Cox hazard regressions in which the age at first marriage or first birth is the outcome, and surviving sibship size is the independent variable.[15]

---

[15] Results for ever-born sibship size are similar, so I do not report them to conserve space.

The hazard ratio from this regression conveys the proportional change in the hazard of either event from an additional sibling. Figure 6 presents the results in a format analogous to Figure 5, plotting the hazard ratio against the sibsize-education association on the left-hand side, and plotting residuals net of country and cohort fixed effects on the right-hand side. Both with and without country and cohort fixed effects, Figure 6 reveals a clear negative slope for both outcomes. When women with more siblings have *more* education, they marry and begin having children *later*, as reflected in hazard ratios below 1. In contrast, when such women have *less* education, they marry and begin having children *earlier*. The explanation for this pattern follows the same logic as that for Figure 5. Most women in the sample came of age in an era in which more education translated to later marriage and childbearing. As a result, for those born in the Malthusian era, sibship size is negatively related with marriage and childbearing risk, while for those born in the modern era, the opposite result obtains.

5 Composition Effects on Average Fertility

What do these strengthening intergenerational associtaions imply for the evolution of average fertility? This section estimates the composition effect of differential fertility among cohort t's parents on the average fertility of cohort t. It focuses in particular on how the composition effect varies with average fertility in cohort t, to shed light on how the dynamics of population composition affect the time series of aggregate fertility rates. To address issues of endogeneity, it uses a deweighted measure of average fertility—which undoes the reweighting of the population due to differential fertility—to track the stage of the fertility transition.

5.1 Methods

For empirical application, the composition effect in Section 2 requires three modifications. First, the model considered a population of one sex for analytical convenience, but the empirical measure must account for the reality of two sexes. As such, I refine the definition the composition effect to be the average number of children born to *women* minus the average number of grandchildren per *daughter* born to their mothers.[16] Second, Section 2 wrote the composition effect over the distribu-

---

[16] A more detailed definition might explicitly consider men and the role of assortative mating, but the DHS only offers data on women's sibship sizes, making such an extension impossible to estimate.

tion of fertility in the grandparent generation, but the DHS offers data on the women themselves. Thus, the definition of the composition effect we take to the data must involve two ingredients: the average fertility of women from cohort ct, $\bar{y}^a_{ct}$, as well as their deweighted average fertility, $\tilde{y}^a_{ct}$, which undoes the reweighting of the population due to differential fertility among their mothers. Let K be the maximum number of surviving females in a woman's sibship (including herself); η_{ctk} be the share of women from country-cohort ct with k surviving females in their sibships; and $\bar{y}^a_{ctk}$ be the average fertility of women with k surviving females in their sibships. Then:

$$\bar{y}^a_{ct} = \sum_{k=1}^{K} \eta_{ctk} \bar{y}^a_{ctk} \qquad \text{and} \qquad \tilde{y}^a_{ct} = \sum_{k=1}^{K} \left(\frac{\eta_{ctk}/k}{\sum_{l=1}^{K} \eta_{ctl}/l} \right) \bar{y}^a_{ctk}$$

Third, because changing fertility levels complicate comparisons of composition effects at different stages of the fertility transition and at different ages, I divide the absolute composition effect by the average fertility that would obtain in the absence of reweighting from differential fertility in the previous generation. This new measure, the *relative composition effect*, captures how the reweighting of the population affects average fertility in proportional terms.

For country-cohort ct, the relative composition effect on mean fertility at age a is:

$$\delta^a_{ct} = \frac{\bar{y}^a_{ct} - \tilde{y}^a_{ct}}{\tilde{y}^a_{ct}} \tag{11}$$

This expression is a simple non-linear combination of the weighted and deweighted averages. The non-parametric estimator for the relative composition effect plugs in the empirical analogs of those averages. Keeping with Section 4, the measure of fertility is cumulative fertility at age 25. For conciseness, I report results only for children ever born; unreported results for surviving fertility are extremely similar. Notably, however, $\tilde{y}^a_{ct}$ reweights on the number of *surviving* sisters only, as deceased sisters are irrelevant to the cohort's composition in adulthood.

Lemma 1 showed that in the case of a linear conditional expectation function, the composition effect equals the intergenerational association times the ratio of the variance to the mean of sibship size. To link the composition effect estimates with the fertility transmission results of the previous section, I report such decompositions using slopes from regressions of ever-born fertility on surviving sibship size.[17] For county-cohort ct, the decomposition is:

$$\delta^a_{ct} = \left(\frac{\beta_{ct}}{\tilde{y}_{ct}}\right)\left(\frac{\tilde{\sigma}^2_{ct}}{\tilde{s}_{ct}}\right) \tag{12}$$

where $\tilde{\sigma}^2_{ct}$ and $\tilde{s}_{ct}$ are the deweighted variance and mean of sibship size, respectively. The relative composition effect has two components: a relative intergenerational association and a ratio. If the intergenerational association becomes positive during fertility decline, but the ratio stays constant, then the composition effect rises. Each of these components is separately estimable, and their product provides a parametric estimator for the composition effect under the assumption of linearity.

5.2 Results

Figure 6 plots the composition effect and its components against deweighted average fertility, along with local linear regression estimates. In the top-left corner, non-parametric estimates of equation (11) show the composition effect averaging near zero for levels of deweighted average fertility above 1.5. But below 1.5, the composition effect rises, such that at the lowest observed levels of fertility, differential fertility in the previous generation raises current mean fertility by 4 percent on average and by as much as 10 percent. In the top-right corner, parametric estimates based on equation (12) point to a similar but more pronounced rise below 1.5. These results suggest that endogenous changes in population composition push back against aggregate fertility decline late in the fertility transition. As in Section 4, the growth of composition effects late in the fertility transition is especially apparent in the LAC region.

The bottom panels of Figure 6 decompose equation (12) into the association and ratio components, revealing that the rise in composition effects can be attributed entirely to rising intergenerational associations. Like the composition effect, the relative intergenerational fertility association ($\beta_{ct}/\tilde{y}_{ct}$) remains close to zero when deweighted average fertility is above 1.5, but it becomes positive as the population approaches low fertility. In contrast, the ratio of the variance to the mean of sibship size falls slightly as deweighted average fertility declines. Apparently, this fall is too weak to keep rising intergenerational associations from increasing the composition effect.

Consistent with the theory, some estimates of the intergenerational association and the compo-

---

[17] Results are extremely similar if one changes the covariate to the surviving number of female siblings, which would more appropriately capture the dynamics of a two-sex population but would have a less obvious relation to the intergenerational associations presented in Section 4.

sition effect are negative. However, the local linear regression estimates do not fall far below zero. This result is consistent with a positive transmission component that keeps both parameters near zero even when the human capital mechanism alone would cause them to be negative. Regardless, the general logic of the model still holds. Rising composition effects push back against fertility decline late in the fertility transition.

To associate magnitudes and significance levels with the local linear regression estimates, Table 4 runs a series of ordinary least regressions, in which the dependent variable is the non-parametric composition effect, and the independent variable is deweighted average fertility. Columns (1)-(2) include a single linear term in deweighted average fertility. Both with and without country and cohort fixed effects, the relationship is negative and significant at the 10% level. While these regressions are useful as summaries, the non-linearities apparent in Figure 6 suggest that they are mis-specified. To address these non-linearities, I divide the distribution of deweighted average fertility into bins $\frac{1}{3}$ of a child wide. With and without country and cohort fixed effects, the binned specification leads to highly significant results. All of the bin coefficients are negative (relative to the lowest bin), and all but one are significant at the 5 percent level. Thus, the rise in composition effects late in the fertility transition is statistically significant. Although Figure 6 suggests an important role for the LAC region, no single region drives these findings. As with the regression results in Table 3, the coefficients continue to be negative and statistically significant if any one of the three regions is omitted from the sample. Additionally, as indicated by the R^2 terms, this course semi-parametric specification explains a considerable share of the variation in composition effects. It explains 12 percent of the overall variation, and 21 percent of the variation net of country and cohort fixed effects.

6 Implications for Period Fertility Rates

While the cohort analysis has the ability to shed light on the evolution of composition effects within countries over time, it has an important limitation. Cumulative fertility at age 25 is an uncommon measure, making the cohort results somewhat difficult to map onto more familiar fertility metrics. This section estimates the implications for the total fertility rate (TFR), or the number of children a woman would expect to bear if she experienced current age-specific fertility rates throughout her reproductive lifecycle. TFR is a period measure of fertility, summing over the cross-sectional distribution of age-specific fertility rates. As a result, each DHS survey provides exactly one estimate of TFR and its composition effect.

6.1 Methods

To analyze TFR, redefine the outcome y^a_{ict} as an indicator for whether woman i of age a from country c and birth cohort t gave birth in the year preceding the survey. Let the τ be the survey year, so that the cohort average $\bar{y}^a_{ct}$ is an age-specific fertility rate for age a in year $\tau = t + a$. Similarly, the deweighted cohort average $\tilde{y}^a_{ct}$ is the age-specific fertility rate that would have obtained absent the influence of differential fertility on population composition. The total fertility rate and deweighted total fertility rate sum over these age-specific rates. For country c in year τ:

$$TFR_{c\tau} = 5 \textstyle\sum_a \bar{y}^a_{c,\tau-a} \qquad \text{and} \qquad dTFR_{c\tau} = 5 \textstyle\sum_a \tilde{y}^a_{c,\tau-a}$$

The summations are multiplied by five because I use five-year age groups, from 15-19 to 45-49.[18] The composition effect of differential fertility on TFR is the relative difference between actual TFR and the TFR that would obtain if women from different sibship sizes were equally represented within their age group:

$$\delta_{c\tau} = \frac{TFR_{c\tau} - dTFR_{c\tau}}{dTFR_{c\tau}} \tag{13}$$

As in the cohort analysis, this expression is a non-linear combination of weighted and deweighted averages. The empirical analog provides a non-parametric estimator for the composition effect on TFR.

For the linear case, the cohort analysis decomposed the composition effect into a relative intergenerational association and variance-mean ratio. For TFR, the decomposition becomes complicated because equation (13) sums 35 age-specific composition effects, each of which has a separate multiplicative decomposition. However, one can approximate the decomposition by assuming that either the association component or the ratio component is constant across age groups. Because the paper focuses on variation in intergenerational associations, I hold the ratio at its mean:

$$\delta_{c\tau} = \frac{\sum_a \beta^a_{c,\tau-a} \frac{\sigma^2_{c,\tau-a}}{\mu_{c,\tau-a}}}{dTFR_{c\tau}} \approx \left(\frac{5 \sum_a \beta^a_{c,\tau-a}}{dTFR} \right) \left(\frac{1}{7} \sum_a \frac{\tilde{\sigma}^2_{c,\tau-a}}{\tilde{\mu}_{c,\tau-a}} \right) \tag{14}$$

---

[18] Specifically, for age group a with lower bound $\underline{a}$ and upper bound $\overline{a}$, the birth cohort $\tau - a$ would be born between the year $\tau - \underline{a}$ and the year $\tau - \overline{a}$.

where $\beta^a_{c,\tau-a}$ is the coefficient from a regression of an indicator for a birth in the year preceding the survey on surviving sibship size for respondents aged a, while $\tilde{\sigma}^2_{c,\tau-a}$ and $\tilde{\mu}_{c,\tau-a}$ are the deweighted variance and mean of sibship size, respectively. In the approximate decomposition of equation (14), $\delta_{c\tau}$ has two parts: a relative association and a ratio. In the first, the total fertility association $5\sum_a \beta^a_{c,\tau-a}$, which mimics the total fertility rate by summing over the cross-sectional distribution of age-specific fertility associations, is divided by the deweighted TFR. In the second, the ratio of the variance to the mean of sibship size is averaged across the seven age groups. As in Section 5, the product of these components provides a parametric estimator for the composition effect, which in this case is approximate in the linear case.

6.2 Results

For each of the 81 surveys in the sample (not including Nepal 1996, which lacks information on younger, unmarried women), Figure 7 plots the composition effect and its components against deweighted TFR, along with local linear regression estimates. The results mirror the cohort results in the previous section. Non-parametric estimates of the composition effect average near zero when deweighted TFR exceeds 4, but below this threshold, they turn positive. At the lowest observed levels of deweighted average fertility, differential fertility in the previous generation raises average fertility by 3 percent on average, and by as much as 6 percent. Parametric estimates based on the approximation in equation (14) display a similar pattern, albeit with larger magnitudes. Thus, although the approximation overstates the composition effect, it captures the important features of the evolution of the composition effect over the fertility transition. Here again, the association component rather than the ratio compnent drives the shape of this evolution. When deweighted TFR falls below 4 in the bottom panels of Figure 7, the total fertility association rises dramatically as a share of deweighted TFR, but the average ratio of the variance to the mean of sibship size does not appreciably change.

To capture magnitudes and significance, Table 4 carries out an analysis of these period parameter estimates that closely follows the cohort analysis of Table 3. In columns (1) and (2), it runs regressions in which the dependent variable is the composition effect on TFR and the independent variable is deweighted TFR. In the specification without country and cohort fixed effects, the coefficient is negative and statistically significant, implying that a one-child reduction in deweighted

TFR raises the composition effect by nearly 1 percentage point. The addition of country and cohort fixed effects reduces precision considerably—the fixed effects use 59 of the available 81 degrees of freedom— but the point estimate remains similar. Because graphical evidence again suggested non-linearities, columns (3)-(4) divide the distribution of deweighted TFR into bins one child wide. The specification without fixed effects again reveals that composition effects are significantly lower at higher levels of deweighted TFR, relative to the reference category of fewer than 3 children. The R^2 indicates that this coarse binned specification explains more than a quarter of the variation in the composition effect on TFR. And although the standard errors grow and the within R^2 shrinks with the addition of fixed effects, the point estimates remain negative and are sometimes marginally significant.

Overall, the analysis of period fertility rates has similar conclusions to the analysis of cohort age-25 fertility rates. In the later phases of the process of aggregate fertility decline, the composition effect slows the process by raising average fertility. With magnitudes of 3-4 percent on average and 6-10 percent in the most pronounced cases, these results explain a small but systematic feature of aggregate fertility dynamics.

7 Conclusion

The question of whether girls from large families tend to bear many children in adulthood has drawn the attention of social scientists and statisticians for over a century. Nevertheless, existing research has shed limited light on how this association evolves over the course of the fertility transition, and on what implications it has for aggregate fertility outcomes. This paper fills that gap on two levels. First, it first theoretically explores the evolution of the association in a quality-quantity framework drawn from the literature on demography and long-run economic growth. The theoretical results highlight the importance of the changing relationship between human capital and fertility. Intergenerational fertility associations—and therefore also composition effects—turn positive when skill differentials in fertility are aligned across generations, which occurs within the Malthusian or modern regimes but not during the transition between them. As a result, differential fertility initially speeds the decline of aggregate fertility but then slows it. Second, the empirical analysis confirms these predictions in data on mother-daughter pairs from 40 developing countries over half a century. In these data, intergenerational associations grow during the transition to a modern fertility regime, as the link between human capital and fertility flips from positive to negative, and the education-fertility association becomes more negative. Due to this change in intergenerational associations, the composition effect of differential fertility on average fertility turns positive as the fertility transition progresses. The paper's approach provides a new demographic lens for understanding the global decline of fertility, the subject of a large literature in economics (Schultz 1997; Galor 2011; de la Croix 2012).

From a broader economic perspective, the results are of interest not just because they clarify aggregate fertility dynamics but also because they demonstrate a unique way in which intergenerational associations aggregate up to population-level dynamics. Mother-daughter associations in reproductive behavior are the subject of a long-standing literature, but few contributions to this literature have formally considered their aggregate implications. In this sense, this paper is related to the literature exploring how the intergenerational transmission of economic status affects the dynamics of the income distribution (Becker and Tomes 1979; Goldberger 1989).

While the paper breaks new ground on the changing causes and consequences of fertility transmission, several questions remain open. First, although fertility differentials by female education clearly influence intergenerational fertility associations, other sources of intergenerational persistence, such as fecundity and culture, may also play a role. A fuller model of the evolution of intergenerational transmission would also take these factors into account, although data limitations hinder tests of such a model. Second, although the paper provides evidence that changes in differential fertility and population composition interact to increase the composition effect of differential fertility, future research could further draw out the connections to the pace of fertility decline, and especially to the recent "stalls" Bongaarts (2006, 2008) has identified in many populations. Finally, the analysis here has not considered other demographic influences on the time path of aggregate fertility, such as assortative mating, or intergenerational influences that span more than two generations, as Mare (2011) proposes. Investigations into the consequences of these other mechanisms may shed light on the micro-level demographic phenomena influencing the pace of aggregate change.

References

Anderton, Douglas L., Noriko O. Tsuya, Lee L. Bean, and Geraldine P. Mineau. (1987). "Intergenerational Transmission of Relative Fertility and Life Course Patterns." *Demography* 24(4): 467-480.

Axinn, William G., Marin E. Clarkberg, and Arland Thornton. (1994). "Family Influences on Family Size Preferences." *Demography* 31(1): 65-79.

Becker, Gary S., and Nigel Tomes. (1979). "An Equilibrium Theory of rhe Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87(6): 1153-1189.

Ben-Porath, Yoram. (1975). "First-Generation Effects on Second-Generation Fertility." *Demography* 12(3): 397-405.

Bengtsson, Tommy, and Martin Dribe. (2014). "The Fertility Transition at the Micro Level: Socioeconomic Status, Income and Fertility in Southern Sweden 1815-1939." *Demographic Research* 30(17): 493-534.

Bhalotra, Sonia, and Samantha Rawlings. (2013). "Gradients of the Intergenerational Transmission of Health in Developing Countries." *Review of Economics and Statistics* 95(2): 660-672.

Björklund, Anders, and Markus Jäntti. (1997). "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review* 87(5): 1009-1018.

Bongaarts, John. (2006). "The Causes of Stalling Fertility Transitions." *Studies in Family Planning* 37(1): 1-16.

Bongaarts, John. (2008). "Fertility Transitions in Developing Countries: Progress Or Stagnation?" *Studies in Family Planning* 39(2): 105-110.

Caldwell, John C. (1982). *Theory of Fertility Decline*. New York: Academic Press.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. (2014). "Where Is The Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics* 129(4): 1553-1623.

Chevalier, Arnaud, Kevin Denny, and Dorren McMahon. (2009). "A Multi-Country Study Of Intergenerational Educational Mobility." In Peter Dolton, Rita Asplundh and Erling Barth, eds., *Education and Inequality Across Europe*. London: Edward Elgar.

Corak, Miles. (2013). "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives* 27(3): 79-102.

Clark, Gregory. (2007). *A Farewell to Alms: A Brief Economic History of the World*. Princeton: Princeton University Press.

Clark, Gregory, and Neil Cummins. (2015). "Malthus to Modernity: Wealth, Status, and Fertility in England, 1500–1879." *Journal of Population Economics* 28: 3-29.

De la Croix, David and Matthias Doepke. (2003). "Inequality and Growth: Why Differential Fertility

Matters." *American Economic Review* 93: 1091-1113.

De la Croix, David. (2013). *Fertility, Education, Growth, and Sustainability*. Cambridge: Cambridge University Press.

Doepke, Matthias. (2005). "Child Mortality and Fertility Decline: Does the Barro-Becker Model Fit the Facts?" *Journal of Population Economics* 18(2): 337-366.

Duncan, Otis Dudley, Ronald Freedman, J. Michael Coble, and Doris P. Slesinger. (1965). "Marital Fertility and Size of Family of Orientation." *Demography* 2(1): 508-515.

Fernandez, Raquel, and Richard Rogerson. (2001). "Sorting and Long-Run Inequality." *Quarterly Journal of Economics* 101: 1305-1341.

Galor, Oded. (2011). *Unified Growth Theory*. Princeton, NJ: Princeton University Press.

Galor, Oded, and Omer Moav. (2002). "Natural Selection and the Origin of Economic Growth." *Quarterly Journal of Economics* 117: 1133-1192.

Galor, Oded, and Stelios Michalopoulos. (2012). "Evolution and the Growth Process: Natural Selection of Entrepreneurial Traits." *Journal of Economic Theory* 147(2): 759-780.

Goldberger, Arthur S. (1989). "Economic and Mechanical Models of Intergenerational Transmission." *American Economic Review* 79(3): 504-513.

Haider, Steven, & Solon, Gary. (2006). "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4): 1308-1320.

Hazan, Moshe, and Binyamin Berdugo. (2002). "Child Labour, Fertility, and Economic Growth." *Economic Journal* 112(482): 810-828.

Huestis, R.R., and Aline Maxwell. (1932). "Does Family Size Run in Families?" *Journal of Heredity* 23(2): 77-79.

Jones, Larry E., Alice Schoonbroodt, and Michèle Tertilt. (2010). "Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?" In J.B. Shoven, ed., *Demography and the Economy*. Chicago: University of Chicago Press, pp. 43-100.

Kahn, Joan R., and Kay E. Anderson. (1992). "Intergenerational Patterns of Teenage Fertility." *Demography* 29(1): 39-57.

Keyfitz, Nathan. (1971). "On the Momentum of Population Growth." *Demography* 8: 71-80.

Kolk, Martin, Daniel Cownden, and Magnus Enquist. (2014). "Correlations in Fertility Across Generations: Can Low Fertility Persist?" *Proceedings of the Royal Society B: Biological Sciences* 281(1779): 20132561.

Kremer, Michael. (1997). "How Much Does Sorting Increase Inequality?" *Quarterly Journal of Economics* 97: 115-139.

Lam, David. (1986). "The Dynamics of Population Growth, Differential Fertility, and Inequality."

American Economic Review 76(5): 1103-1116.

Lam, David, and Letícia Marteleto. (2013). “Family Size of Children and Women During the Demographic Transition.” Mimeo, University of Michigan.

Mare, Robert D. (2011). “A Multigenerational View of Inequality.” *Demography* 48(1): 1-23.

Moav, Omer. (2005). “Cheap Children and the Persistence of Poverty” *Economic Journal* 115(500): 88-110.

Murphy, Michael. (1999). “Is The Relationship Between Fertility of Parents snd Children Really Weak?” *Biodemography and Social Biology* 46(1-2): 122-145.

Murphy, Michael. (2012). “Intergenerational Fertility Correlations in Contemporary Developing Counties.” *American Journal of Human Biology* 24(5): 696-704.

Murphy, Michael, and Duolao Wang. (2003). “The Impact of Intergenerationally-Transmitted Fertility and Nuptuality on Population Dynamics in Contemporary Populations.” In Joseph L. Rodgers and Hans-Peter Kohler, eds., *Biodemography of Human Reproduction and Fertility*. Boston: Kluwer Academic Publishers, pp. 209–228.

Pearson, Karl, Alice Lee, and Leslie Bramley-Moore. (1899). “Inheritance of Fertility in Man, and of Fecundity in Thoroughbred Racehorses.” *Philosophical Transactions of the Royal Society of London, Series A* 192(1899): 257-330.

Preston, Samuel H. (1976). “Family Sizes of Children and Family Sizes of Women.” *Demography* 13(1): 105-114.

Schultz, T. Paul. (1997). “Demand for Children in Low Income Countries.” In Mark Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*. Amsterdam: Elsevier, pp. 349-430.

Solon, Gary. (1999). “Intergenerational Mobility in the Labor Market.” In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics Vol. 3*. Amsterdam: North-Holland.

Treloar, S.A., and N.G. Martin (1990). “Age at Menarche as a Fitness Trait: Nonadditive Genetic Variance Detected in a Large Twin Sample.” *American Journal of Human Genetics* 47(1): 137-148.

Van den Akker, O.B.A., G.S. Stein, M.C. Neale, and R.M. Murray. (1987). “Genetic and Environmental Variation in Menstrual Cycle: Histories of Two British Twin Samples.” *Acta Geneticae Medicae Et Gemellologiae: Twin Research* 36(4): 541-548.

Vogl, Tom. (2016). “Differential Fertility, Human Capital, and Development.” *Review of Economic Studies*, forthcoming.

Wise, Jeffrey M., and Spencer J. Condie. (1975). “Intergenerational Fertility Throughout Four Generations." *Biodemography and Social Biology* 22(2): 144-150.

Theory Appendix

Proof of Lemma 1

Let $E\left(n_{i,t+1}|n_{it}\right) \equiv a + bn_{it}$. Then:

$$\begin{aligned}
\Delta_{t+1} &= E_t\left[\left(\tfrac{n_{it}}{\mu_t}\right) E\left(n_{i,t+1}|n_{it}\right)\right] - E_t\left[E\left(n_{i,t+1}|n_{it}\right)\right] \\
&= E_t\left[\left(\tfrac{n_{it}}{\mu_t}\right)\left(a + bn_{it}\right)\right] - E_t\left[a + bn_{it}\right] \\
&= \tfrac{b}{\mu_t} E_t\left[n_{it}^2\right] - bE_t\left[n_{it}\right] \\
&= b\frac{E_t\left[n_{it}^2\right] - bE_t\left[n_{it}\right]^2}{\mu_t} \\
&= b\frac{\sigma_t^2}{\mu_t}
\end{aligned}$$

Since β_{t+1} is the coefficient from the linear projection of $n_{i,t+1}$ on n_{it}, $\beta_{t+1} = b$. Then $\Delta_{t+1} = \beta_{t+1}\frac{\sigma_t^2}{\mu_t}$.

Proof of Lemma 2

Differentiation of equation (5) leads to:

$$\frac{\partial n_{it}}{\partial h_{it}} = \begin{cases} \frac{(1-\alpha)\kappa A}{\left(\kappa+\tau A h_{it}\right)^2} > 0 & \text{if } h_{it} \leq \widetilde{h} \\ \frac{(1-\alpha)(1-\eta)(\kappa-p\theta)A}{\left(\kappa-p\theta+\tau A h_{it}\right)^2} < 0 & \text{if } h_{it} > \widetilde{h} \end{cases}$$

Proof of Lemma 3

$\phi\left(\cdot\right)$ has domain $\mathbb{R}_+$, is a constant function for $h_{it} < \widetilde{h}$, and is a power function for $h_{it} > \widetilde{h}$, implying continuity for all $h_{it} \neq \widetilde{h}$. Further, $\lim_{h_{it}\downarrow\widetilde{h}}\left\{\left(\frac{\eta}{1-\eta}\right)\left(\frac{\kappa+\tau A h_{it}}{p} - \theta\right)\right\}^{\eta} = \theta^{\eta}$, proving continuity on $\mathbb{R}_+$. Differentiation of equation (6) leads to:

$$\phi'\left(h_{it}\right) = \begin{cases} 0 & \text{if } h_{it} \leq \widetilde{h} \\ \eta\left(\frac{\tau A}{p}\right)\left(\frac{\eta}{1-\eta}\right)^{\eta}\left(\frac{\kappa+\tau A h_{it}}{p} - \theta\right)^{\eta-1} > 0 & \text{if } h_{it} > \widetilde{h} \end{cases}$$

where the second line is positive because $\frac{\kappa+\tau A h_{it}}{p} - \theta > \frac{\kappa+\tau A \widetilde{h}}{p} - \theta = \theta\left(\frac{1}{\eta} - 1\right) > 0$. In the limit, $\lim_{h_{it}\to\infty}\phi'\left(h_{it}\right) = \lim_{h_{it}\to\infty}\eta\left(\frac{\tau A}{p}\right)\left(\frac{\eta}{1-\eta}\right)^{\eta}\left(\frac{\kappa+\tau A h_{it}}{p} - \theta\right)^{\eta-1} = 0$.

Further differentiation leads to:

$$\phi''\left(h_{it}\right) = \begin{cases} 0 & \text{if } h_{it} \leq \widetilde{h} \\ \eta\left(\eta-1\right)\left(\frac{\tau A}{p}\right)^{2}\left(\frac{\eta}{1-\eta}\right)^{\eta}\left(\frac{\kappa+\tau A h_{it}}{p}-\theta\right)^{\eta-2} < 0 & \text{if } h_{it} > \widetilde{h} \end{cases}$$

Proof of Proposition 1

To prove that at least one steady state exists, note:

- If $\widetilde{h} > \theta^{\eta}$, then $\phi\left(\cdot\right)$ must cross the 45 degree line from above at $h_{it} = \theta^{\eta}$ because $\phi\left(h_{it}\right) = \theta^{\eta}$ for all $h_{it} \leq \widetilde{h}$.
- If $\widetilde{h} \leq \theta^{\eta}$, then $\phi\left(\cdot\right)$ must cross the 45 degree line from above at some $h_{it} > \theta^{\eta}$ because $\phi\left(\cdot\right)$ is continuous, $\phi\left(\theta^{\eta}\right) > \theta^{\eta}$, and $\lim_{h_{it}\to\infty}\phi'\left(h_{it}\right) = 0$.

In each of these steady states, $\phi\left(\cdot\right)$ crosses the 45 degree line from above, so at least one steady state is stable. To prove that the number of steady states cannot exceed three, note:

- For $h_{it} > \widetilde{h}$, $\phi\left(\cdot\right)$ is strictly increasing and strictly concave, so it can cross or be tangent to the 45 degree line at most twice, once from below with $\phi' > 1$ and once from above with $\phi' < 1$.
- For $h_{it} \leq \widetilde{h}$, $\phi\left(\cdot\right)$ is constant, so it can cross the 45 degree line from above or be tangent to it only once.

Because $\phi\left(\cdot\right)$ can cross the 45 degree line from above at most once for $h_{it} > \widetilde{h}$ and at most once for $h_{it} \leq \widetilde{h}$, at most two steady states are stable.

Proof of Proposition 2

An increase in A or decrease in p reduces the corner solution threshold $\frac{1}{\tau A}\left(\frac{p\theta}{\eta}-\kappa\right)$. Call this threshold $\hat{h}$, so that $\hat{h} < \widetilde{h}$. Then differentiation of equation (5) leads to:

$$\frac{\partial n_{it}}{\partial A} = \begin{cases} \frac{(1-\alpha)\kappa h_{it}}{(\kappa+\tau A h_{it})^{2}} > 0 & \text{if } h_{it} \leq \hat{h} \\ \frac{(1-\alpha)(1-\eta)h_{it}(\kappa-p\theta)}{(\kappa-p\theta+\tau A h_{it})^{2}} < 0 & \text{if } h_{it} > \widetilde{h} \end{cases}$$

and:

$$\frac{\partial n_{it}}{\partial p} = \begin{cases} 0 & \text{if } h_{it} \leq \hat{h} \\ \frac{\theta(1-\alpha)(1-\eta)A h_{it}}{(\kappa-p\theta+\tau A h_{it})^{2}} > 0 & \text{if } h_{it} > \widetilde{h} \end{cases}$$

Differentiation of equation (6) leads to:

$$\frac{\partial \phi}{\partial A} \equiv \begin{cases} 0 & \text{if } h_{it} \leq \hat{h} \\ \left(\frac{\eta}{1-\eta}\right)^{\eta} \left(\frac{\kappa+\tau A h_{it}}{p} - \theta\right)^{\eta-1} \frac{\eta \tau h_{it}}{p} > 0 & \text{if } h_{it} > \widetilde{h} \end{cases}$$

and:

$$\frac{\partial \phi}{\partial p} \equiv \begin{cases} 0 & \text{if } h_{it} \leq \hat{h} \\ -\frac{\eta}{p} \left(\frac{\eta}{1-\eta}\right)^{\eta} \left(\frac{\kappa+\tau A h_{it}}{p} - \theta\right)^{\eta-1} \left(\frac{\kappa+\tau A h_{it}}{p}\right) < 0 & \text{if } h_{it} > \widetilde{h} \end{cases}$$

First consider lineages initially in the low human capital stable steady state, $h_{it} = \theta^{\sigma}$. Since the population attains a new non-degenerate fertility regime, $\hat{h} > \theta^{\eta}$. Thus, for these lineages, $\phi(\cdot)$ and therefore h_{it} do not change. In the case of higher A, these lineages choose higher n_{it}; in the case of lower p, n_{it} is unchanged.

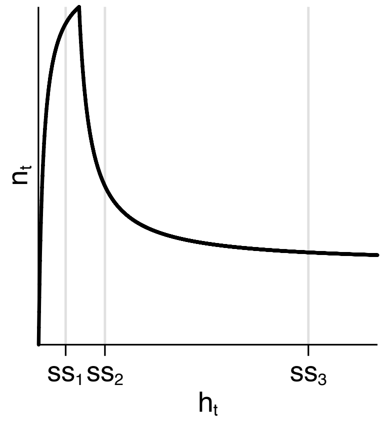
Next consider lineages initially in the high human capital stable steady state, with $h_{it} > \widetilde{h}$. $\phi(\cdot)$ rises from higher A or lower p, so h_{it} rises with t, and $h_{it} > \widetilde{h}$ in every period. Holding h_{it} constant, n_{it} falls from higher A and lower p. Since h_{it} rises with t and $\frac{\partial n_{it}}{\partial h_{it}} < 0$ for $h_{it} > \widetilde{h}$, n_{it} falls further. Thus, for these lineages, n_{it} is lower in the transition and the new regime.

Proof of Proposition 3

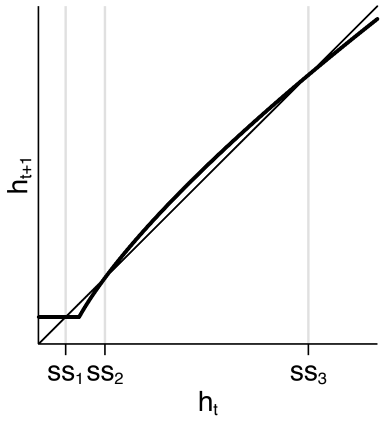
Denote the stable steady state human capital levels in generation t as h_{0t} and h_{1t}, with $h_{0t} < h_{1t}$, and let n_{0t} and n_{1t} denote the associated fertility levels. Let $h_{0,t+1}$, $h_{1,t+1}$, $n_{0,t+1}$, and $n_{1,t+1}$ be the same outcomes for the next generation of the 0 and 1 lineages. Then the intergenerational transmission coefficient is $\beta_{t+1} = \frac{n_{1,t+1} - n_{0,t+1}}{n_{1t} - n_{0t}}$. If the child cost parameters and the human capital production function is the same in both t and $t+1$, then $h_{i,t+1} = h_{it}$ and $n_{i,t+1} = n_{it}$ for all lineages, so $\beta_{t+1} = 1$. In the transition between two Malthusian regimes, $n_{0t} < n_{1t}$, $n_{0,t+1} < n_{1,t+1}$, $n_{0,t+1} \geq n_{0t}$, and $n_{1,t+1} < n_{1t}$, so $\beta_{t+1} \in (0,1)$. In the transition between two modern regimes, $n_{0t} > n_{1t}$, $n_{0,t+1} > n_{1,t+1}$, $n_{0,t+1} \geq n_{0t}$, and $n_{1,t+1} < n_{1t}$, so $\beta > 1$. During a transition between Malthusian and modern regimes, there exists t such $n_{0t} > n_{1t}$, $n_{0,t+1} \leq n_{1,t+1}$, $n_{0,t+1} \geq n_{0t}$, and $n_{1,t+1} < n_{1t}$, so $\beta_{t+1} \leq 0$. Because the skill distribution has two mass points in any non-degenerate regime and any transition between non-degenerate regimes, $E(n_{t+1} | n_t)$ can be fully characterized by a linear function, and $\sigma_t^2 > 0$. Therefore, $\Delta_{t+1} = \beta_{t+1} \frac{\sigma_t^2}{\mu_t}$. Since $\mu_t > 0$, $\text{sgn}(\Delta_{t+1}) = \text{sgn}(\beta_{t+1})$.

Figure 1: Fertility Determination and the Dynamics of Human Capital

A. Optimal fertility

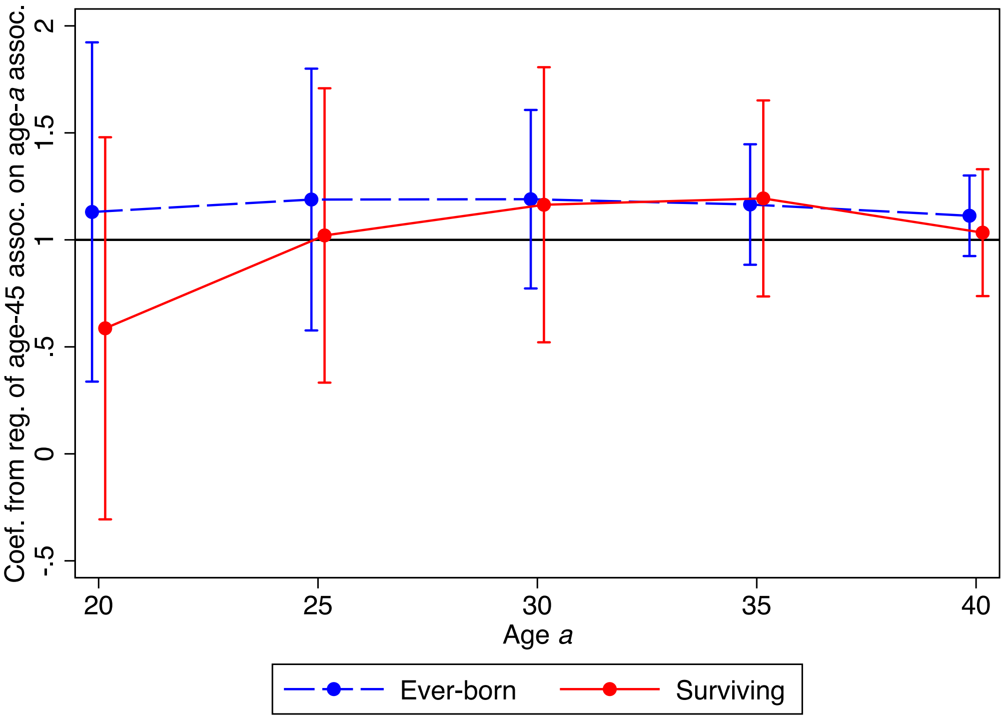


B. Dynamical system



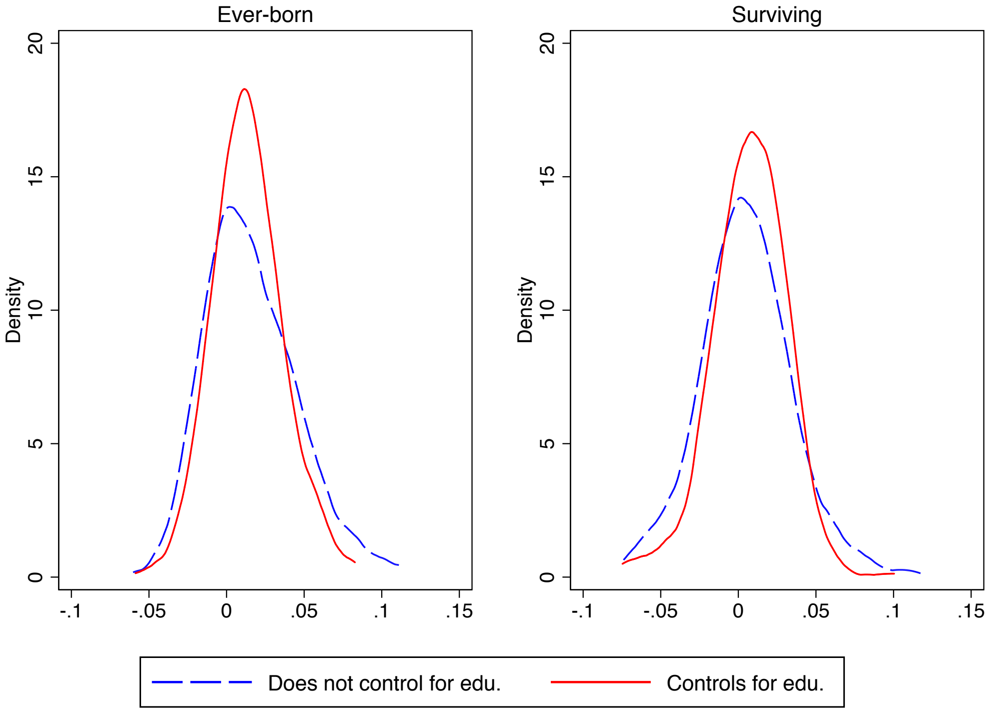
Note: The figures depict a parameterization of the model in which steady states with higher human capital have lower fertility. A modern fertility regime would obtain if all lineages were located at ss_1 or ss_3.

Figure 2: Choosing the Age at Fertility Measurement



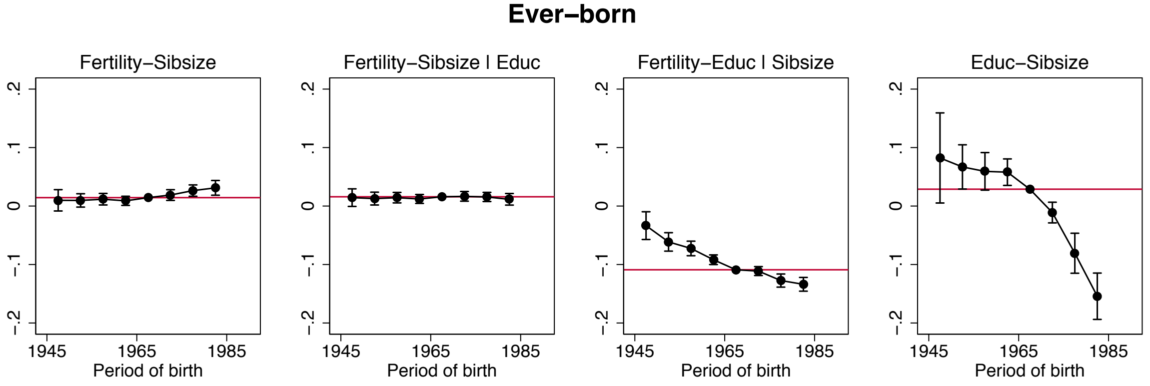
Note: Sample includes 96 country-cohort cells. Each coefficient and 95% confidence interval comes from a separate regression of the fertility association at age 45 on the fertility association at an earlier age. Capped spikes represent 95% confidence intervals based on standard errors clustered at the country level.

Figure 3: Distribution of Intergenerational Fertility Associations

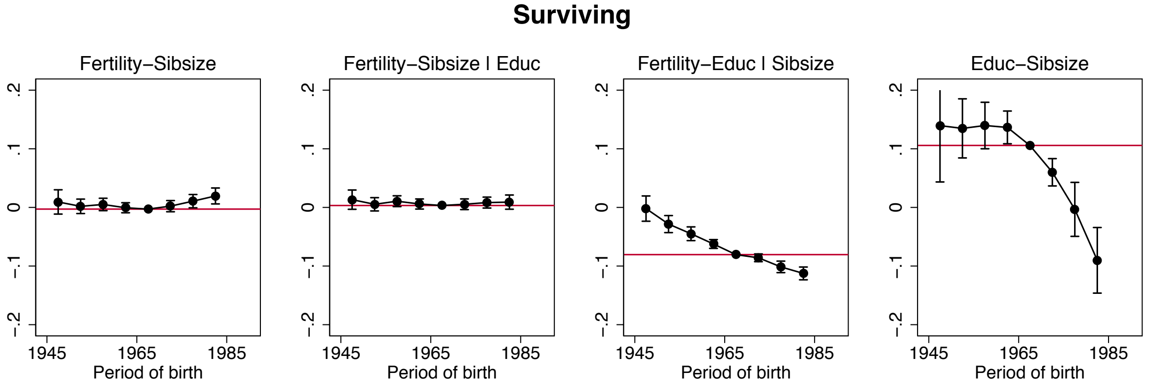


Note: Kernel density estimates with a bandwidth of 0.01. Sample includes 554,182 women from 257 country-cohort cells. Fertility associations are based on cumulative fertility at age 25.

Figure 4: Cohort Effects in Parameter Estimates



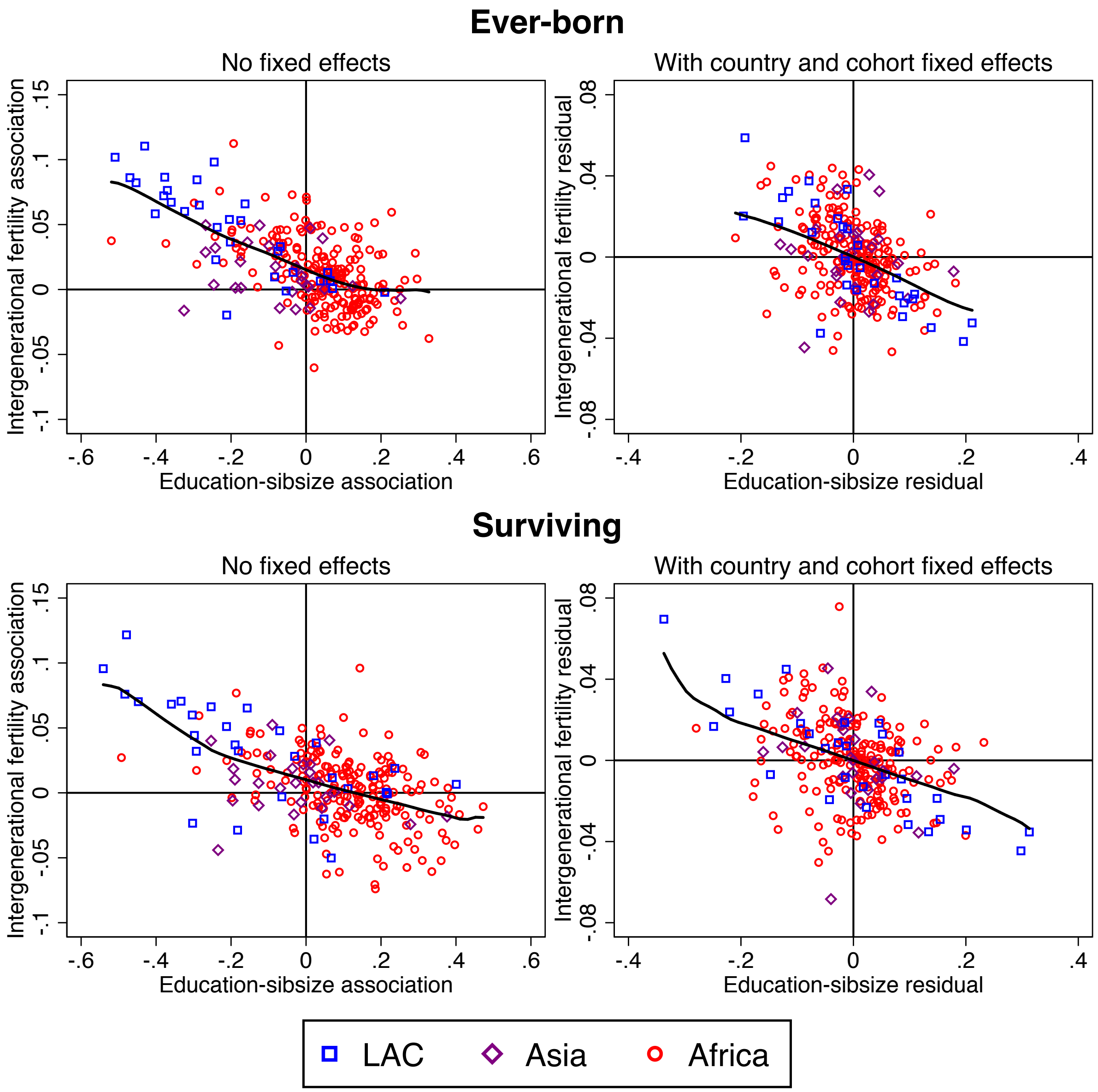
Ever–born



Surviving

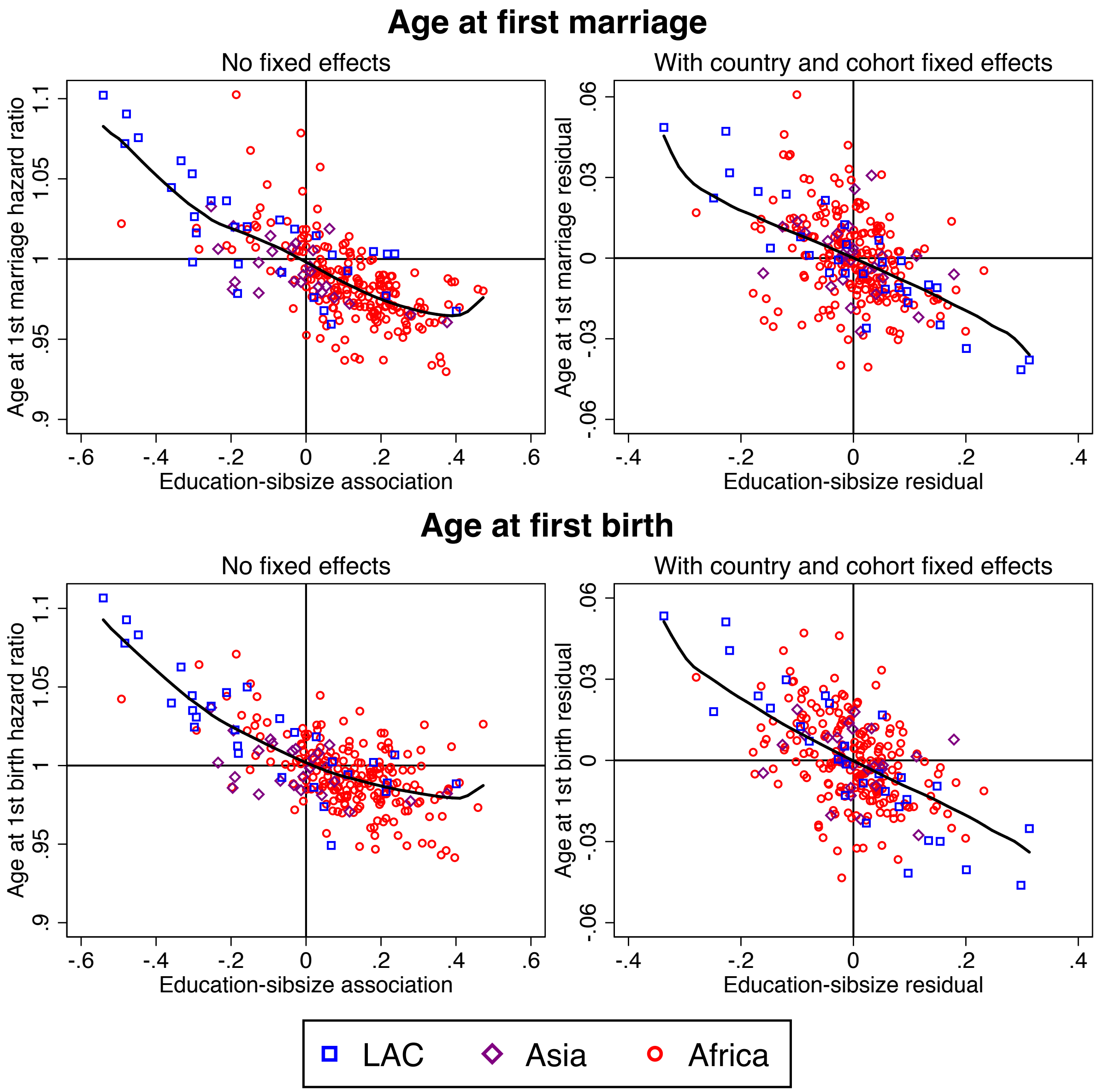
Note: Sample includes 259 county-cohorts cells. Plots represent cohort effects from regressions of the estimated country-cohort parameters on country and cohort fixed effects. Capped spikes represent 95% confidence intervals based on standard errors clustered at the country level. Omitted category is 1965-69, the only cohort with data available from all countries. Horizontal lines are average parameter values across countries for the 1965-69 cohort. Fertility associations are based on cumulative fertility at age 25.

Figure 5: Intergenerational Fertility Associations vs. Sibsize-Education Associations



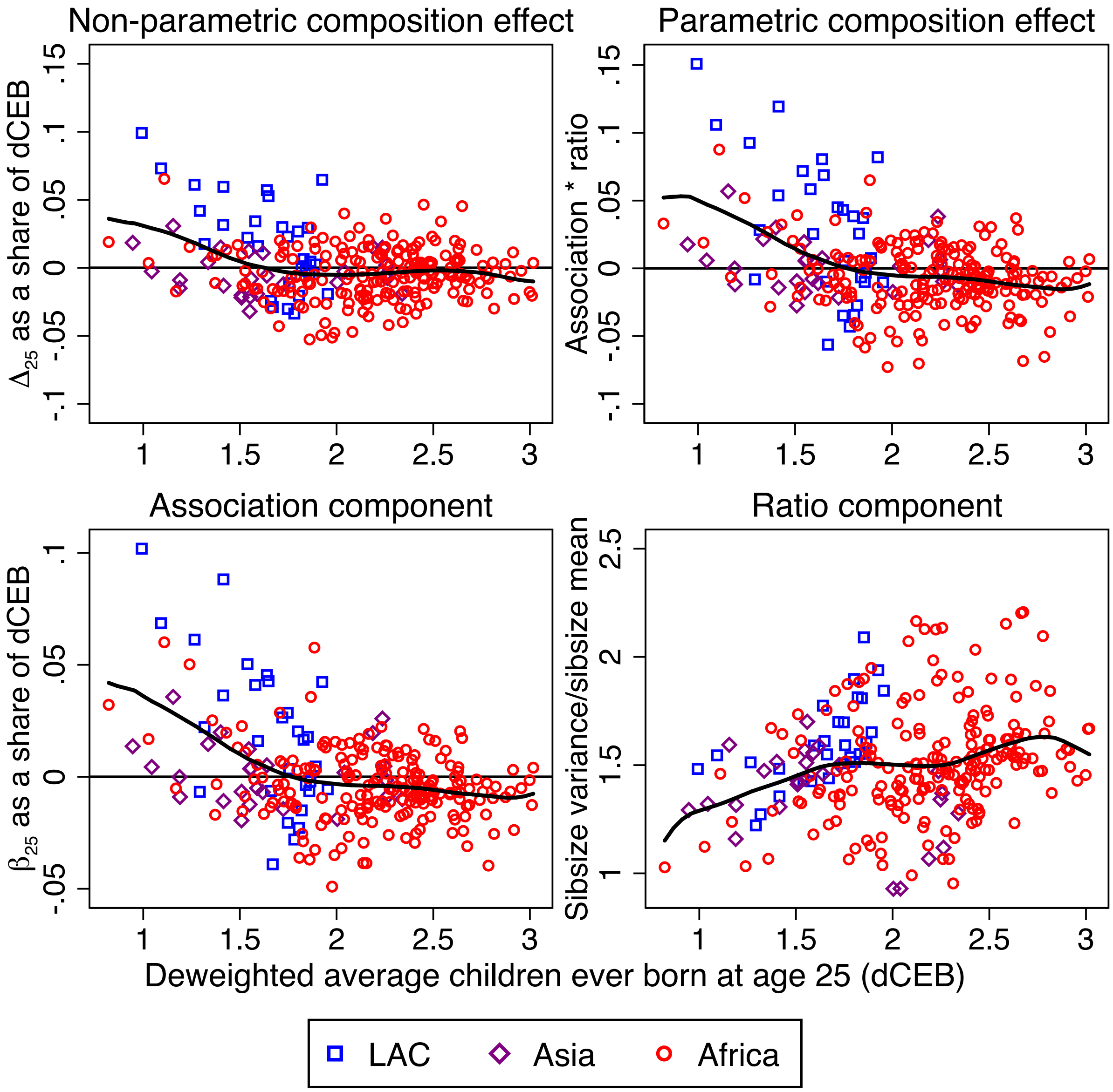
Note: Sample includes 259 county-cohorts cells. Black curves are local linear regressions with bandwidths of 0.1. Fertility associations are based on cumulative fertility at age 25.

Figure 6: Role of Early Earlier Marriage and Childbearing



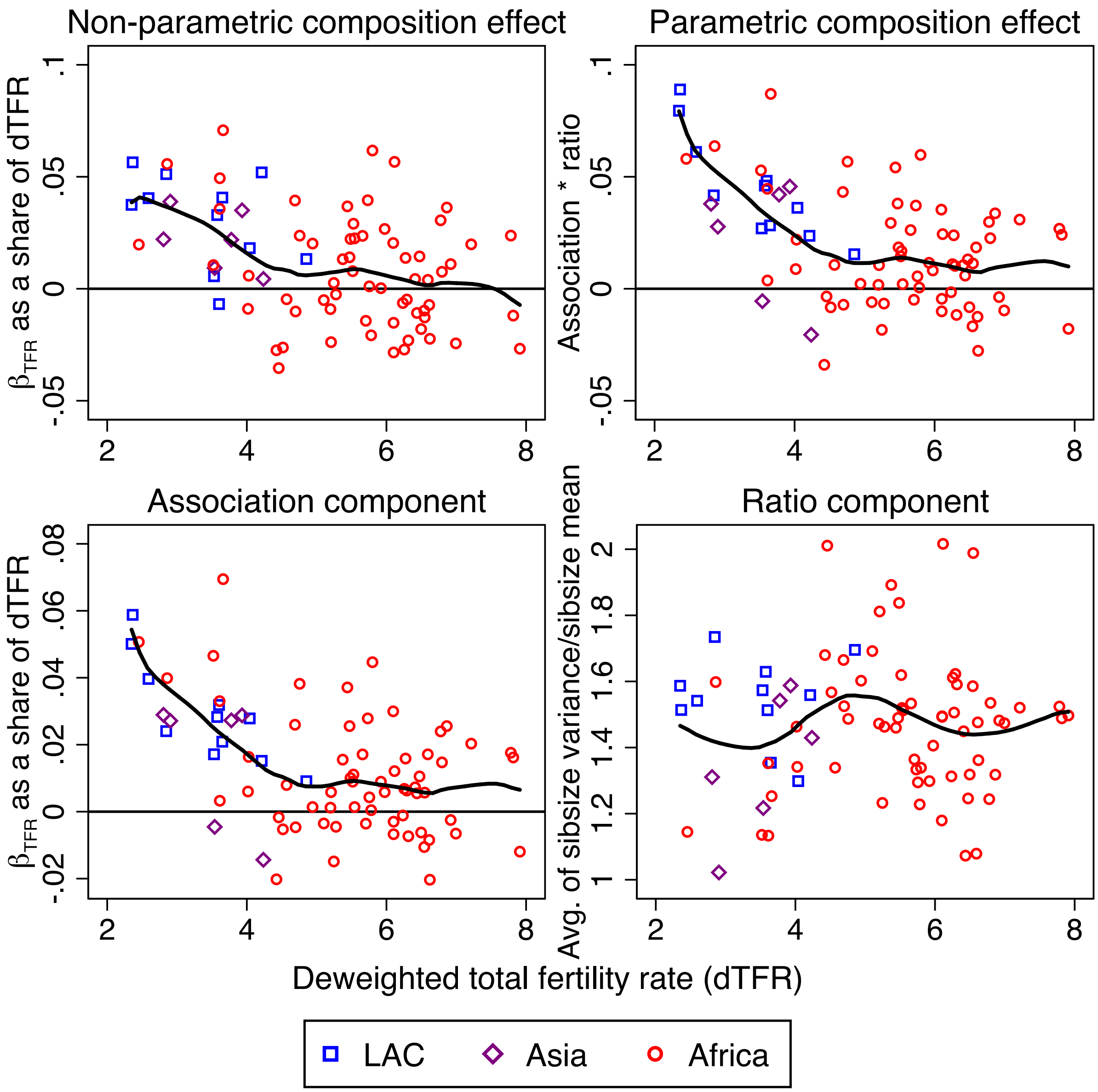
Note: Sample includes 259 county-cohorts cells. Black curves are local linear regressions with bandwidths of 0.1. All estimations use surviving sibship size.

Figure 7: Composition Effects over the Fertility Transition



Note: Sample includes 259 county-cohort cells. Curves are local linear regressions with a bandwidth of 0.2. Estimates are based on based on cumulative ever-born fertility at age 25.

Figure 8: Composition Effects on the Total Fertility Rate



Note: Sample includes 81 surveys. Curves are local linear regressions with a bandwidth of 0.5.

Table 1: Summary of Parameter Estimates

		Distribution of *t*-statistics		
	Mean [S.E.] (S.D.)	$t \le -1.96$	$-1.96 < t < 1.96$	$t \ge 1.96$
	(1)	(2)	(3)	(4)
A. Ever-born				
$\boldsymbol{\beta}^{25}$	.016 [.020] (.029)	3 (1%)	183 (71%)	73 (28%)
B^{25}	.013 [.019] (.022)	3 (1%)	197 (76%)	59 (23%)
$\boldsymbol{\Gamma}^{25}$	-.098 [.018] (.049)	230 (89%)	27 (1%)	2 (1%)
$\boldsymbol{\varphi}^{25}$	.002 [.043] (.157)	68 (26%)	93 (36%)	98 (38%)
B. Surviving				
$\boldsymbol{\beta}^{25}$	.005 [.021] (.030)	24 (9%)	195 (76%)	40 (15%)
B^{25}	.007 [.021] (.023)	11 (4%)	211 (81%)	37 (14%)
$\boldsymbol{\Gamma}^{25}$	-.070 [.016] (.051)	204 (79%)	50 (19%)	5 (2%)
$\boldsymbol{\varphi}^{25}$	.076 [.052] (.179)	39 (15%)	87 (34%)	132 (51%)

Note: Sample includes 586,283 women from 259 country-cohort cells with at least 250 observations. S.E. = standard error of the mean parameter. S.D. = standard deviation of the parameter. Fertility associations are based on cumulative fertility at age 25.

Table 2: Relating Intergenerational Fertility Associations to Other Associations

	Ever-born fertility				Surviving fertility			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sibsize-edu. association	-0.11	-0.12	0.03	-.004	-0.09	-0.10	.004	-0.01
	[0.02]	[0.03]	[0.03]	[0.04]	[0.01]	[0.02]	[0.02]	[0.03]
Edu.-fertility association			-0.04	-0.11			-0.02	-0.12
			[0.05]	[0.07]			[0.05]	[0.10]
Interaction			1.23	1.13			1.09	0.95
			[0.23]	[0.23]			[0.19]	[0.23]
Overall R-squared	0.37	0.66	0.44	0.69	0.28	0.60	0.37	0.64
Within R-squared		0.19		0.27		0.13		0.25
Country & Cohort FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: Dependent variable is the intergenerational association in cumulative fertility at age 25. Sample 259 county-cohort cells. Brackets contain standard errors clustered at the country level.

Table 3: Composition Effects over the Fertility Transition

	(1)	(2)	(3)	(4)
Deweighted average children ever born	-0.010	-0.021		
	[0.006]	[0.011]		
Bins (reference category < 1⅓)				
1⅓-1⅔			-0.024	-0.030
			[0.012]	[0.010]
1⅔-2			-0.033	-0.045
			[0.014]	[0.014]
2-2⅓			-0.031	-0.037
			[0.014]	[0.016]
2⅓-2⅔			-0.028	-0.035
			[0.013]	[0.017]
2⅔-3			-0.035	-0.047
			[0.014]	[0.019]
Overall R-squared	0.04	0.45	0.12	0.55
Within R-squared		0.04		0.21
Country & Cohort FE	No	Yes	No	Yes

Note: Dependent variable is the composition effect on the average number of children ever born. Fertility is measured at age 25. Sample includes 259 county-cohort cells. Brackets contain standard errors clustered at the country level.

Table 4: Results for Total Fertility Rates

	(1)	(2)	(3)	(4)
Deweighted total fertility rate	-0.008	-0.011		
	[0.002]	[0.019]		
Bins (reference category < 3)				
3-4			-0.013	-0.065
			[0.008]	[0.033]
4-5			-0.036	-0.064
			[0.008]	[0.043]
5-6			-0.029	-0.066
			[0.007]	[0.046]
6-7			-0.041	-0.077
			[0.007]	[0.063]
7-8			-0.039	-0.088
			[0.008]	[0.076]
Overall R-squared	0.18	0.81	0.28	0.86
Within R-squared		0.04		0.07
Country & Cohort FE	No	Yes	No	Yes

Note: Dependent variable is the composition effect on the total fertility rate. Sample includes 81 countries. Brackets contain standard errors clustered at the country level.

Appendix Table 1: Demographic and Health Surveys Included in the Sample

Benin: 1996, 2006
Bolivia: 1994, 2003, 2008
Burkina Faso: 1999
Burundi: 2010
Cambodia: 2000, 2005, 2010
Cameroon: 1998, 2004
Central African Republic: 1995
Chad: 1996, 2004
Congo, Democratic Republic: 2007
Congo, Republic: 2005
Cote d'Ivore: 1994
Dominican Republic: 2002, 2007
Ethiopia: 2000, 2005, 2010
Gabon: 2000
Guinea: 2000, 2005
Haiti: 2000, 2005
Indonesia: 2012
Kenya: 2003, 2008
Lesotho: 2004, 2009
Madagascar: 1992, 1997, 2004, 2008
Malawi: 1992, 2000, 2004, 2010
Mali: 1995, 2001, 2006
Morocco: 1992, 2003
Mozambique: 1997, 2003
Namibia: 1992, 2000
Nepal: 1996, 2006
Nigeria: 2008
Peru: 1992, 1996, 2000, 2004
Philippines: 1993, 1998
Rwanda: 2000, 2005, 2010
São Tomé & Príncipe: 2008
Senegal: 1992, 2005
Sierra Leone: 2008
South Africa: 1998
Sudan: 2010
Swaziland: 2007
Tanzania: 1996, 2004, 2010
Togo: 1998
Zambia: 1996, 2001, 2007
Zimbabwe: 1994, 1999