Intergenerational Dynamics and the Fertility Transition

Tom S. Vogl*

Princeton University, BREAD, and NBER

April 2015

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 the intergenerational transmission of fertility—the association of a mother's fertility with her daughter's fertility—evolves over the fertility transition, as well as how it feeds back into the time series of aggregate fertility rates. Microdata from 40 developing countries over the second half of the 20^{th} century show that intergenerational transmission strengthens during fertility decline, in large part because socioeconomic differentials in fertility flip during the transition from Malthusian to modern fertility regimes. As a result, intergenerational transmission raises aggregate fertility rates as populations approach low fertility, pushing back against aggregate fertility decline.

*E-mail: tvogl@princeton.edu. I am grateful for comments from seminar participants at Bocconi University, Lund University, Northeastern University, Stockholm University, Princeton University, UC Berkeley, USC, and the AEA/ASSA meetings.

1 Introduction

Associations in the traits, behaviors, and outcomes of parents and children have attracted interest across the social sciences, both for their relevance to social mobility and for their implications regarding the dynamics of aggregate social and economic phenomena.[1] Some recent work aims to disentangle the mechanisms underlying these associations, including the causal effect of parents' on children's outcomes (Currie and Moretti 2003; Black et al. 2005) and the role of pre-birth factors like genes (Sacerdote 2007; Björklund et al. 2010). But whatever their drivers, the associations are in themselves important determinants of processes that span generations. Indeed, much of the recent literature on long-run economic growth and the demographic transition emphasizes intergenerational linkages within families, especially as they relate to fertility, child investment, and bequests (Galor 2011). Continuing this focus on intrafamily associations and aggregate change, this paper studies one of the earliest intergenerational associations to draw researchers' attention, dating as far back as Pearson et al. (1899): that involving the fertility of mothers and daughters.

Intergenerational associations in fertility are theoretically distinct from those in other behaviors and outcomes because fertility heterogeneity affects the composition of the population: lineages with higher fertility in one generation comprise a larger share of the next generation. This insight may help shed light on aggregate fertility dynamics during the demographic transition, yet it has received surprisingly little attention in the literature. Two recent papers in biodemography (Murphy and Wang 2003; Kolk et al. 2014) do investigate the issue, but only in the context of highly stylized micro-simulation models. Though they advance our understanding of how the intergenerational transmission of fertility aggregates up to the population, these simulations provide neither general theoretical results nor an immediate link to the data. Similarly, in the century since Pearson et al.'s pioneering work, a large literature in demography has developed around the estimation of intergenerational associations in fertility, mostly in Western populations.[2] But little of this work draws clear-cut links to aggregate change, theoretically or empirically.

In economics, while both intergenerational transmission and differential fertility have received

[1] Income and educational attainment have received special attention (Solon 1999; Black and Devereux 2011), but the literature also considers a wide range of abilities, attitudes, and behaviors: intelligence (Black et al. 2009), risk preferences (Dohmen et al. 2012), occupational choice (Hellerstein and Morrill 2011), labor supply (Couch and Dunn 1997; Altonji and Dunn 2000), fertility (Anderton et al. 1987), and divorce (Wolfinger 1999), among others.

[2] 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).

much attention, rarely has this attention covered the intergenerational transmission *of* fertility. Classic theoretical studies of intergenerational transmission in other behaviors or outcomes tend to assume fixed fertility, thus abstracting from how differential fertility (in the cross-section) interacts with intergenerational transmission to shape the distribution of behaviors or outcomes in the next generation (Becker and Tomes 1979; Goldberger 1989; Bowles and Gintis 2002).[3] In recent research seeking to model the joint evolution of economic and demographic forces (Galor and Moav 2002; de la Croix and Doepke 2003; Moav 2005), the interplay between differential fertility and intergenerational transmission has received more attention, but this literature does not explicitly consider intergenerational transmission in demographic behavior. A separate, primarily empirical literature on how differential fertility affects the distribution of outcomes in the next generation (Lam 1986; Preston and Campbell 1993; Mare 1997; Mare and Maralani 2006; Vogl 2014) is similarly silent on fertility transmission. In both of these recent literatures, major questions revolve around the role of differential fertility in the emergence of modern economic growth and the demographic transition. Drawing on data on more than half a million mother-daughter pairs from developing countries, this paper studies how the intergenerational transmission of fertility changes over the demographic transition, as well as how it feeds back into the evolution of aggregate fertility rates.

Much of the analysis focuses on documenting and understanding the evolution of mother-daughter fertility associations in developing countries over the past half-century. According to the conventional interpretation of the literature, intergenerational associations in fertility are positive but small. However, a recent meta-analysis (Murphy 1999) shows that the predictive power of parental fertility in explaining offspring fertility has grown in recent decades, now approaching that of other commonly-studied determinants of fertility. This finding may suggest that economic development or demographic transition strengthens the intergenerational transmission of fertility. Indeed, in a separate analysis of developing countries, Murphy (2012) finds much lower intergenerational correlations than one currently finds in industrialized countries, although among the countries he studies, intergenerational transmission is if anything highest in the least developed (as measured by the Human Development Index). This paper extends Murphy's sample to cover a greater number of birth cohorts, allowing for a fuller understanding of variation in intergenerational transmission, both across countries and within them over time.

[3]The same is true of the classic sociological texts on the topic (McGinnis 1968; McFarland 1970; and Spilerman 1972).

Because the paper defines intergenerational transmission broadly—incorporating any source of persistence in relative fertility across generations—its variation across space and time is likely to be influenced by the transformation of socioeconomic differentials in fertility in developing countries over the past half century. Building on a large literature that mainly consists of single-country studies, Vogl (2014) analyzes retrospective data from 48 developing countries, documenting how fertility gaps between the rich or highly-educated and their worse-off counterparts have flipped. In the past, when Malthusian population dynamics prevailed, richer or more educated parents had higher fertility than their less fortunate counterparts; largely as a result, children from big families obtained more education than those from small. Over the course of the demographic transition, these associations flipped from positive to negative, most likely due to an increase in the return to investing in children. The history of differential fertility therefore spans two regimes, one Malthusian and one modern. In the Malthusian regime, better-off parents bear *more* children, and their children obtain more education, which in turns promotes *higher* fertility. In the modern regime, better-off parents bear *fewer* children, and their children obtain more education, which in turns promotes *lower* fertility. Thus, within either stable regime, differential fertility strengthens the intergenerational transmission of fertility. But during the transition between the regimes, the intergenerational transmission of fertility is muted and may even turn negative. Better-off parents bear more children, and their children obtain more education, which promotes lower fertility. The intergenerational transmission of fertility follows a u-shape over the fertility transition.

To empirically document these patterns, I draw on the Demographic and Health Surveys (DHS), combining sibling history data (in which respondents list their mothers' children ever born) with fertility history data (in which respondents list their own children ever born) to form a mother-daughter dataset spanning 40 developing countries. A limitation of the DHS is that it interviews women of childbearing age, most of whom have not completed childbearing. However, I show that the association between a mother's completed fertility and her daughter's cumulative fertility at age 25 strongly predicts the association between both generation's completed fertilities. Building on this result, I assemble a panel dataset by estimating, for each country and 5-year birth cohort (from 1945-9 to 1980-4), the association between a mother's completed fertility and her daughter's cumulative fertility at age 25.

This dataset brings to view clear evidence that the intergenerational transmission of fertility

strengthened significantly during the transition to the modern regime, although it contains only limited signs of the Malthusian regime. In their childbearing decisions, women born in the 1980s were more strongly linked to the mothers than were their counterparts born in any of the previous three decades. The strengthening is strongly linked with changes in the correlates of growing up in a large family; country-cohort cells with more negative associations between sibship size and education have more positive intergenerational fertility associations. Education negatively predicts fertility for all of these cohorts, so when the sibsize-education association flipped from positive to negative, the link between a daughter's fertility and her mother's fertility increased significantly. These results, which hold for counts of both children ever born and surviving children, suggest that intergenerational transmission may be evolving in developing countries in a way similar to its historical evolution in currently industrialized countries.

Although these findings by themselves contribute to the literature on intergenerational associations in reproductive behavior, their potential implications for aggregate fertility rates make them especially interesting. To shed light on these implications, I non-parametrically estimate the extent to which average fertility differs from the counterfactual average that would arise if all mothers had the same number of daughters, but their daughters bore children as if they were born into their actual sibships (of heterogeneous size). This gap, which I term the *composition effect*, captures the property that distinguishes fertility transmission from other intergenerational transmission: the endogenous change in the next generation's composition.[4] Estimating this effect at the country-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 this finding to a more common measure of fertility, I also estimate composition effects on total fertility rates (TFR): the expected number of children a woman would expect to bear if she experienced current age-specific fertility rates throughout her lifetime. Here too, the composition effect becomes significantly stronger as the TFR declines. In populations with TFRs less than 3, differential fertility raises the 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

[4]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.

dynamics. In this sense, they relate to the literature on how micro-level demographic phenomena like differential fertility and assortative mating aggregate up to the population or the economy (Kremer 1993; Lam 1997; 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 well-known puzzles in aggregate fertility dynamics. In particular, Bongaarts (2006, 2008) provides evidence that following a period of sustained fertility decline, many populations "stall" at fertility levels well above the replacement level. One explanation for such "stalls" is that intergenerational transmission and population composition interact to slow (and even reverse) the decline of average fertility.[5]

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 overrepresented 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. Both lines of research highlight the value of considering how the cross-section interacts with the aggregate time series of important behaviors and outcomes.

Finally, the results contribute to a recent economic literature that takes interest in how intergenerational transmission varies across space and over time. Much of this literature focuses on intergenerational earnings mobility and its variation both across countries (Björklund and Jantti 1997; Solon 1999; Corak 2013) and across sub-national areas within a country (Chetty et al. 2014). International comparisons are also available for the transmission of other outcomes, including educational attainment (Hertz et al. 2007; Chevalier et al. (2009) and health (Bhalotra and Rawlings 2013). This paper provides a wider-ranging analysis of the intergenerational transmission of fertility than has previously been available across any group of countries, as well as a wider-ranging analysis of transmission in any outcome than has previously been available across developing countries.

[5]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.

2 Intergenerational Transmission and Aggregate Fertility Dynamics

To guide the empirical work, this section provides a conceptual framework for understanding intergenerational transmission in fertility. First, it describes how the dynamics of average fertility depend on the relationship between childhood family size and fertility in adulthood. Second, it embeds this relationship in an economic model of differential fertility, asking how it may change over the course of the fertility transition.

2.1 Aggregating Lineages

To clarify the link between intergenerational transmission and the time series of average fertility, first consider the dynamics of a population with one sex. The population consists of a set of lineages, whose size may vary from generation to generation. Let $N_{ijt} \geq 0$ be the net fertility (the number of offspring surviving to reproductive age) of member i from lineage j in generation t. To capture the relationship between fertility in generations $t-1$ and t, define the conditional expectation function $n_t\left(N_{ij,t-1}\right) \equiv E\left[N_{ijt}|N_{ij,t-1}\right]$.

Given this setup, one can express average fertility in generation t as a function of the distribution of fertility in generation $t-1$:

$$E\left[N_{ijt}\right] = E\left[\left(\frac{N_{ij,t-1}}{E\left[N_{ij,t-1}\right]}\right) n_t\left(N_{ij,t-1}\right)\right] \tag{1}$$

where $N_{ij,t-1}$ is the net fertility of individual ijt's parent or, equivalently, individual ijt's surviving sibship size. The distribution of $n_t\left(N_{ij,t-1}\right)$ is reweighted by the factor $N_{ij,t-1}/E\left[N_{ij,t-1}\right]$ to reflect the changing composition of the population from $t-1$ to t. This reweighting is the main theoretical feature that distinguishes the intergenerational transmission of fertility from the intergenerational transmission of other behaviors and traits. If $n_t\left(\cdot\right)$ is an increasing function, then the added weight given to high fertility lineages raises average fertility. Equation (1) has the counterintuitive implication that a reduction of fertility levels in one generation can cause fertility to trend upward in future generations, as the population adjusts to a new steady state family size distribution.

With data linking the fertility behavior of two generations, one can quantify how intergenerational transmission interacts with population composition to affect the evolution of average fertility. In particular, one can compare actual average fertility with a counterfactual average that would ob-

tain in the absence of reweighting. As in Vogl (2014), I refer to this comparison as the *composition effect* of differential fertility on average fertility in the next generation. It reflects the mechanical effect of fertility heterogeneity in generation $t-1$ on the composition of the population in generation t. For the one-sex population, the composition effect on generation t's average fertility is:

$$\Delta_t = E\left[\frac{N_{ij,t-1}}{E\left[N_{ij,t-1}\right]} n_t\left(N_{ij,t-1}\right)\right] - E\left[n_t\left(N_{ij,t-1}\right)\right] \tag{2}$$

The first term is actual average fertility as given by Equation (1), whereas the second term corresponds to a counterfactual world in which members of generation t were all born into families of the same size but bore children as they do in reality. When Δ_t is positive, fertility heterogeneity in generation $t-1$ raises average fertility in generation t.

Like all model-free decomposition methods (Blinder 1973; Oaxaca 1973; DiNardo et al. 1996), the composition effect of differential fertility does not account for parents' endogenous responses to changes in relative fertility or for general equilibrium effects. These excluded channels may alter the shape of the intergenerational transmission function $n_t\left(N_{ij,t-1}\right)$. For example, a reduction in the fertility of poor parents might lead them to invest more in their children's human capital, with eventual effects on their children's fertility.

Both for the empirical work and for insight into the properties of the composition effect, an important case involves a linear intergenerational transmission function: $n_t\left(N_{ij,t-1}\right) = \alpha_t + \beta_t N_{ij,t-1}$. The coefficient β_t is the *intergenerational transmission coefficient.* In this case:

$$\Delta_t = \beta_t\left(\frac{V\left[N_{ij,t-1}\right]}{E\left[N_{ij,t-1}\right]}\right) \tag{3}$$

Δ_t increases in β_t and $V\left[N_{ij,t-1}\right]$, while decreasing in $E\left[N_{ij,t-1}\right]$. In other words, the composition effect of differential fertility is more positive when intergenerational transmission is stronger or when the variance of childhood family size is larger relative to its mean. The role of the ratio $V\left[N_{ij,t-1}\right]/E\left[N_{ij,t-1}\right]$ in linking the cross-section with the aggregate 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.

2.2 Intergenerational Transmission and Demographic Transition

In Equation (1), two quantities determine average fertility: (1) the reweighting factor $\frac{N_{ij,t-1}}{E\left[N_{ij,t-1}\right]}$ and (2) the intergenerational transmission function $n_t\left(N_{ij,t-1}\right)$. What are the properties of $n_t\left(N_{ij,t-1}\right)$, and how do they vary over the demographic transition? Section 2.2.1 first studies these questions in an overlapping generations model of differential fertility, building on previous work in the Unified Growth Theory (Galor 2011) tradition. The model generates several hypotheses regarding the evolution of $n_t\left(N_{ij,t-1}\right)$ over the demographic transition. With an eye toward taking these hypotheses to the data, Section 2.2.2 then generalizes them to allow for other mechanisms of intergenerational transmission.

2.2.1 Intergenerational Transmission in a Model of Differential Fertility

Parents maximize a log-linear utility function over their own consumption (c_t), the number of children (n_t), and human capital per child (h_{t+1}):

$$U\left(c_t, n_t, h_{t+1}\right) = \alpha \log\left(c_t\right) + \left(1-\alpha\right)\left(\log\left(n_t h_{t+1}\right)\right) \tag{4}$$

$\alpha \in (0,1)$ indexes the weight the parents place on their own consumption relative to the combined quantity and quality of children. Child quality, or human capital, is produced from education spending e_t by means of an increasing, concave, and twice-differentiable production function:

$$h_{t+1} = h(e_t) \tag{5}$$

with the additional properties that $h\left(0\right) > 0$ and $h'(0) < \infty$ These additional assumptions imply that even if fertility declines with parental human capital in the interior solution, there will be a corner solution in which fertility rises with parental human capital..

Apart from education spending, each child costs $\tau \in (0,1)$ units of time and $\kappa > 0$ goods. Parents are endowed with human capital h_t, which is drawn from a nondegenerate distribution and receives wage rate w. These assumptions result in the following budget constraint:

$$c_t + \kappa n_t + e_t n_t \leq w_t h_t\left(1-\tau n_t\right) \tag{6}$$

Parents can spend their full income wh_t on their own consumption (c_t), on the quantity costs of children (κn_t and $\tau w h_t n_t$), and on the quality of children ($e_t n_t$).

On can gain much insight into the model's predictions regarding differential fertility by examining the first order condition for fertility, which sets:

$$n_t = (1-\alpha)\left(\frac{h_t}{\kappa + \tau h_t + e_t}\right)$$

If parents choose $e_t = 0$, then n_t increases in h_t. Thus, for parents at a child investment corner solution, fertility increases with parental human capital. If parents choose positive e_t, fertility can either increase or decrease with parental human capital, with the sign of the comparative static depending on how rapidly e_t rises with h_t. In particular, n_t decreases with h_t only when e_t rises rapidly with h_t. As Jones et al. (2011) point out, the elasticity of human capital with respect to child investment $\eta(e_t) \equiv \frac{e_t h'(e_t)}{h(e_t)}$ plays a key role in determining how rapidly e_t rises with h_t.

Proposition 1. *If $\eta'(\cdot) > 0$ and κ is sufficiently small, then there exists a threshold $\tilde{h}$ such that n_t increases in h_t for all $h_t < \tilde{h}$ and decreases in h_t for all $h_t > \tilde{h}$.*

The Theory Appendix contains all proofs. Under these assumptions for the elasticity and the goods cost, the model thus predicts a hump-shaped relationship between fertility and parental skill. Below $\tilde{h}$, parents choose a child investment corner solution, so the income effect of higher skill dominates the substation effect, and fertility rises with skill. Under the stated assumptions, the substation effect dominates above $\tilde{h}$, so fertility declines with skill. When the stated assumptions do not hold, the declining portion of the hump-shape will disappear in most circumstances, so that fertility everywhere increases in parental skill. Using data from 48 developing countries, Vogl (2014) documents a hump-shaped skill-fertility profile that leads the least-skilled to exhibit lower fertility in the early stages of development and higher fertility in the later stages of development. I refer to the earlier regime as "Malthusian" and the later regime as "modern."

One can gain much insight into the model's implications for intergenerational transmission by studying its steady states—with constant levels of fertility and human capital across generations—and the transitions between them. To explore this issue, one must first establish the possibility of multiple steady state levels of human capital, which generate the heterogeneity that is crucial to the model.

Proposition 2. *For any τ and κ, there exist human capital production functions $h(\cdot)$ with $\eta'(\cdot) > 0$ that generate 1, 2, and 3 steady state levels of human capital.*

This result has much in common with Proposition 1 of Moav (2005), who considers a human capital production function with slightly different properties. The model has up to three steady states, one with zero child investment and two with positive child investment. 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 the other two also exist) is unstable. For the remainder of the section, I will focus on the two stable steady states of an environment with three steady states.

In the model, two forces are particularly natural for generating fertility transitions: rising wages and rising returns to child investment. Rising wages increase the fertility of low-skill parents, while decreasing the fertility of high-skill parents. Rising returns decrease the fertility of high-skill parents, while leaving the low-skill steady state unchanged. Consider a population that experiences a one-time, permanent increase in either of these parameters, such that two stable steady states exist both before and after the increase. Then the following results hold for the intergenerational transmission coefficient, β.

Proposition 3. *In any multigenerational steady state distribution of human capital, $\beta = 1$. During a transition between two Malthusian steady state distributions, $\beta \in (0, 1)$. During a transition between two modern steady state distributions, $\beta > 1$. During a transition between Malthusian and modern steady state distributions, $\beta < 0$ for at least one generation.*

The first result is a basic property of a steady state: fertility is constant within lineages. In contrast, the second and third results come from considering the dynamics of the population as optimal fertility moves along the hump shape. When the population switches from a Malthusian regime to a modern regime, relative fertility flips. Low-skill parents bear more children than high-skill parents, despite having grown up in larger childhood families. As a result, the intergenerational transmission coefficient follows a u-shape over the fertility transition.

2.2.2 Other Mechanisms

The model provides one source of intergenerational transmission in fertility, due to linkages across generations in the socioeconomic determinants of fertility. Although the literature on demography

and long-run growth (Galor 2011) suggests that this source is likely to be important during the demographic transition, other sources are possible. Chief among them are the genetic heritability of fecundity and cultural influences on fertility preferences and behavior, and linkages across generations in the socioeconomic determinants of fertility. To combine the model's predictions with these alternative sources of intergenerational transmission, one might more generally express the intergenerational transmission of fertility as:

$$\frac{\partial E\left[N_{ijt}|N_{ij,t-1}\right]}{\partial N_{ij,t-1}} = \frac{\partial E\left[N_{ijt}|N_{ij,t-1}, H_{ijt}\right]}{\partial N_{ij,t-1}} + \frac{\partial E\left[N_{ijt}|N_{ij,t-1}, H_{ijt}\right]}{\partial H_{ij,t}} \times \frac{\partial E\left[H_{ijt}|N_{ij,t-1}\right]}{\partial N_{ij,t-1}} \tag{7}$$

where H_{ijt} is the human capital of individual i from lineage j in generation t. Note that if the conditional expectation is linear, $\beta = \frac{\partial E[N_t|N_{t-1}]}{\partial N_{t-1}}$. This equation merely expands the overall intergenerational transmission of fertility, $\frac{\partial E[N_t|N_{t-1}]}{\partial N_{t-1}}$, into sub-components using the chain rule. These subcomponents include a 'pure' intergeneration transmission term, $\frac{\partial E[N_t|N_{t-1},H_t]}{\partial N_{t-1}}$, the relationship between skill and fertility, $\frac{\partial E\left[N_{ijt}|N_{ij,t-1},H_{ijt}\right]}{\partial H_{ij,t}}$, and the relationship between childhood family size and skill, $\frac{\partial E\left[H_{ijt}|N_{ij,t-1}\right]}{\partial N_{ij,t-1}}$. The model in Section 2.1.1 focused on the second and third components, which change in a staggered fashion during the demographic transition. In both the Malthusian, both components are positive; in the modern regime, both components are negative; in the transition, they are of opposite sign. Thus, their product follows a u-shape, as predicted by the model.

However, the model does not capture the 'pure' transmission component, which reflects reflect genetic and cultural transmission.[6] The behavioral genetics literature documents some genetic heritability in various 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 heritability in 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 theory for the transmission of fertility, positing that growing up in a large family shapes preferences for a large family of one's own. In their economic model of cultural transmission across generations, Bisin and Verdier (2000, 2010) distinguish between "direct vertical socialization," in which parents consciously shape their children's preferences and behavior, and "oblique horizontal socialization," in which children learn from or imitate their surroundings. Both

[6]Galor and Moav (2002) also posit a model with genetic or cultural transmission in the *taste* for child quantity or quality, which has similar properties to the model in Section 2.1.1 and can thus be characterized by the product discussed in the previous paragraph.

sources of cultural transmission could lead to the intergenerational transmission of fertility. Children may behave like their parents because of parental indoctrination or because of the preference-shaping effects of their family and community experiences in childhood. Alternatively, families and communities may reinforce fertility behavior across generations through social norms—as emphasized by Munshi and Myaux (2006)—or through social learning—as emphasized by researchers Coale and Watkins (1986) and Spolaore and Wacziarg (2014).

If the 'pure' transmission component $\frac{\partial E[N_t|N_{t-1},H_t]}{\partial N_{t-1}}$ is small or varies little over time, then one would still expect β to follow the trajectory predicted by the model in the previous section. In the Malthusian regime, higher-skill parents bear more children, so $\frac{\partial E[N_t|N_{t-1},H_t]}{\partial H_t} > 0$, and women from larger sibships obtain more education, so $\frac{\partial E[H_t|N_{t-1}]}{\partial N_{t-1}} > 0$. In the modern regime, both patterns are flipped, so $\frac{\partial E[N_t|N_{t-1},H_t]}{\partial H_t} < 0$ and $\frac{\partial E[H_t|N_{t-1}]}{\partial N_{t-1}} < 0$. In the transition between these two regimes, a single generation experiences long and auxiliary parameters of opposite sign: $\frac{\partial E[N_t|N_{t-1},H_t]}{\partial H_t} < 0$ and $\frac{\partial E[H_t|N_{t-1}]}{\partial N_{t-1}} > 0$. Thus, if the 'pure' transmission component is small or varies little over time, then $\frac{\partial E[N_t|N_{t-1}]}{\partial N_{t-1}}$ (in the linear model, β) follows a u-shape over the transition. I will argue that in the postwar era, developing countries experienced the upward-sloping portion of this process, moving from the transition to the modern regime.

3 Data on the Fertility Behavior 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.[7] First, they must include information on the survival of both offspring and siblings,

[7]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.

to allow estimation of intergenerational transmission coefficients for both ever-born and surviving fertility. Second, they must sample all women, not only ever-married women.[8] As of November 2013, 80 surveys in 40 African, Asian, Caribbean, Latin American, and Pacific countries satisfied these criteria (see Appendix Table 1 for a list of countries and survey years). Together, these surveys provide data on sibship size and fertility for 966,498 women aged 15-49.

Most of the analyses below group these women into cells defined by country and 5-year birth cohort. To maintain precision in country-cohort parameter estimates, I drop country-cohort cells with fewer than 100 observations. Some analyses (described below) restrict further to cells with at least 250 observations. The analyses use sampling weights provided by the DHS, but because I combine multiple surveys per country and then estimate parameters at the country-cohort level, I rescale the survey weights to maximize efficiency. Specifically, suppose 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.

4 Intergenerational Transmission of Fertility: Cohort Variations

This section analyzes the intergenerational transmission of fertility, a key determinant of the composition effect of differential fertility (the subject of Section 5 below). As a first step, it estimates the parameters of the short, long, and auxiliary models by country and birth cohort, providing summary statistics on their distributions. Using these coefficients, it then shows how changes in cross-sectional fertility patterns affect the association of mothers' and daughters' fertilities.

4.1 Method

I estimate intergenerational transmission coefficients for cells defined by country and 5-year birth cohort. In this estimation exercise, two issues of measurement arise. First, should one count all offspring and siblings ever born, or only those who survived to the interview? If one were purely interested in the intergenerational transmission of fertility *behavior* (the number of pregnancies), then ever-born (E) fertility would be more relevant. If one were instead interested in the intergen-

[8] 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.

erational transmission of fertility *outcomes* (the effective number of offspring), then surviving (S) fertility would be more relevant. But for those interested in the intergenerational transmission of 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. 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. Because both measures are of interest, I estimate intergenerational transmission in ever-born and surviving fertility.

A second methodological question relates to the age of fertility measurement. The DHS interviews most respondents midway through their childbearing years, so the analysis will involve a regression of the respondent's fertility at some age $a < 50$ on her sibship size, or her mother's completed fertility. The choice of a involves a tradeoff. An earlier choice allows for a larger sample and coverage of later birth cohorts, while a later choice covers more of the reproductive lifespan. As an antecedent to the rest of the analysis, Section 4.2 studies the choice of a.

Given the choice of a, I estimate intergenerational transmission of fertility as follows:

$$fertility_{ict}^{ja} = \alpha_{ct}^{ja} + \beta_{ct}^{ja} sibsize_{ict}^{j} + u_{ict}^{ja} \tag{8}$$

where i indexes the individual, c indexes her country, t indexes her birth cohort, and j indexes the type of fertility (ever-born or surviving). β_{ct}^{ja} is country-cohort ct's intergenerational transmission coefficient for type-j fertility at age a. As specified in Section 2, β_{ct}^{ja} captures *all* mechanisms linking a mother's fertility behavior with her daughter's. Section 4.3 describes the distribution of β_{ct}^{ja} across countries and birth cohorts.

Motivated by Equation (7), one can decompose β_{ct}^{ja} into its driving forces: $\frac{\partial E[N_t|N_{t-1},H_t]}{\partial N_{t-1}}$, $\frac{\partial E[N_t|N_{t-1},H_t]}{\partial H_t}$ and $\frac{\partial E[H_t|N_{t-1}]}{\partial N_{t-1}}$. In a linear system, two equations summarize these three partial derivatives. The first expresses fertility as a function of sibship size and education:

$$fertility_{ict}^{ja} = \mathrm{A}_{ct}^{ja} + \mathrm{B}_{ct}^{ja} sibsize_{ict}^{j} + \Gamma_{ct}^{ja} edu_{ict} + U_{ict}^{ja} \tag{9}$$

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

$$edu_{ict} = \psi^{j}_{ct} + \varphi^{j}_{ct} sibsize^{j}_{ict} + v^{j}_{ict} \tag{10}$$

In the language Goldberger (1991) uses to characterize omitted variables bias, Equation (8) is a "short" regression, while Equations (9) and (10) are the corresponding "long" and "auxiliary" regressions. By Equation (7), the slope parameters are related by the identity $\beta^{ja}_{ct} = \mathrm{B}^{ja}_{ct} + \varphi^{j}_{ct}\Gamma^{ja}_{ct}$. This correspondence will guide Section 4.4, which seeks to explain heterogeneity in β^{ja}_{ct}.

Both for choosing a and for analyzing the relationships between Equations (9)-(10) and Equation (8), an issue is that the parameter estimates, which will be used as regressors and regressands, approximate the true parameters with error. As a result, OLS estimates of the linkages among $\hat{\beta}^{ja}_{ct}$, $\hat{\mathrm{B}}^{ja}_{ct}$, $\hat{\Gamma}^{ja}_{ct}$, and $\hat{\varphi}^{j}_{ct}$ will be biased. In earlier work based on the same dataset (Vogl 2014), I confront the same issue with similar cell sizes and find that Fuller's (1987) error-correction model produces estimates very close to those produced by OLS. To make the results as transparent as possible, I report only the OLS estimates here. To give a sense of the extent of bias, Appendix Tables 2-3 report the results of analyses that double the minimum cell size. Estimates from these larger cells have lower error variances, so if estimation error substantially biases the main results, then the results in Tables 2-3 should differ from the main results. In practice, they differ little, suggesting that errors-in-variables bias is not a major concern.

4.2 Choosing the Age of Fertility Measurement

Ideally, one would regress a daughter's completed fertility on her mother's completed fertility. Unfortunately, the DHS interviews most respondents before they have completed their fertility, so that such a regression would leave out much of the sample. My solution to this problem is to choose the earliest age a such that β^{ja}_{ct} is representative of the intergenerational transmission coefficient for completed fertility. To find this age, I study how transmission coefficients at later ages relate to transmission coefficients at earlier ages. If those at earlier ages can serve as proxies for those at later ages, then one can take advantage of their greater coverage of birth cohorts.

Questions surrounding the age of measurement in the second generation also arise in the estimation of intergenerational income elasticities (Solon and Haider 2006), so the literature on that

topic provides some precedent on the issue. 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). In their research on intergenerational income mobility in the United States, Chetty et al. (forthcoming) address this issue by using the age at which the intergenerational association becomes stable for the rest of the lifecycle. This approach works well for estimating intergenerational associations in income, which appear to level off when the child reaches age 30.

If fertility gaps are concentrated early in the reproductive lifespan, then one can take a similar approach to analyzing intergenerational associations in fertility.[9] Along these lines, I choose the earliest age a for which β_{ct}^{ja} approximates the completed fertility transmission coefficient. I draw on the full fertility histories of women at least 45 years old to estimate intergenerational transmission coefficients at ages 20, 25, 30, 35, 40, and 45. Taking fertility at age 45 to be "complete," I then regress the age-45 transmission coefficient on transmission coefficients at earlier ages:

$$\hat{\beta}_{ct}^{j45} = \theta \hat{\beta}_{ct}^{ja} + \delta_c + \tau_t + \varepsilon_{ct} \tag{11}$$

for ages $a = 20$, 25, 30, 35, and 40. The parameter of interest is θ, the association between transmission coefficients at different ages. If θ is close to 1, then one can on average interpret $\hat{\beta}_{ct}^{ja}$ in the same way as $\hat{\beta}_{ct}^{j45}$. The specification has country (δ_c) and birth cohort (τ_t) fixed effects, but for completeness, I also report estimations without them. Standard errors are clustered by country.

The results, which appear in Table 1, indicate that transmission coefficients at earlier ages are strongly informative about the intergenerational transmission coefficient for completed fertility (as measured at age 45). None of the estimates of θ are significantly different from 1, although precision is low at younger ages of measurement a. When $a = 20$, so that θ measures the relationship between a cohort's age-20 transmission coefficient and its age-45 transmission coefficient, the estimates of θ are unstable across specifications. But for all $a \geq 25$, θ is stable at values close to 1, both for ever-born and for surviving fertility, both with and without country and cohort fixed effects. These

[9] More precisely, the approach is appropriate if women attain lower fertility be delaying the first pregnancy, rather than by lowering the rate of childbearing following the first pregnancy.

conclusions do not change when I double the minimum cell size (see Appendix Table 2).[10]

These estimations suggest that one can use age-25 transmission coefficients to track the evolution of intergenerational correlations over the fertility transition. This approach covers younger birth cohorts (which have not yet completed childbearing) while still allowing one to interpret the coefficients (on average) as the intergenerational transmission of completed fertility. In the remainder of this section, I focus on age-25 transmission coefficients. Because the sample includes multiple surveys for many countries, data on age-25 fertility are available for a larger number of women per cohort than data on age-45 fertility. This feature of the data allows me to set a higher minimum cell size. For all age-25 estimates, I use a minimum cell size of 250, which excludes 3.6 percent of the country-cohort cells that include women over 25.

4.3 Intergenerational Transmission Coefficients

Figure 2 describes the distribution of $\hat{\beta}_{ct}^{j25}$ and $\hat{\mathrm{B}}_{ct}^{j25}$ across country-cohort cells, showing considerable dispersion in the parameter estimates. In separate panels for ever-born and surviving fertility, the figure draws a kernel density estimate for each coefficient, using a bandwidth of 0.01. The darker curves correspond to $\hat{\beta}_{ct}^{j25}$ from the "short" regression, while the lighter curves correspond to $\hat{\mathrm{B}}_{ct}^{j25}$ from the "long regression." Two patterns emerge for both ever-born and surviving fertility. First, the central tendencies of $\hat{\beta}_{ct}^{j25}$ and $\hat{\mathrm{B}}_{ct}^{j25}$ are close to zero, and the supports of the distributions include both positive and negative values. Second, the dispersion of $\hat{\mathrm{B}}_{ct}^{j25}$—from the "long" regression, which controls for education—is smaller than that of $\hat{\beta}_{ct}^{j25}$ from the "short" regression. This second finding implies that heterogeneity in the role of education explains part of the observed heterogeneity in intergenerational transmission. Perhaps more interestingly, the distribution of $\hat{\mathrm{B}}_{ct}^{j25}$ has shorter right *and* left tails than the distribution of $\hat{\beta}_{ct}^{j25}$, suggesting that the association of sibship size and education amplifies the intergenerational transmission of fertility both when it is positive and when it is negative.

One interpretation of Figure 2 is that the intergenerational transmission of fertility is 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 interpreta-

[10] In Appendix Table 2, estimations with country and cohort fixed effects have more unstable results because the higher minimum cell size reduces the number of countries with multiple cohorts.

tion. Using a significance level of 5 percent, the ever-born transmission coefficient ($\hat{\beta}_{ct}^{E25}$) is significantly positive in 27 percent of cells and significantly negative in 2 percent of cells; the surviving transmission coefficient ($\hat{\beta}_{ct}^{S25}$) is significantly positive in 15 percent of cells and significantly negative in 9 percent of cells. Notably, the distribution of $\hat{\beta}_{ct}^{S25}$ contains too many positive values *and* too many negative values to be explained by randomness. Further evidence against the "sampling variability" hypothesis can be found in Appendix Figure 1, which plots two time series of coefficient estimates (one each for ever-born and surviving fertility) for each of the 40 countries in the sample. To the extent that the coefficients change across successive cohorts within a country, that change is relatively smooth. Many countries experienced increasing intergenerational transmission over time, but this trend is not universal. Some countries exhibit no trend at all, some show evidence of a declining trend, and some display a moderate u-shape over time. Even more notably, some estimates of the intergenerational transmission coefficient are significantly negative, a result consistent with the transition between the first and send regimes described in Section 2.2.

To further characterize the evolution of the intergenerational transmission coefficients, on average across countries, I estimate the regression:

$$\hat{\beta}_{ct}^{j25} = \delta_c + \tau_t + \varepsilon_{ct} \tag{12}$$

The dependent variable is the estimated intergenerational transmission coefficient, and the independent variables are country and cohort fixed effects. Figure 3 plots the cohort effects (τ_t), estimated relative to the omitted 1945-9 cohort. Transmission coefficients for ever-born fertility trend significantly upward across cohorts, while those for surviving fertility exhibit a moderate u-shape. For ever-born fertility, the transmission coefficients for the 1980-4 cohort are significantly more positive ($p < 0.05$) than those for the 1950-4 cohort; for surviving fertility, those for the 1980-4 cohort are significantly more positive ($p < 0.01$) than those for the 1965-9 cohort.

All analyses until this point have shown results for ever-born and surviving fertility separately. How do these two sets of results relate to each other? Appendix Figure 2 explores this issue by plotting transmission coefficients for surviving fertility ($\hat{\beta}_{ct}^{S25}$) against those for ever-born fertility ($\hat{\beta}_{ct}^{E25}$). The two transmission coefficients are highly correlated (with a correlation coefficient of 0.77), but the coefficients for surviving fertility tend to be less strongly positive than those for ever-born

fertility. This finding suggests that, when the intergenerational transmission of fertility is positive, it is amplified by the intergenerational transmission of child mortality.

4.4 Understanding Heterogeneity in Intergenerational Transmission

The fact that the $\hat{\beta}_{ct}^{j25}$ distribution has more mass in both tails than the $\hat{B}_{ct}^{j25}$ distribution means that the role or education varies across time and space. This variation could stem from two sources: (1) heterogeneity in the association of education with fertility (Γ_{ct}^{j25}) and (2) heterogeneity in the association of sibship size with education (φ_{ct}^{j}). Although both the SES-fertility association and the sibsize-education association flip from positive to negative over the fertility transition, the flips are one generation apart. The SES-fertility association flips in one generation, inducing a flip in the next generation's sibsize-education association. This section finds that the second of these flips is crucial to understanding global variation in intergenerational transmission in the second half of the twentieth century. Education is negatively associated with fertility in most of the country-cohort cells, but mirroring the results of Vogl (2014), the association of sibship size with education takes on a wide range of values, both positive and negative.

Basic features of the distributions of $\hat{\Gamma}_{ct}^{j25}$, and $\hat{\varphi}_{ct}^{j}$ are visible in Table 2. For both ever-born and surviving fertility, more than three-quarters of the education-fertility association estimates are significantly negative. Meanwhile, the ever-born sibsize-education association is significantly positive in 39 percent of cells and significantly negative in 25 percent of cells. Similarly, the surviving sibsize-education association is significantly positive in 51 percent of cells and significantly negative in 15 percent of cells. Hence, most women in the sample came of age in the modern fertility regime, with a negative association between female education and fertility. However, some were born in the Malthusian regime, in which richer or higher skill parents had more children, so the estimated sibsize-education associations ($\hat{\varphi}_{ct}^{j}$) flip from positive and negative, while the estimated education-fertility associations ($\hat{\Gamma}_{ct}^{j25}$) are generally negative.

On average, both parameters become more negative across successive birth cohorts. In addition to plotting cohort effects in $\hat{\beta}_{ct}^{j25}$, Figure 3 also plots cohort effects in the other parameters of the linear system (still based on Equation (12)). Net of country fixed effects, $\hat{\Gamma}_{ct}^{j25}$ and $\hat{\varphi}_{ct}^{j}$ become significantly more negative in later cohorts, both for ever-born and surviving fertility. Trends in $\hat{\Gamma}_{ct}^{j25}$ and $\hat{\varphi}_{ct}^{j}$ work to make the intergenerational transmission coefficient more positive.

Because the education-fertility association is generally negative, the overall intergenerational transmission of fertility should be more positive when the sibsize-education association is more negative. One can test this hypothesis by running a regression of the estimated intergenerational transmission coefficient on the estimated sibsize-education association:

$$\hat{\beta}_{ct}^{j25} = \theta \hat{\varphi}_{ct}^{j} + \delta_c + \tau_t + \varepsilon_{ct} \tag{13}$$

As in Equation (11), the regression specification includes both country (δ_c) and birth cohort (τ_t) fixed effects, but I also report estimations without them for completeness. Because the education-fertility associations are largely negative, θ too should be negative.

Estimates of Equation (13)—reported in Table 3, columns (1)-(2) and (5)-(6)—confirm this prediction. Both for ever-born fertility and for surviving fertility, both with and without country and cohort fixed effects, the sibsize-education association bears a significantly negative relationship with the intergenerational transmission coefficient. The estimates imply that a 1-unit decrease in the sibsize-education association leads to a 0.1-unit increase in the intergenerational transmission coefficient. Figure 4 portrays these patterns graphically. The left-hand panels plot the intergenerational transmission coefficient as a function of the sibsize-education association; the right-hand panels plot the the relationship between the residuals after removing country and cohort fixed effects. Scatterplots in all four panels show clearly declining relationships, which are confirmed in the local linear regressions drawn in black. As indicated by the R^2 terms in Table 3, variation in the sibsize-education association explains roughly one-third of the overall variance of the intergenerational transmission coefficient. Perhaps due to a smaller signal-to-noise ratio, *changes* in the sibsize-education association explain a smaller share, about 15 percent, of changes in the intergenerational transmission coefficient. Regardless, variation in the association of sibship size and education appears to play an important role in shaping the intergenerational transmission of fertility.

The prediction of a negative θ depended on the assumption of a negative education-fertility association. Although the data are broadly consistent with this assumption, one can modify Equation (13) to more precisely account for heterogeneity in the education-fertility association:

$$\hat{\beta}_{ct}^{j25} = \theta_1 \hat{\varphi}_{ct}^{j} + \theta_2 \hat{\Gamma}_{ct}^{j25} + \theta_3 \left(\hat{\Gamma}_{ct}^{j25} \times \hat{\varphi}_{ct}^{j} \right) + \delta_c + \tau_t + \varepsilon_{ct} \tag{14}$$

This specification adds the fertility-education association ($\hat{\Gamma}_{ct}^{j25}$) and its interaction with the sibsize-education association ($\hat{\varphi}_{ct}^{j}$) to Equation (13). The omitted variables bias formula predicts that $\theta_3 = 1$. Indeed, estimates of Equation (14)—reported in Table 3, columns (3)-(4) and (7)-(8)—place θ_3 very close to 1, for both types of fertility.

As Section 4.1 pointed out, estimations of Equations (13)-(14) may suffer from bias because the regressand and regressors $\{\hat{\beta}_{ct}^{j25}, \hat{\Gamma}_{ct}^{j25}, \hat{\varphi}_{ct}^{j}\}$ are themselves estimated with error. To explore the importance of this issue, Appendix Table 3 doubles the minimum cell size from 250 to 500, with little change to the parameter estimates. These findings suggest that the initial minimum cell size is large enough to make errors-in-variables bias unimportant.

5 Composition Effects over the Fertility Transition

What are the implications of strengthening intergenerational transmission for the evolution of population fertility rates? This section estimates the composition effect of differential fertility in generation $t-1$ on average fertility in generation t.

5.1 Method

5.1.1 Non-Parametric Estimator of the Composition Effect on Age-25 Fertility

Equation (2) defines the composition effect Δ_t as the levels difference between average fertility and de-weighted average fertility for a one-sex population. For estimation, I modify this definition in two ways. First, to better capture the reality of a two-sex population, I rewrite the effect as a function of the number of female offspring $F_{ij,t-1}$ instead of the number of all offspring $N_{ij,t-1}$:

$$\tilde{\Delta}_t = E\left[\frac{F_{ij,t-1}}{E\left[F_{ij,t-1}\right]}\tilde{n}_t\left(F_{ij,t-1}\right)\right] - E\left[\tilde{n}_t\left(F_{ij,t-1}\right)\right]$$

where $\tilde{n}_t\left(F_{ij,t-1}\right) \equiv E\left[N_{ijt}|F_{ij,t-1}\right]$ is the expected fertility of a woman with $F_{ij,t-1}-1$ sisters. Second, because changing fertility levels complicate comparisons of $\tilde{\Delta}_t$ at different stages of the fertility transition and at different ages, I divide $\tilde{\Delta}_t$ by deweighted average fertility $E\left[\tilde{n}_t\left(F_{ij,t-1}\right)\right]$:

$$\tilde{\delta}_t = \frac{\tilde{\Delta}_t}{E\left[\tilde{n}_t\left(F_{ij,t-1}\right)\right]} \tag{15}$$

$\tilde{\delta}_t$ reflects the *relative* composition effect: the composition effect as a share of baseline fertility. Since it is unit-less, it allows easier comparisons of composition effects at different stages of the fertility transition and at different ages. For linear conditional expectation functions, $\tilde{\delta}_t$ has the same properties as $\tilde{\Delta}_t$: increasing in the slope of $\tilde{n}_t(\cdot)$ and in the ratio of the variance to the mean of $F_{ij,t-1}$.

$\tilde{\delta}_t$ captures three generations: grandmothers (who were adults in $t-1$), mothers (who are adults in t), and children (who will become adults in $t+1$). $F_{ij,t-1}$ is the number of mothers born to a grandmother, while $\tilde{n}_t\left(F_{ij,t-1}\right)$ is the expected number of children born to a mother with $F_{ij,t-1}-1$ sisters. Equation (15) therefore integrates over the distribution of grandmothers. In the DHS, the unit of observation is a mother, so one cannot directly estimate Equation (15). However, applying the law of iterated expectations, one can rewrite $\tilde{\delta}_t$ over the distribution of mothers in generation t:

$$\tilde{\delta}_t = \frac{\sum_{k=1}^{K}\left(\eta_{kt} - \frac{\eta_{kt}/k}{\sum_{l=1}^{K}\eta_{lt}/l}\right)\mu_{kt}}{\sum_{k=1}^{K}\left(\frac{\eta_{kt}/k}{\sum_{l=1}^{K}\eta_{lt}/l}\right)\mu_{kt}} \tag{16}$$

where K is the maximum number of surviving females in a mother's sibship (including herself); η_{kt} is the share of mothers with k surviving females in their sibships; and μ_{kt} is the mean fertility of mothers with k surviving females in their sibships. In theory, the terms "grandmothers" and "mothers" refer to all female members of generations $t-1$ and t, not just those who have children. In practice, however, data only exist for actual grandmothers, as childless members of generation $t-1$ have no offspring in generation t.

Using the empirical analogues of η_{kt} and μ_{kt}, I estimate $\tilde{\delta}_t$ non-parametrically for mean children ever born (CEB) at age 25 among women in birth cohort t. By redefining t as a birth cohort rather than a generation, I slightly abuse notation. As estimated in equation (16), $\tilde{\delta}_t$ is a synthetic cohort measure, corresponding to a hypothetical cohort of families whose size is distributed according to the cross-sectional distribution of sibship size in birth cohort t. $\tilde{\delta}_t$ thus lands somewhere between a period measure and a cohort measure. The true cohort version of the composition effect would follow the fertility behavior of all offspring born to a single cohort of parents; the "cohort" would be the first generation.

5.1.2 Non-Parametric Estimator of the Composition Effect on the Total Fertility Rate

While the structure of the data precludes estimation of the true cohort measure, it does permit estimation of a true period measure: the composition effect of differential fertility on the total fertility rate (TFR). The TFR is the expected number of children a woman would bear if she experienced current age-specific fertility rates throughout her lifetime. To obtain the relative composition effect of differential fertility on this quantity, I estimate 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:

$$\tilde{\delta}_{TFR} = \frac{\sum_{a=15}^{49}\sum_{k=1}^{K}\left(\eta_{ka} - \frac{\eta_{ka}/k}{\sum_{l=1}^{K}\eta_{la}/l}\right)f_{ka}}{\sum_{a=15}^{49}\sum_{k=1}^{K}\left(\frac{\eta_{ka}/k}{\sum_{l=1}^{K}\eta_{la}/l}\right)f_{k\text{ł}}} \tag{17}$$

where f_{ka} is the age-specific fertility rate for women in age group a and sibship size k, and η_{ka} is defined as in equation (16). The term $\frac{\eta_{ka}/k}{\sum_{l=1}^{K}\eta_{la}/l}$ re-weights each age-specific fertility rate to undo the effects of differential fertility. In empirically implementing Equation (17), I use five-year age groups, from 15-19 to 45-49, and an exposure period of one year before the survey. I compute one estimate of $\tilde{\delta}_{TFR}$ per DHS survey, so some countries have multiple estimates.

5.1.3 Decomposition of the Composition Effect in the Linear Case

Recall from Equation (3) that in the case of a linear transmission function with slope $\boldsymbol{\beta}$ the composition effect Δ_t equals $\boldsymbol{\beta}$ times the ratio of the variance to the mean of sibship size. This result provides a useful decomposition of the composition effect into a transmission component and a heterogeneity component. 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. 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 fertility transmission results.

For age-25 fertility, the decomposition of the relative composition effect is exact. For cohort t:

$$\delta_t = \left(\frac{\beta_t}{dCEB_t}\right)\left(\frac{\sigma_t^2}{\mu_t}\right)$$

where $dCEB$ is de-weighted mean children ever born, while σ_t^2 and μ_t are the variance and mean

of sibship size, respectively. Thus, the relative composition effect has two components: a relative transmission coefficient and a ratio. If transmission coefficients become positive during fertility decline, but the ratio stays constant, then the composition effect rises.

For TFR, the decomposition becomes complicated because Equation (17) sums 35 age-specific composition effects. One must approximate the decomposition by either assuming that the relative coefficient or the ratio is constant across age groups. Because the paper focuses on variation in transmission coefficients, I hold the ratio of the variance to the mean of sibship size at its mean:

$$\delta_{TFR} = \frac{\sum_{a=15}^{49} \frac{\beta_a \sigma_a^2}{\mu_a}}{dTFR} \approx \left(\frac{\sum_{a=15}^{49} \beta_a}{dTFR} \right) \left(\frac{1}{35} \sum_{a=15}^{49} \frac{\sigma_a^2}{\mu_a} \right)$$

where β_a is the coefficient from a regression of an indicator for a birth in the year preceding the survey on sibship size for respondents aged a. $dTFR$ is the de-weighted total fertility rate, while σ_a^2 and μ_a are the age-specific variances and means of sibship size, respectively.

5.2 Results

Panel A of Figures 4 and 5 report estimates of $\tilde{\delta}_t$ and $\tilde{\delta}_{TFR}$, each plotted against its corresponding deweighted average fertility measure (deweighted mean age-25 CEB and deweighted TFR, respectively). The x-axis is deweighted to eliminate any reverse causality stemming from the endogenous reshaping of the population. At high levels of fertility, relative composition effects are on average close to zero and invariant to deweighted average fertility. However, both scatterplots reveal an uptick in composition effects at low levels of average fertility, suggesting a link with fertility decline. Composition effects become positive when children ever born at age 25 dips below 1.5, and when the total fertility rate dips below 4. At the lowest observed levels of average fertility for both measures, the predicted composition effect is approximately 0.04, implying that average fertility is 4 percent higher than it would have been in the absence of endogenous changes in the composition of the population. Consistent with previous results, however, Figures 4 and 5 reveal little evidence of a u-shaped pattern over the course of the fertility transition.

Panel B of Figures 4 and 5 shifts attention to the case of a linear fertility transmission function. In both cases, the ratio of the variance to the mean of sibship size stays flat or even decline slightly as average fertility falls. Meanwhile, mirroring the relative composition effect estimates in Panel A, the relative transmission coefficients show a sharp uptick at lower levels of fertility. The rise

of composition effects in the later stages of fertility decline thus appears to be driven largely by increasing intergenerational transmission.

Is this uptick in the composition effect at lower levels of average fertility statistically significant? To shed light on this question, I categorize both measures of average fertility into deciles and then regress the composition effect on indicators for 9 of these 10 bins, with and without country and time fixed effects. The results, appearing in Figure 6, indicate a decline in the composition effect as average fertility increases past the first decile. This conclusion applies to both $\tilde{\delta}_t$ and $\tilde{\delta}_{TFR}$, and the magnitudes are similar with and without fixed effects. Differences from the base category are significant for $\tilde{\delta}_t$ with and without fixed effects, but for $\tilde{\delta}_{TFR}$, significance levels decline in the specification with fixed effects because most countries have just one or two survey years. In both within-country and cross-country variation, declining average fertility is associated with a rising composition effect.

6 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 fertility aggregates. Using data on mother-daughter pairs from 40 developing countries, this paper tracks the evolution of the intergenerational transmission of fertility and explores its aggregate implications, with two main results. First, the intergenerational transmission of fertility grows over the transition to a modern fertility regime, as the sibsize-education association flips from positive to negative, and the education-fertility association becomes more negative. Second, due to this change in intergenerational transmission, 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 burgeoning 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 transmission aggregates 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.[11] In this sense, this paper expands on 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 the intergenerational transmission of fertility, 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 promises to shed much light on the micro-level demographic phenomena influencing the pace of aggregate change.

[11] Exceptions include Murphy and Wang (2003) and Kolk et al. (2014), but these authors use hypothetical models to explore the aggregate consequences of intergenerational transmission. They do not use data to estimate composition effects for real populations.

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.

Altonji, Joseph G., and Thomas A. Dunn. (2000). "An Intergenerational Model of Wages, Hours, and Earnings." *Journal of Human Resources* 35(2): 221-258.

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.

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.

Black, Sandra E., and Paul J. Devereux. (2011). "Recent Developments in Intergenerational Mobility." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics Vol. 4*. Amsterdam: North Holland.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. (2005). "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review* 95(1): 437-449.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. (2009). "Like Father, Like Son? A Note on The Intergenerational Transmission of IQ Scores." *Economics Letters* 105(1): 138-140.

Blinder, Alan S. (1973). "Wage Discrimination, Reduced Form, and Structural Estimates." *Journal of Human Resources* 8(4): 436-455.

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.

Bowles, Samuel, and Herbert Gintis. (2002). "The Inheritance of Inequality." *Journal of Economic Perspectives* 16(3): 3-30.

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), forthcoming.

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.

Coale, Ainsley J., and Susan Cotts Watkins. (1986). *The Decline of Fertility in Europe. The Revised Proceedings of a Conference on the Princeton European Fertility Project*. Princeton: Princeton University Press.

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

Couch, Kenneth A., and Thomas A. Dunn. (1997). "Intergenerational Correlations in Labor Market Status: A Comparison of the United States and Germany." *Journal of Human Resources* 32(1): 210-232.

Currie, Janet and Enrico Moretti. (2003). "Mother's Education and The Intergenerational Transmission of Human Capital: Evidence From College Openings." *Quarterly Journal of Economics* 118(4): 1495-1532.

De la Croix, David and Matthias Doepke. (2003). "Inequality and Growth: Why Differential Fertility Matters." *American Economic Review* 93: 1091-1113.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64(5): 1001-1044.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. (2012). "The Intergenerational Transmission of Risk and Trust Attitudes." *Review of Economic Studies* 79(2): 645-677.

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.

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.

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

Hellerstein, Judith K., and Melinda Sandler Morrill. (2011). "Dads and Daughters: The Changing Impact of Fathers on Women's Occupational Choices." *Journal of Human Resources* 46(2): 333-372.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith and Alina Verashchagina. (2007). "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends." *B.E. Journal of Economic Analysis & Policy* 7(2): 10.

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

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.

Kye, Bongoh, and Robert D. Mare. (2012). "Intergenerational Effects of Shifts in Women's Educational Distribution in South Korea: Transmission, Differential Fertility, and Assortative Mating." *Social Science Research* 41(6): 1495-1514.

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. (1997). "Differential Fertility, Intergenerational Educational Mobility and Racial Inequality." *Social Science Research* 26: 263-291.

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

Mare, Robert D., and Vida Maralani. (2006). "The Intergenerational Effects of Changes in Women's Educational Attainments." *American Sociological Review* 71(4): 542-564.

Martin, Teresa Castro. (1995). "Women's Education and Fertility: Results from 26 Demographic and Health Surveys." *Studies in Family Planning* 26(4): 187-202.

McFarland, David D. (1970). "Intragenerational Social Mobility as a Markov Process: Including a Time-Stationary Markovian Model that Explains Observed Declines in Mobility Rates Over Time." *American Sociological Review* 35(3): 463-476.

McGinnis, Robert. (1968). "A Stochastic Model of Social Mobility." *American Sociological Review* 33(5): 712-722.

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.

Oaxaca, Ronald L. (1973). "Male-Female Wage Differences in Urban Labor Markets." *International Economic Review* 14(3): 693-709.

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.

Preston, Samuel H., and Cameron Campbell. (1993). "Differential Fertility and the Distribution of Traits: The Case of IQ." *American Journal of Sociology* 98(5): 997-1019.

Schultz, T.P. (1993). "Returns to Women's Education." In E.M. King and M.A. Hill, eds., *Women's Education in Developing Countries: Barriers, Benefits, and Policies*. Baltimore: John Hopkins University Press.

Spilerman, Seymour. (1972). "The Analysis of Mobility Processes by the Introduction of Independent Variables into a Markov Chain." American Sociological Review 37(3): 277-294.

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.

Spalaore, Enrico, and Romain Wacziarg. (2014). "Fertility and Modernity." Mimeo, UCLA.

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. (2014). "Differential Fertility, Human Capital, and Development." Mimeo, Princeton.

Wilhelm, Mark Ottoni, Eleanor Brownb, Patrick M. Rooneya, and Richard Steinberga. (2008). "The Intergenerational Transmission of Generosity." *Journal of Public Economics* 92(10): 2146-2156.

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

Wolfinger, Nicholas H. (1999). "Trends in the Intergenerational Transmission of Divorce." *Demography* 36(3): 415-420.

Theory Appendix

Proof of Proposition 1 The first order condition for e_t leads to:

$$\frac{h(e_t)}{h'(e_t)} - e_t \begin{cases} \geq \kappa + \tau w h_t & \text{if } e_t = 0 \\ = \kappa + \tau w h_t & \text{if } e_t > 0 \end{cases}$$

The left hand side of this equation is increasing in e_t, so that in the interior solution, e_t increases with h_t. A corner solution with $e_t = 0$ occurs when $h_t \leq \tilde{h} = \frac{1}{\tau w}\left(\frac{h(0)}{h'(0)} - \kappa\right)$, which is strictly positive when κ is small. For $h_t \leq \tilde{h}$:

$$\left.\frac{\partial n_t}{\partial h_t}\right|_{h_t \leq \tilde{h}} = \frac{(1-\alpha) w\kappa}{(\kappa + \tau w h_t)^2} > 0$$

so that fertility increases in parental skill. To find the comparative static for $h_t > \tilde{h}$, rewrite the first order condition for e_t in terms of the elasticity $\eta(e_t)$:

$$e_t = \frac{\kappa + \tau w h_t}{\frac{1}{\eta(e_t)} - 1}$$

and substitute into the first order condition for n_t:

$$n_t = (1-\alpha)(1-\eta(e_t))\left(\frac{w h_t}{\kappa + \tau w h_t}\right)$$

Total differentiation implies:

$$\left.\frac{\partial n_t}{\partial h_t}\right|_{h_t > \tilde{h}} = \left(\frac{1-\alpha}{\kappa + \tau w h_t}\right)\left[(1-\eta(e_t))\left(\frac{w\kappa}{\kappa + \tau w h_t}\right) - w h_t \eta'(e_t)\frac{\partial e}{\partial h_t}\right]$$

so that $\left.\frac{\partial n_t}{\partial h_t}\right|_{h_t > \tilde{h}} < 0$ if and only if:

$$\frac{1 - \eta(e_t) - h_t \eta'(e_t)\frac{\partial e}{\partial h_t}}{h_t \eta'(e_t)\frac{\partial e}{\partial h_t}} < \frac{\tau w h_t}{\kappa}$$

By assumption, $\eta'(\cdot) > 0$, so this inequality is automatically satisfied when $1 - \eta(e_t) - h_t \eta'(e_t)\frac{\partial e}{\partial h_t} \leq 0$. When $1 - \eta(e_t) - h_t \eta'(e_t)\frac{\partial e}{\partial h_t} > 0$, differentiate the first order condition for e_t and substitute for $\frac{\partial e}{\partial h_t}$ to obtain:

$$\kappa < \tau w h_t \left(\frac{h_t \eta'(e_t)\tau w \eta(e_t)}{(1-\eta(e_t))^2 + \eta'(e_t)(\eta(e_t) - 1)(\kappa + e_t) - \eta'(e_t)\tau w h_t}\right)$$

As κ goes to 0, the limit of the right hand side is strictly positive. Therefore, with $\eta'(e_t) > 0$ and small enough κ, $\left.\frac{\partial n_t}{\partial h_t}\right|_{h_t > \tilde{h}} < 0$. ■

Proof of Proposition 2 Consider the human capital production function:

$$h_{t+1} = h(e_t) = (\theta_0 + \theta_1 e_t)^{\sigma}$$

$\theta_0 > 0$ is a human capital endowment, $\theta_1 > 0$ is the return to investment in children, and $\sigma \in (0,1)$ determines the curvature of the production function. It is easy to verify that this production function has $h(0) > 0$; finite first, second, and third derivatives; and $\eta'(e_t) > 0$. Parents choose $e_t = 0$ if $h_t \leq \tilde{h}$ and $e_t = \frac{\sigma(\kappa+\tau w h_t)-\theta_0/\theta_1}{1-\sigma}$ if $h_t > \tilde{h}$, where $\tilde{h} = \frac{1}{\tau w}\left(\frac{\theta_0/\theta_1}{\sigma} - \kappa\right)$. This behavior generates a dynamical system $h_{t+1} = \psi(h_t)$ such that:

$$\psi(h_t) \equiv \begin{cases} \theta_0^{\sigma} & \text{if } h_t \leq \tilde{h} \\ \left(\frac{\sigma\theta_1(\kappa+\tau w h_t)-\sigma\theta_0}{1-\sigma}\right)^{\sigma} & \text{if } h_t > \tilde{h} \end{cases}$$

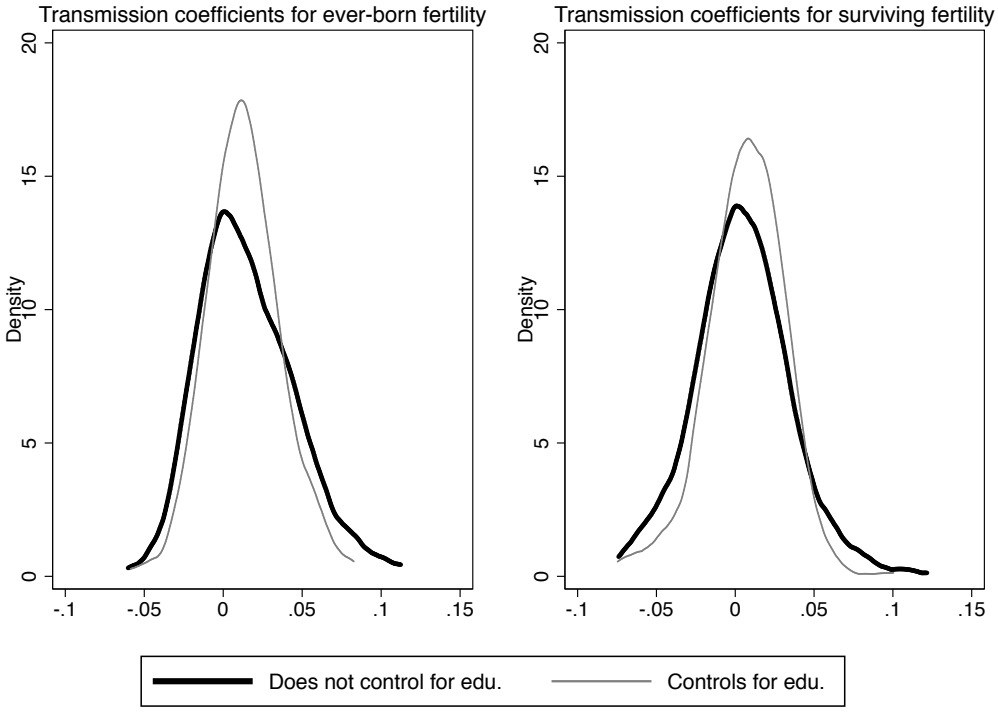
The goal is to choose parameters θ_0, θ_1, and σ to generate 1, 2, or 3 steady state levels of h_t, given the cost parameters κ and τ. First note that if $\theta_0^{\sigma} \leq \tilde{h}$, a low human capital steady state h_L obtains in which parents with $h_t = h_L = \theta_0^{\sigma}$ set $e_t = 0$, so that $h_{t+1} = h_L = \theta_0^{\sigma}$. Consider the case in which $\theta_0^{\sigma} = \tilde{h}$ holds exactly, so that θ_0 solves the equation $(\tau w)^{-1/\sigma}\left(\frac{\theta_0/\theta_1}{\sigma} - \kappa\right)^{1/\sigma} - \theta_0 = 0$. Assume $\sigma = \frac{1}{2}$, so that the polynomial has roots $\theta_0 = \kappa\sigma\theta_1 + (\tau w\sigma\theta_1)^2\left[\frac{1}{2} \pm \left(\frac{\kappa}{(\tau w)^2\theta_1\sigma} + \frac{1}{4}\right)^{1/2}\right]$, and select the larger root, which is guaranteed to be positive. Then a steady state exists at $h_L = \tilde{h}$, where $\psi(\cdot)$ is tangent to the 45 degree line. The existence of additional steady states depends on $\psi'(h_t) = \frac{\sigma^{1+\sigma}}{(1-\sigma)^{\sigma}} \cdot \frac{\theta_1\tau w}{[\theta_1(\kappa+\tau w h_t)-\theta_0]^{1-\sigma}}$. Because $\lim_{h_t\to\infty}\psi'(h_t) = 0$ and $\psi''(h_t) < 0$ for all $h_t > \tilde{h}$, $\psi(h_t)$ intersects the 45 degree line from above at another steady state $h_H > \tilde{h}$ if and only if $\psi'(\tilde{h}) > 1$. Setting $\theta_1 > \frac{(1-\sigma)\theta_0^{1-\sigma}}{\sigma^2\tau w}$ guarantees that $\psi'(\tilde{h}) > 1$. Therefore, with $\sigma = \frac{1}{2}$, the equation $\theta_0 = \kappa\sigma\theta_1 + (\tau w\sigma\theta_1)^2\left[\frac{1}{2} + \left(\frac{\kappa}{(\tau w)^2\theta_1\sigma} + \frac{1}{4}\right)^{1/2}\right]$ and the inequality $\theta_1 > \frac{2\theta_0^{1/2}}{\tau w}$ can solve for values of θ_0 and θ_1 that generate two steady human capital levels. A single steady state obtains for $\theta_1 < \frac{2\theta_0^{1/2}}{\tau w}$.

To obtain three steady states, define a new human capital production function $\hat{h}(\cdot)$ with parameters $\hat{\theta}_0$, $\hat{\theta}_1$, and $\hat{\sigma}$. Set $\hat{\theta}_0 = \theta_0 - \epsilon$ for $\epsilon > 0$, and set $\hat{\theta}_1$ and $\hat{\sigma}$ such that $\hat{\theta}_1 = \theta_1 + \frac{1}{\tau w}\left(2\theta_0^{1/2} - 2(\theta_0 - \epsilon)^{\hat{\sigma}}\right)$ and $\hat{\sigma} = \frac{\hat{\theta}_0/\hat{\theta}_1}{2\theta_0/\theta_1}$, such that $\tilde{h}$ remains unchanged, and the resulting dynamical system $\hat{\psi}(h_t)$ continues to have $\hat{\psi}'(h_t) > 1$ in the vicinity of $\tilde{h}$. A low steady state $\hat{h}_L$ still exists because $\hat{\theta}_0^{\hat{\sigma}} < \tilde{h}$. As $\epsilon \to 0$, $\hat{\psi}(h_t)$ intersects the 45 degree line from below to generate a second steady state $\hat{h}_M$. By the reasoning above, $\hat{\psi}(h_t)$ also intersects the 45 degree line from above to form a third steady state $\hat{h}_H$. Therefore, given the cost parameters κ and τ, one can choose θ_0, θ_1, and σ to guarantee three steady state human capital levels, $\hat{h}_L < \hat{h}_M < \hat{h}_H$. Since $\hat{\psi}(h_t)$ intersects the 45 degree line from above at $\hat{h}_L$ and $\hat{h}_H$ and from below at $\hat{h}_M$, $\hat{h}_L$ and $\hat{h}_H$ are stable, while $\hat{h}_M$ is unstable. ■

Proof of Proposition 3 Denote the stable steady state human capital levels in generation t as h_t^L and h_t^H, with $h_t^L < h_t^H$, and let n_t^L and n_t^H denote the associated fertility levels. Let h_{t+1}^L, h_{t+1}^H, n_{t+1}^L, and n_{t+1}^H be the same outcomes for the next generation of the L and H lineages. Then the

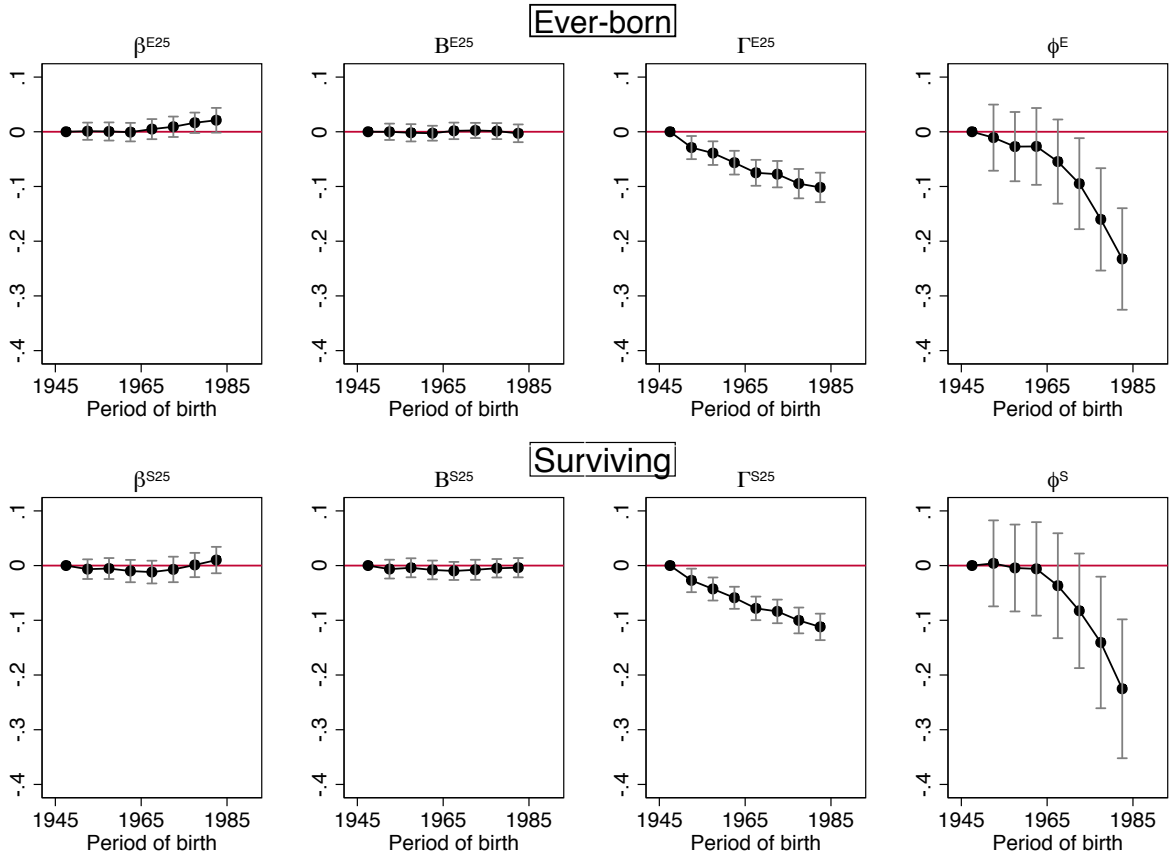
intergenerational transmission coefficient is $\beta = \frac{n_{t+1}^H - n_{t+1}^L}{n_t^H - n_t^L}$. If the child cost parameters and the human capital production function is the same in both t and $t+1$, then $h_{t+1}^i = h_t^i$ and $n_{t+1}^i = n_t^i$ for all dynasties, so $\beta = 1$. In the transition between two Malthusian steady state distributions, $n_t^L < n_t^H$, $n_{t+1}^L < n_{t+1}^H$, $n_{t+1}^L \geq n_t^L$, and $n_{t+1}^H < n_t^H$, so $\beta \in (0,1)$. In the transition between two modern steady state distributions, $n_t^L > n_t^H$, $n_{t+1}^L > n_{t+1}^H$, $n_{t+1}^L \geq n_t^L$, and $n_{t+1}^H < n_t^H$, so $\beta > 1$. In a transition between Malthusian and modern steady state distributions, $n_t^L > n_t^H$, $n_{t+1}^L < n_{t+1}^H$, $n_{t+1}^L \geq n_t^L$, and $n_{t+1}^H < n_t^H$, so $\beta < 0$.■

Figure 1: Distribution of Age-25 Intergenerational Transmission Coefficients



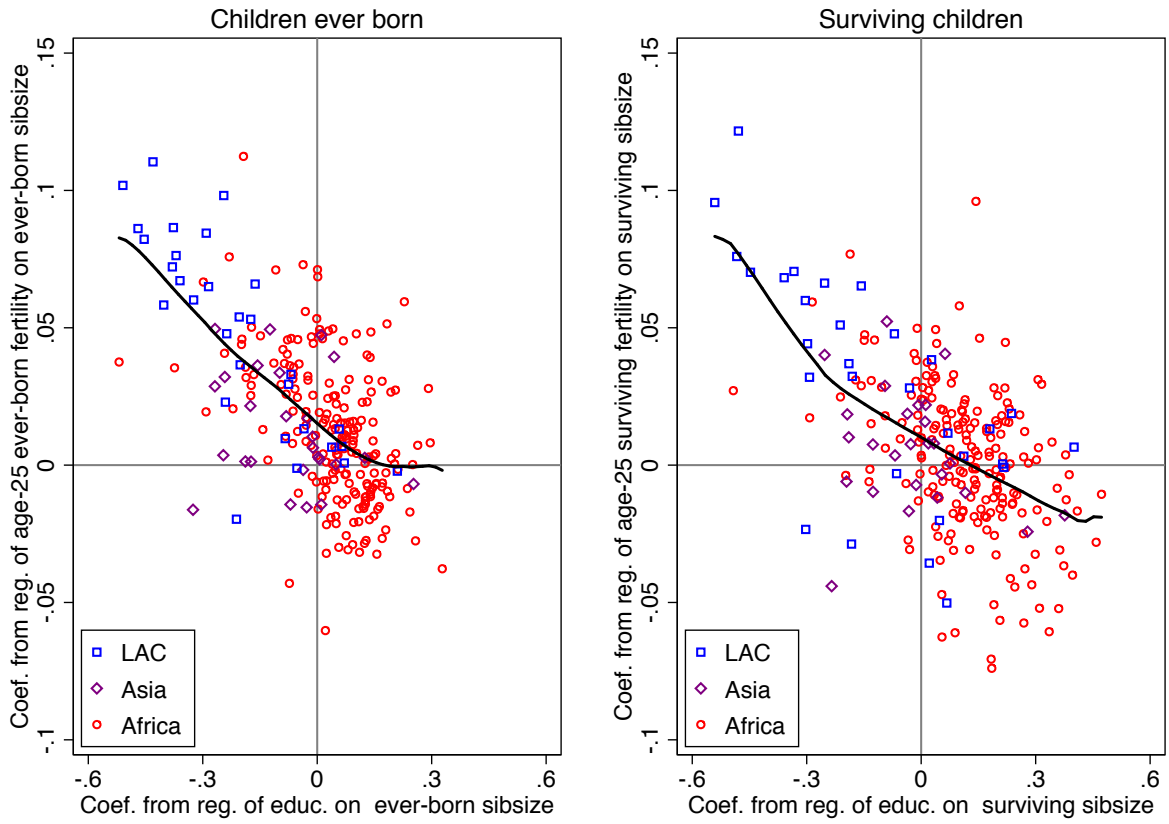
Note: Kernel density estimates with a bandwidth of 0.01. Sample includes 554,182 women from 257 country-cohort cells.

Figure 2: Cohort Effects in Age-25 Parameter Estimates



Note: Sample includes 554,182 women from 257 county-cohorts cells. Plots represent cohort effects from regressions of the estimated country-cohort parameters on country and cohort fixed effects. Gray bars represent 95% confidence intervals (based on standard errors clustered at the country level).

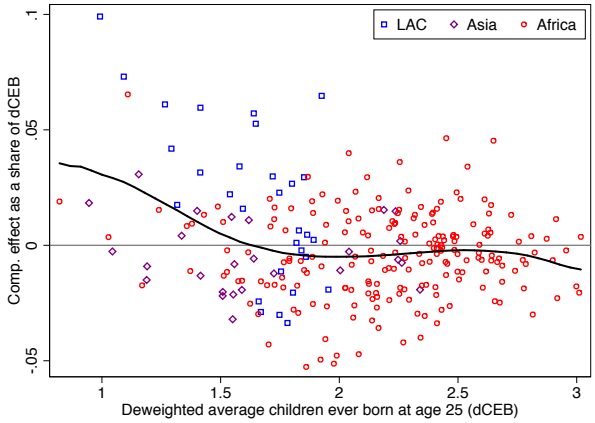
Figure 3: Age-25 Intergenerational Transmission Coefs. vs. Sibsize-Education Coefs.



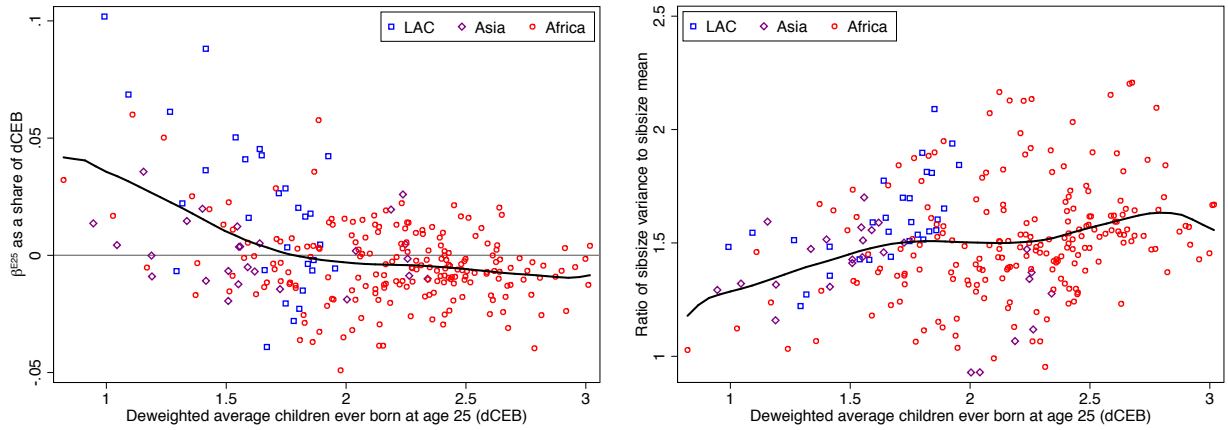
Note: Sample includes 554,182 women from 257 county-cohorts cells. Thick curves are local linear regressions with bandwidths of 0.1.

Figure 4: Composition Effects on Age-25 Fertility

A. Overall Effect



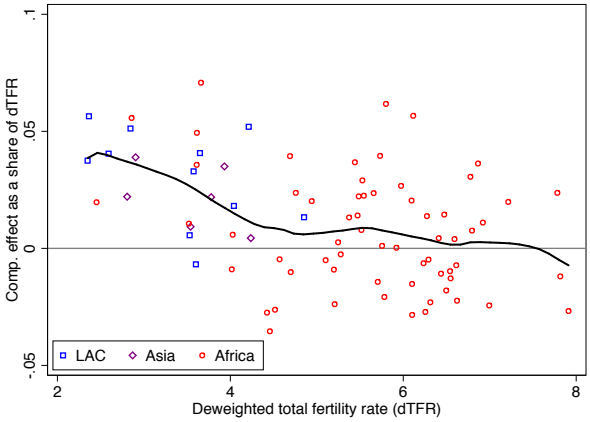
B. Decomposition, Linear Case



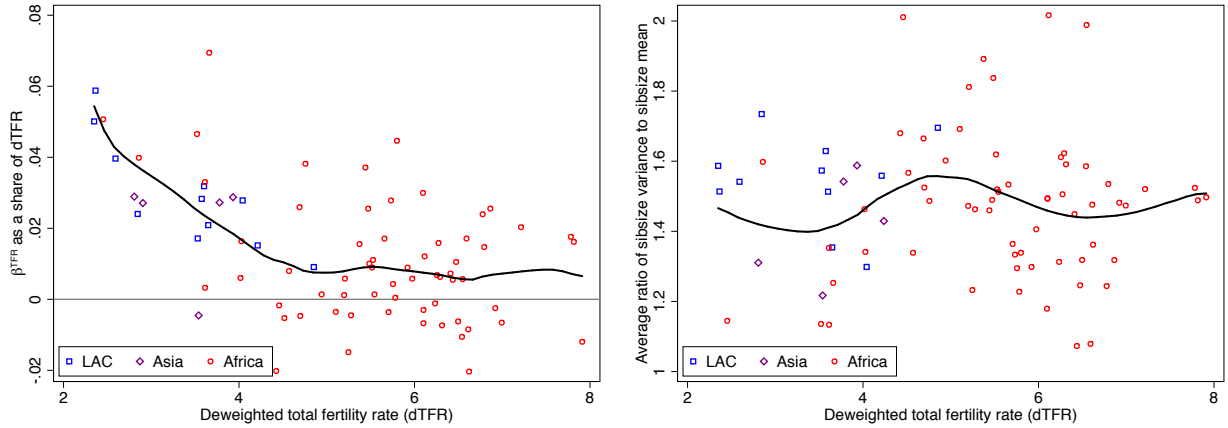
Note: Sample includes 554,182 women from 257 county-cohort cells. Curves are local linear regressions with a bandwidth of 0.5.

Figure 5: Composition Effects on the Total Fertility Rate

A. Overall Effect



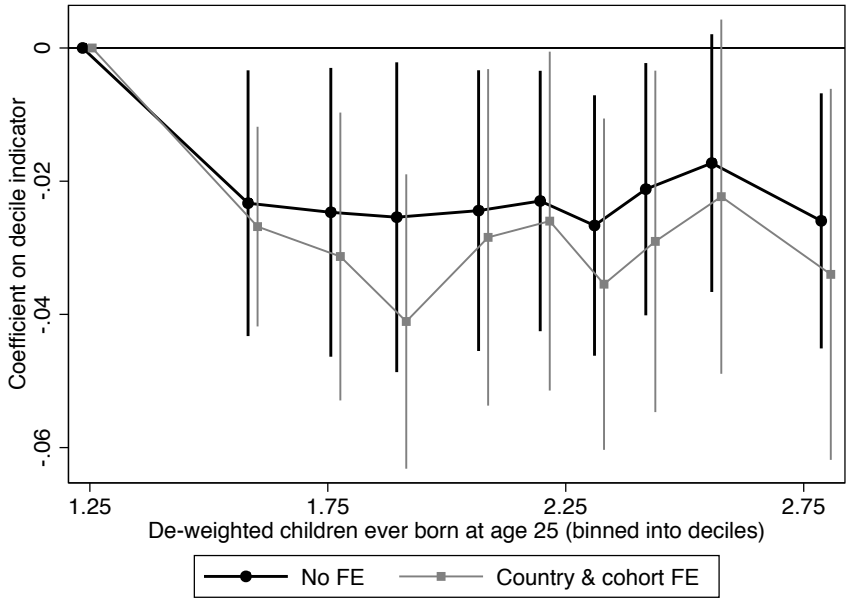
B. Decomposition, Linear Case



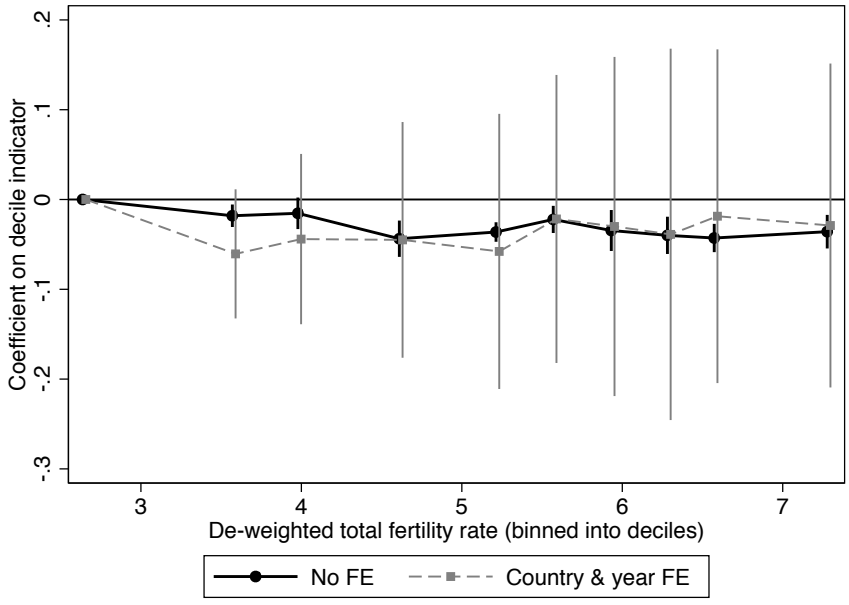
Note: 966,498 women from 80 surveys. Curves are local linear regressions with a bandwidth of 0.5.

Figure 7: Composition Effects over the Fertility Transition, Binned Models

A. Age-25 Mean Children ever Born



B. Total Fertility Rate



Note: Panel A reports coefficients and 95% confidence intervals from regressions of age-25 composition effects on indicators for deciles in the distribution of de-weighted children ever born at age 25. Sample includes 554,182 women from 257 county-cohort cells. Panel B reports coefficients and 95% confidence intervals from regressions of TFR composition effects on indicators for deciles in the distribution of de-weighted TFR. Sample includes 554,182 women from 257 county-cohort cells.

Table 1: Regression of Age-45 Transmission Coef. on Transmission Coef. at Age *a*

	Ever-born fertility		Surviving fertility	
Age *a*	(1)	(2)	(3)	(4)
20	1.29	0.71	0.77	0.96
	[0.32]	[0.57]	[0.34]	[0.27]
25	1.14	1.05	1.04	0.99
	[0.19]	[0.28]	[0.30]	[0.22]
30	1.05	0.95	1.11	0.74
	[0.11]	[0.16]	[0.23]	[0.23]
35	1.04	1.06	1.17	0.82
	[0.09]	[0.15]	[0.15]	[0.19]
40	1.07	1.04	1.03	0.80
	[0.06]	[0.12]	[0.11]	[0.14]
Country & Cohort FE	No	Yes	No	Yes

Note: Dependent variable is the age-45 transmission coefficient. Sample includes 110 country-cohort cells. Brackets contain standard errors clustered at country level.

Table 2: Summary Statistics on Age-25 Parameter Estimates

		Distribution of *t*-statistics		
	Mean [S.E.] (S.D.)	$t \leq -1.96$	$-1.96 < t < 1.96$	$t \geq 1.96$
	(1)	(2)	(3)	(4)
A. Ever-born				
$\boldsymbol{\beta}^{E25}$	.015 [.020] (.030)	4 (2%)	183 (71%)	70 (27%)
B^{E25}	.013 [.019] (.022)	3 (1%)	198 (77%)	56 (22%)
$\boldsymbol{\Gamma}^{E25}$	-.098 [.018] (.050)	227 (88%)	28 (11%)	2 (1%)
$\boldsymbol{\varphi}^{E25}$	.006 [.044] (.158)	65 (25%)	93 (36%)	99 (39%)
B. Surviving				
$\boldsymbol{\beta}^{S25}$	.005 [.022] (.030)	23 (9%)	195 (76%)	39 (15%)
B^{S25}	.007 [.021] (.024)	11 (4%)	210 (82%)	36 (14%)
$\boldsymbol{\Gamma}^{S25}$	-.070 [.016] (.051)	200 (78%)	52 (20%)	5 (2%)
$\boldsymbol{\varphi}^{S25}$	.078 [.053] (.179)	38 (15%)	87 (34%)	132 (51%)

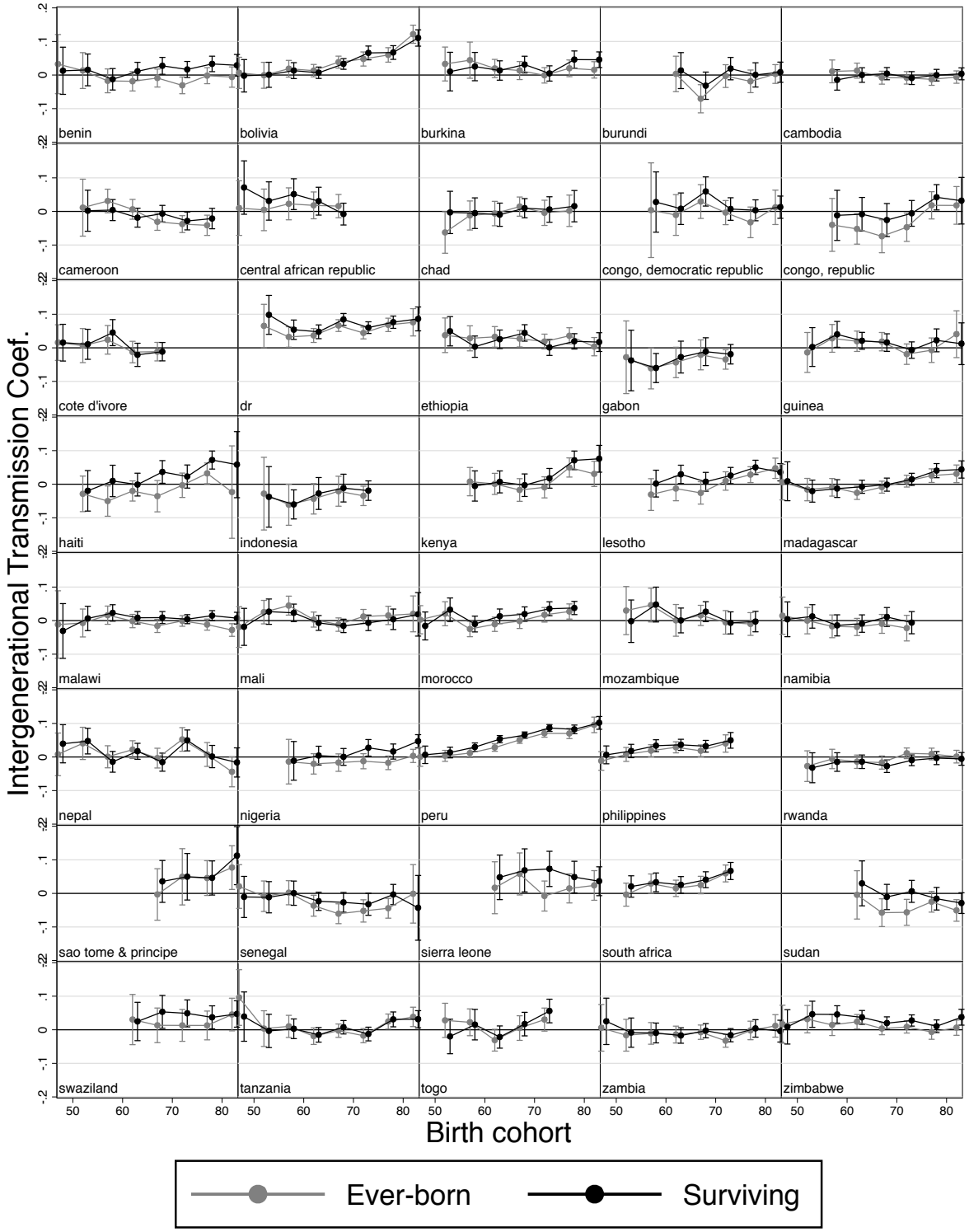
Note: Sample includes 554,182 women from 257 country-cohort cells with at least 250 observations. S.E. = standard error of the mean parameter. S.D. = standard deviation of the parameter.

Table 3: Relating Age-25 Transmission Coefs. to Other Associations

	Ever-born fertility				Surviving fertility			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sibsize-edu. association	-0.12	-0.12	0.02	-.002	-0.09	-0.10	.002	-0.01
	[0.02]	[0.03]	[0.03]	[0.04]	[0.01]	[0.02]	[0.02]	[0.03]
Edu.-fertility association			-0.04	-0.09			-0.02	-0.11
			[0.05]	[0.07]			[0.04]	[0.09]
Interaction			1.15	1.14			1.08	0.98
			[0.23]	[0.22]			[0.19]	[0.22]
Overall R-squared	0.38	0.67	0.44	0.71	0.28	0.61	0.37	0.64
Within R-squared		0.18		0.27		0.12		0.24
Country & Cohort FE	No	Yes	No	Yes	No	Yes	No	Yes

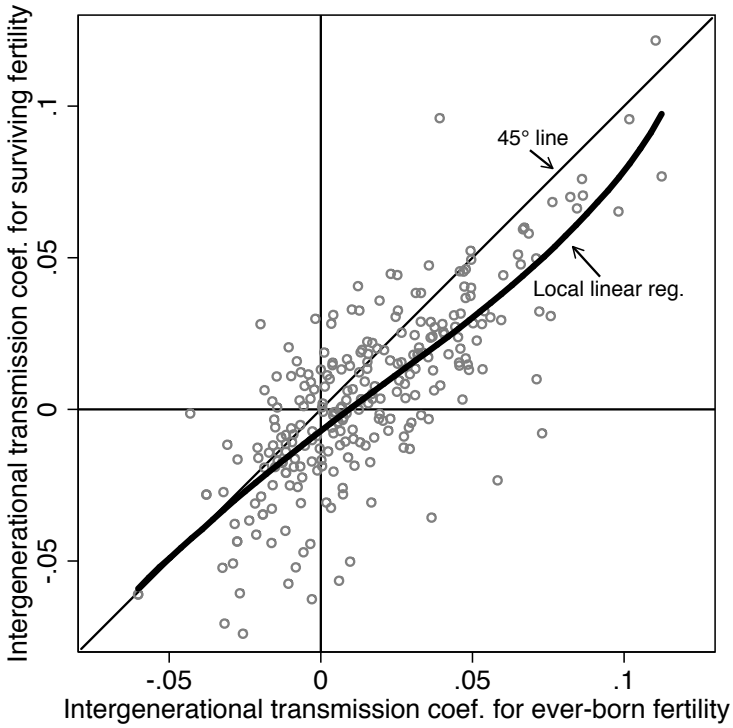
Note: Dependent variable is the age-25 transmission coefficient. Sample includes 554,182 women from 257 county-cohort cells. Brackets contain standard errors clustered at the country level.

Appendix Figure 1: Intergenerational Transmission Coefficients for Age-25 Fertility



Note: Coefficients and 95% confidence intervals (adjusted for cluster-based survey design). Sample includes 554,182 women from 257 county-cohort cells.

Appendix Figure 2: Age-25 Intergenerational Transmission Coefs.: Surviving vs. Ever-born



Note: Sample includes 554,182 women from 257 county-cohort cells. The local linear regression uses a bandwidth of 0.03. The correlation between the coefficients is 0.77.

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

Appendix Table 2: Regression of Age-45 Transmission Coef. on Transmission Coef. at Age *a*
Country-Cohort Cells with at Least 200 Observations

	Ever-born fertility		Surviving fertility	
Age *a*	(1)	(2)	(3)	(4)
20	1.11	0.61	0.30	0.67
	[0.36]	[0.59]	[0.31]	[0.37]
25	1.14	0.98	0.81	0.75
	[0.20]	[0.32]	[0.37]	[0.33]
30	1.08	0.83	1.05	0.47
	[0.15]	[0.24]	[0.36]	[0.40]
35	1.08	0.99	1.19	0.75
	[0.11]	[0.22]	[0.21]	[0.35]
40	1.09	1.01	1.04	0.87
	[0.07]	[0.14]	[0.14]	[0.19]
Country FE	No	Yes	No	Yes
Cohort FE	No	Yes	No	Yes

Note: Dependent variable is the age-45 transmission coefficient (from the "short" regression). Sample includes 99 country-cohort cells. Brackets contain standard errors clustered at country level.

Appendix Table 3: Regression of Age-25 Trans. Coef. on Other Associations
Country-Cohort Cells with at Least 500 Observations

	Ever-born fertility				Surviving fertility			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sibsize-edu. association	-0.12	-0.15	0.03	-0.02	-0.09	-0.11	0.020	-0.02
	[0.02]	[0.03]	[0.03]	[0.04]	[0.01]	[0.02]	[.021]	[0.03]
Edu.-fertility association			-0.06	-0.14			-0.04	-0.10
			[0.05]	[0.09]			[0.05]	[0.11]
Interaction			1.27	1.20			1.28	0.96
			[0.24]	[0.25]			[0.19]	[0.27]
Overall R-squared	0.41	0.70	0.49	0.74	0.31	0.66	0.43	0.68
Within R-squared		0.24		0.34		0.17		0.28
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Cohort FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: Dependent variable is the age-25 transmission coefficient (from the "short" regression). Sample includes 225 county-cohort cells. Brackets contain standard errors clustered at the country level.