Policy Research Working Paper 10285

Growing Up Together

Sibling Correlation, Parental Influence, and Intergenerational Educational Mobility in Developing Countries

Md. Nazmul Ahsan
M. Shahe Emran
Hanchen Jiang
Qingyang Han
Forhad Shilpi

WORLD BANK GROUP
Development Economics
Development Research Group
January 2023

Policy Research Working Paper 10285

Abstract

This paper presents credible and comparable evidence on intergenerational educational mobility in 53 developing countries using sibling correlation as a measure, and data from 230 waves of Demographic and Health Surveys. It is the first paper to provide estimates of sibling correlation in schooling for a large number of developing countries using high quality standardized data. Sibling correlation is an omnibus measure of mobility as it captures observed and unobserved family and neighborhood factors shared by siblings when growing up together. The estimates suggest that sibling correlation in schooling in developing countries is much higher (average 0.59) than that in developed countries (average 0.41). There is substantial spatial heterogeneity across regions, with Latin America and Caribbean having the highest (0.65) and Europe and Central Asia the lowest (0.48) estimates. Country level heterogeneity within a region is more pronounced. The evolution of sibling correlation suggests a variety of mobility experiences, with some regions registering a monotonically declining trend from the 1970s birth cohort to the 1990s birth cohort (Latin America and the Caribbean and East Asia and Pacific), while others remained trapped in stagnancy (South Asia and Sub-Saharan Africa). The only region that experienced monotonically increasing sibling correlation is the Middle East and North Africa. The recent approach of Bingley and Cappellari (2019) is used to estimate the share of sibling correlation due to intergenerational transmission. The estimates show that when the homogeneity and independence assumptions implicit in the standard model of intergenerational transmission are relaxed, the estimated share is much larger. In the sample of countries, on average 74 percent of sibling correlation can be attributed to intergenerational transmission, while there are some countries where the share is more than 80 percent (most in Sub-Saharan Africa). This suggests a dominant role for parents in determining the educational opportunities of their children. Evidence on the evolution of the intergenerational share, however, suggests a declining importance of the intergenerational transmission component in many countries, but the pattern is diverse. In some cases, the trend in the intergenerational share is opposite to the trend in sibling correlation.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp.

The authors may be contacted at fshilpi@worldbank.org.

Growing Up Together: Sibling Correlation, Parental Influence, and Intergenerational Educational Mobility in Developing Countries

Md. Nazmul Ahsan, St. Louis University[1]

M. Shahe Emran, IPD at Columbia University

Hanchen Jiang, University of North Texas

Qingyang Han, Johns Hopkins University

Forhad Shilpi, DECRG, World Bank

JEL Codes: J0, D3, J62

Key Words: Sibling Correlation, Intergenerational Mobility, Education, Years of Schooling, Developing Countries, Intergenerational Share, Decomposition, DHS

[1]Emails for correspondence: shahe.emran.econ@gmail.com. (M. Shahe Emran); fshilpi@worldbank.org (Forhad Shilpi). We would like to thank Lorenzo Cappellari for helpful discussion on econometric issues, and Ira Gang, Kunal Sen, Lewis Anderson and participants at a seminar at University of Windsor for useful comments on an earlier version.

(1) Introduction

A vast literature on intergenerational mobility and inequality of opportunity in economics and sociology focuses on the role of family and neighborhood background in shaping the life chances of a child.[2] This literature, however, primarily concentrates on developed countries, and the evidence base on developing countries remains relatively sparse. There has been a spurt in the interests among development economists in understanding the geography and evolution of intergenerational mobility in the recent years, partly spurred by the evidence of increasing inequality in the 1980s and 1990s in many developing countries (World Bank (2006), Chancel et al. (2022)).[3] Among the policymakers, there has been an increasing emphasis on inequality of opportunity as opposed to inequality of outcome.[4] The observed increase in cross-sectional inequality in socioeconomic outcomes is of serious concern when it reflects increasing inequality of opportunity and declining intergenerational mobility.

It is, however, difficult to build a credible and comparable cross-country evidence base on intergenerational mobility in developing countries because of data limitations. Two essential building blocks are required for such an analysis: a measure of socioeconomic status of children and parents, and a measure of intergenerational mobility or inequality of economic opportunity.[5] Permanent income has been the preferred measure of socioeconomic status in the economic literature on developed countries, but reliable panel data for a long enough period to estimate permanent income remain scarce in developing countries.[6] In the absence of income data, it is feasible to use education data across a large number of developing countries,

[2]For recent surveys of the economic literature see, Black and Devereux (2011), Bjorklund and Salvanes (2011), Solon (1999), Heckman and Mosso (2014), Mogstad and Torsvik (2021). For surveys of the sociology literature see Breen (2010) and Torche (2015a).

[3]The focus on geography of intergenerational mobility also reflects the influential work of Chetty et al. (2014) which constructed an opportunity atlas with estimates of absolute mobility at the zip code level in the United States.

[4]For example, German President Joachim Gauck declared equality of opportunity as the normative policy goal during his inauguration in 2012. US president Barack Obama in his 2014 state of the union address mentioned "opportunity" 10 times.

[5]It has been increasingly emphasized in the literature that intergenerational mobility and inequality of opportunity are closely related, and they focus on the same fundamental question, even though these two strands of the literature developed independently. For discussions, see Bjorklund and Jantti (2020) and Emran and Shilpi (2021).

[6]There are other difficulties in measuring household and individual income in a developing country. First, because of large informal home-based economic activities, it is difficult to measure income. Second, It is difficult to measure individual income in an extended family living and eating together in the same household. For a discussion on this point, see Iversen et al. (2019).

but a major concern here is that parental (usually father's) education may give us only a partial measure of socioeconomic status of a child. To address this, one can include a vector of observable family characteristics in a model of intergenerational mobility by adopting the latent socioeconomic approach developed by Lubotsky and Wittenberg (2006) (henceforth called LW approach following Emran and Shilpi (2021)).[7] A second solution is offered by the literature on inequality of opportunity (henceforth IOP) that grew out of the influential work of Roemer (1998) and Coleman et al. (1966). The IOP approach uses multiple indicators such as father's education and occupation, sex of a child, ethnicity etc. as the "circumstances" for which a child should not be held responsible as they did not choose them, rather inherited by birth. A practical challenge for both of these approaches (IOP and LW) is that the set of family background indicators one can use is dictated by the lowest common denominator across the surveys in different countries. In practice, these approaches thus rely on a limited number of observed family characteristics when studying inequality of opportunity and intergenerational mobility across countries (see the discussion by Bjorklund and Jantti (2020) on IOP). A related important limitation is that these approaches often fail to take into account the effects of unmeasured and difficult to measure family characteristics (e.g., cultural inheritance) and neighborhood characteristics (e.g., school and peer quality). It is especially difficult to get reliable and comparable data on these tacit aspects of family and neighborhood environment when studying a large number of countries.

In a forthcoming paper, Deutscher and Mazumder (2022) discuss 19 measures of intergenerational mobility and find that some of them are only weakly correlated with each other, suggesting that they measure very different concepts of economic mobility. When different studies use different measures of mobility capturing disparate economic concepts, it is impossible to rank countries in terms of intergenerational mobility based on the evidence from individual country studies. Many recent studies on developing countries use measures of intergenerational association based on the conditional expectation function of children's schooling given father's schooling. However, even among the subset of studies focusing on the conditional expectation function of schooling, comparability is compromised in two ways. First, some studies use years of schooling, some use years of schooling normalized by standard de-

[7]Notwithstanding its advantages, the LW approach has not been adopted widely yet. We are aware of only two studies that used this approach: Neidhofer et al. (2018) and Vosters and Nybom (2017).

viation, and some others rely on ranks in the schooling distribution.[8] Second, even when based on the same measures of education and mobility, comparability across countries is often compromised by a lack of standardization of the data. Some surveys include data on years of schooling while others only categorical variables such as primary, secondary, etc. These different measures of schooling are likely to contain different degrees of measurement error and a study based on such data may not yield comparable estimates of intergenerational educational mobility across countries.

To deal with these challenges, we use sibling correlation in education as a measure of mobility, and take advantage of data from Demographic and Health Surveys (DHS). Sibling correlation stands out as a measure on both conceptual and empirical grounds. Conceptually, sibling correlation is a broader measure of intergenerational mobility compared to most of the other measures used in the literature.[9] First, the standard measures such as correlation between parents and children's schooling attainment usually rely solely on father's education because of missing data on mother's education. Sibling correlation reflects the effects of both father and mother along with other family members residing in the household such as grandparents, uncles, and aunts. Second, similarity between sibling's educational outcomes does not only reflect *observable* family characteristics, but also *unobservable* (to the researcher) factors shared by siblings such as family culture. It also captures other factors such as parenting style, aspiration, and risk attitude, even when the data sets do not contain any information on these variables. In this sense, sibling correlation is broader than the IOP and LW approaches which can only use the *observable* characteristics measured in a survey.[10] Third, sibling correlation captures broader neighborhood effects (including school and peers) that are not correlated with parental education but shared by the siblings growing up together in the same neighborhood.

An important advantage of sibling correlation as a measure is that it is less susceptible to coresidency bias because we want to capture the factors shared by siblings when growing

[8]For a comparative analysis of these alternatives measures of intergenerational educational mobility with a focus on possible conflicts among them, please see Ahsan et al. (2022). They use data from China, India, and Indonesia.

[9]For discussions on this point, see Bjorklund and Salvanes (2011), Bjorklund et al. (2010), Deutscher and Mazumder (2022).

[10]IOP can include variables that represent factors not shared by the siblings. See the discussion by Bjorklund and Jantti (2020).

up together, going to the same neighborhood school, socializing with the same cohorts of peers. Thus, we would like to exclude the siblings who grew up far apart in different times (or different places like a sibling who left home for a boarding school), and might have experienced different family, neighborhood, and school environment when growing up.[11] This includes, for example, much older children of the household head who are not coresident because they are in college or working in a different location after graduating from college. We restrict the age of children to 16-28 years in the survey year of a wave for our analysis.[12] This also ensures that possible sample truncation because of coresidency is small.[13]

We provide estimates of sibling correlation in educational attainment for 53 developing countries and trace out the evolution of intergenerational educational mobility for three decade-wise birth cohorts (1970s to 1990s).[14] We use 230 waves of Demographic and Health Surveys (DHS) to build a comparable data base across countries. This ensures that the estimates are not different because of differences in survey instruments or measurement of schooling (years of schooling vs. categorical). Another important advantage is that the information on parental education is not based on recall of the children, and thus are likely to contain much less measurement error. The advantages of the Demographic and Health Surveys for cross-country comparison studies have been well appreciated in the recent literature (see, for example, Bhalotra and Rawlings (2013), and Lleras-Muney et al. (2022)).

Another important contribution of this paper is that we take advantage of recent methodological advances to provide a credible answer to a long-standing policy-relevant question:

[11]This is supported by our cohort-wise estimates that show significant changes in sibling correlation over time. Nicoletti and Rabe (2013) provide evidence that sibling correlation in academic attainment declines sharply as more distantly spaced siblings are considered.

[12]Thus, the maximum age difference between siblings in our sample can be 12 years. Bingley and Cappellari (2019) also use a maximum of 12 years age gap to define their sample. This implies that the cohort variations in our analysis primarily come from different survey years of different waves of DHS for a given country. We note here that the main conclusions are robust to alternative age ranges. These alternative results are available from the authors.

[13]It is well understood in the literature that the degree of sample truncation due to coresidency declines as we focus on younger children. See, for example, the recent influential analysis on intergenerational educational mobility in Africa by Alesina et al. (2021) where they use census data (coresident samples) from IPUMS and restrict age of children to 18-24.

[14]The only other paper we are aware of that provides estimates of sibling correlation for multiple developing countries is Dahan and Gaviria (2001) who report estimates for 16 Latin American countries. But their estimates are not comparable to the other estimates available in the literature because they use a different measure. They focus on the educational failure (lack of grade progression) of a child rather than educational attainment.

how long is the father's shadow cast on the siblings? A major focus of the literature on developed countries has been on the share of sibling correlation accounted for by father's education or income (more broadly, parental education or income). A substantial literature on sibling correlation in income in developed countries suggests that the share of the intergenerational transmission may be small (see Bjorklund and Salvanes (2011) and Solon (1999)). A small intergenerational share would imply that the focus of research and policy should be on the neighborhood and school factors rather than family (Solon (1999)). However, the recent analysis by Bingley and Cappellari (2019) suggests that this low estimate is due to the restrictive assumptions of homogeneity and independence in the estimates of the intergenerational transmission from the parents to the children. They develop an approach that relaxes these assumptions, and find that the share accounted for by the intergenerational transmission of income in sibling correlation in income is much higher. We provide estimates of intergenerational share in sibling correlation in schooling using the classic methods (Solon (1999), Bjorklund et al. (2010), Mazumder (2008)) along with the Bingley and Cappellari (2019) approach.

The evidence suggests four key conclusions. First, the estimates of sibling correlation in our sample of developing countries are, in general, substantially larger than the existing estimates for the developed countries. The average in our 53 country sample is 0.59, and the average for the top half of the distribution is 0.65. Based on 56 estimates available for developed countries, the average sibling correlation in schooling is 0.41.[15] Our estimates thus suggest a considerable gap in educational opportunities between the developing and developed worlds.[16]

Second, there is significant spatial heterogeneity at the regional level and across countries within a region. The Latin America and Caribbean region experienced the worst educational

[15]The 56 estimates are for the same birth cohorts as our sample: 1960s to 1990s birth cohorts. Bjorklund and Salvanes (2011) report a range of 0.40-0.60 for developed countries with the estimates for the United States among the highest. Prag et al. (2019) report an average of 0.49 from a meta analysis of the studies on sibling correlation in income and education published between 1972-2018 (includes both developing and developed countries).

[16]The existing cross-country evidence for developing countries based on father-child correlation in schooling also finds lower mobility in developing countries (see, for example, Neidhofer et al. (2018)). However, the existing evidence cannot answer whether this pattern holds for broader measures of socioeconomic status. We provide the first evidence that this conclusion holds even when we capture a broad set of observed ad unobserved family and neighborhood factors.

opportunities with an estimated average sibling correlation of 0.65, and East Asia and Pacific is the second worst (0.64), while Europe and Central Asia had the lowest sibling correlation estimate (0.48). The within region heterogeneity is also substantial; for example, in Sub-Saharan Africa, the maximum estimate is 0.77 (Madagascar) and the minimum is 0.49 (South Africa).

Third, the evolution of sibling correlation from the 1970s birth cohort to the 1990s birth cohort suggests a rich variety of mobility experiences. At the regional level, some experienced monotonic improvements in intergenerational educational mobility (Latin America and Caribbean, and East Asia and Pacific), while some others faced stagnation (South Asia, and Sub-Saharan Africa). Middle East and North Africa stands out as the only region to have a declining trend in educational opportunities (i.e., monotonically increasing sibling correlation from the 1970s to the 1990s). However, the regional average conceals a lot of country level heterogeneity. For example, notwithstanding a stagnant regional average in South Asia, Bangladesh achieved substantial gains in educational opportunities with sibling correlation declining from 0.67 (1970s) to 0.61 (1990s). In contrast, Pakistan experienced a declining intergenerational educational mobility with sibling correlation increasing from 0.60 (1970s) to 0.68 (1990s).

Fourth, estimates of the role played by intergenerational transmission between father and children vary dramatically depending on the decomposition method used. Consistent with the analysis of Bingley and Cappellari (2019), we find that the estimated share of intergenerational transmission in sibling correlation is considerably higher when we relax the homogeneity and independence assumptions implicit in the standard methods of decomposition. The estimates from the Bingley and Cappelari (2019) approach suggest an average of 74 percent across 53 countries, and in some countries, the share is higher than 80 percent (many of them in Sub-Saharan Africa). In contrast, the average share is only 34 percent according to the estimates from Bjorklund et al. (2010) approach.[17] The estimates across birth cohorts show that the share of intergenerational transmission has declined in many countries from the 1970s to the 1990s birth cohort. But there are 13 countries where the share has increased over decades, many of them (11) are located in the Sub-Saharan Africa region.

[17] The average is 30 percent according to the Solon (1999) method, and 18 percent according to the Mazumder (2008) method.

The rest of the paper is structured as follows. Section (2) discusses the related literature and puts the contributions of this paper in perspective. Section (3) is devoted to the conceptual framework that describes the measure of sibling correlation and the decomposition methods for estimating the share of intergenerational transmission in sibling correlation. A special focus here is on the Bingley and Cappalleri (2019) approach. Section (4) discusses the advantages of the Demographic and Health Surveys for our cross-country analysis and provides a brief discussion of the estimation methods. Section (5), arranged in a number of subsections, reports and discusses the estimates of sibling correlation. Section (6) discusses the estimates of the share of the intergenerational transmission across regions and countries, and traces out the evolution over time from the 1970s birth cohort to the 1990s birth cohort. The paper ends with a summary of the results and the contributions of the paper in the conclusions.

(2) Related Literature

The economics literature on intergenerational mobility is grounded on the seminal contributions of Becker and Tomes (1986) that developed a model of intergenerational persistence in permanent income focusing on the role of human capital. The inequality of opportunity (IOP) strand of the literature builds on the foundation of the theory of distributive justice developed by Roemer (see Roemer (1998), and Roemer and Trannoy (2016)). The inequality of opportunity (IOP) refers to the "circumstances" a child is born into, and emphasizes that inequality due to the circumstances is unjust and should be the focus of policy interventions. Although these two approaches grew largely independently, there has been an increasing appreciation that they deal with fundamentally the same question.[18] These two approaches can be best viewed as complementary. The IOP provides a theory of justice foundation, but does not identify the economic mechanisms which could be the policy levers. The Becker-Tomes model identifies a set of such economic mechanisms. The sociological literature uses occupational prestige and class mobility with a focus on the long-term factors including the role of formal and informal institutions, especially in the labor market (see Torche (2015a), Breen (2010)). In the recent decades, many sociologists adopted the regression-based approach of economists and appeal to the Becker-Tomes model for theoretical underpinning of their results.

[18] See, for example, the discussion by Deutscher and Mazumder (2022), and Bjorklund and Jantti (2020). We discuss the differences between these two approaches later in the paper.

As noted in the introduction, the literature on developing countries mainly focuses on intergenerational educational mobility because of data limitations. Although there is a growing literature studying the persistence of educational attainment across generations at the country level, the studies that attempt to provide comparable estimates across a sample of developing countries remain limited.[19] The most widely known cross-country analysis of intergenerational educational mobility is Hertz et al. (2008) that provides estimates of relative mobility using intergenerational regression coefficient (IGRC) and intergenerational correlation (IGC) between father and children for 42 countries. Neidhofer et al. (2018) report estimates of a variety of measures of absolute and relative educational mobility for 18 Latin American countries. A number of recent papers focus on Sub-Saharan African countries, see, for example, Alesina et al. (2021), Azomahou and Yitbarek (2021), and Razzu and Wambile (2022).[20] The most extensive analysis of intergenerational educational mobility around the world is offered in a recent book by Narayan et al. (2018) covering 153 countries. They provide estimates of a number of absolute and relative educational mobility measures, but their main analysis is based on the IGRC estimates. Perhaps, more important, all the cross-country studies noted above focus on the intergenerational link between parents' and children's educational attainment, and none report estimates of sibling correlation.

In more than two decades following the publication of the handbook of labor economics chapter by Solon (1999), there have been only a few studies on developing countries that use sibling correlation as a measure of educational mobility. This is puzzling because Solon (1999) and the subsequent surveys of the field (e.g., Bjorklund and Salvanes (2011)) provide substantial discussions on the advantages of sibling correlation as a measure, especially in the data scarce environment common in developing countries. The most widely cited is a study by Dahan and Gaviria (2001) who report estimates of sibling correlation in schooling for 16 Latin American countries. But as noted earlier, they do not follow the methodology developed in Solon et al. (1991), and Solon (1999). They use a measure of *educational failure* (lack of grade

[19] At the individual country level, recent contributions on intergenerational educational mobility in developing countries include Kundu and Sen (2022), Azam and Bhatt (2015), Azam (2016), Emran and Shilpi (2015) on India, Fan et al. (2021), Emran and Sun (2015) on China, Torche (2015b) on Mexico, Assaad and Saleh (2018) on Jordan; Ahsan et al. (2023), Ahsan et al. (2021) on Indonesia. For surveys of this literature, please see Iversen et al. (2019), Torche (2019), and Emran and Shilpi (2021).

[20] These studies rely on census data from IPUMS.

progression) rather than *educational attainment* of children, and construct an index different from the measure of mobility used by Solon (1999) and other studies. Their index is based on the index of segregation proposed by Kremer and Maskin (1996). Their estimates are thus not comparable to the other estimates of sibling correlation in the literature, including the estimates reported in this paper. In a meta analysis of sibling correlation estimates published between 1972-2018, Prag et al. (2019) identify only two studies on developing countries including that of Dahan and Gaviria (2001), the second study is on intergenerational educational mobility in post-reform India by Emran and Shilpi (2015).

In contrast, the literature on sibling correlation in education and income in developed countries is substantial with contributions from both economists and sociologists. For surveys of this literature, please see Solon (1999), Bjorklund and Salvanes (2011), Bjorklund and Jantti (2020). Given the focus of the economic literature in developed countries on income, many of the existing studies provide estimates of sibling correlation in income. But the literature on sibling correlation in education is also large. Most of the estimates of sibling correlation in schooling in developed countries fall in the range of 0.40-0.60 (see Bjorklund and Salvanes (2011)). Among recent contributions, Grtz et al. (2021) report estimates of sibling correlation in education for 6 developed countries with an average estimate of 0.44, the lowest estimate of 0.36 (Finland) and the highest 0.51 (United States and Germany).

(3) Conceptual Framework

(3.1) Sibling Correlation

For the estimation and interpretation of sibling correlations, we adopt a conceptual framework that has been the workhorse in the empirical literature on sibling correlations (see, Solon et al. (1991), Solon (1999), Bjorklund et al. (2002), Bjorklund and Lindquist (2010), Mazumder (2008) and (2011)). Following Solon (1999) and Bjorklund et al. (2010), we begin with a simple model of children's educational attainment:

$$\tilde{S}_{if} = \mu + \Gamma X_i + a_f + b_{if} \tag{1}$$

Where $\tilde{S}_{if}$ is measure of educational attainment, usually years of schooling, of sibling i in family f, μ is the country specific component that captures the factors common to all children of a country, and X_i is a set of individual characteristics elements of which depend on the propose of the analysis. Following Bjorklund et al. (2010), we include a gender dummy, and, following Bingley and Cappellari (2019), we include cohort dummies, but no other controls are included in X_i.[21] a_f is the family and neighborhood component shared by all siblings in family f, and b_{if} is the individual specific component for sibling i capturing i's deviation from the common family and country components. We define demeaned years of schooling S_{if} as follows:

$$S_{if} = \tilde{S}_{if} - (\mu + \Gamma X_i) = a_f + b_{if} \tag{2}$$

The focus of the analysis is on the importance of the family and neighborhood component (the family component, for short) a_f in explaining the variance in demeaned years of schooling S_{if}.[22] The country mean μ represents the "growth and structural change" in a country that influence all children the same way irrespective of their family background. The cohort dummies take out the cohort specific effects shared by the children of a cohort, but may vary across different cohorts. The inclusion of the country and cohort specific intercepts in the vector X_i implies that the measure of mobility based on sibling correlation in demeaned schooling refers to *relative* rather than *absolute* mobility. Assuming that a_f is independent of b_{if}, the variance of S_{if} can be expressed as the sum of variances of the family and individual components as:

$$\sigma_s^2 = \sigma_a^2 + \sigma_b^2 \tag{3}$$

The sibling correlation in education (denoted by ρ_s) then can be expressed as:

$$\rho_s = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2} \tag{4}$$

Sibling correlation thus estimates the share of variance of children's education that can be attributed to common family and neighborhood background.

Sibling correlation is a measure of mobility (more precisely immobility) because the fam-

[21] Some studies on intergenerational mobility in education include age and age squared following the literature on intergenerational mobility in income. However, age and age squared are used in studies on income to mop up life-cycle effects. For education, such life-cycle effects are not relevant.

[22] In the variance components analysis, a_f is usually called the family component as it represents the family fixed effect.

ily and neighborhood factors shared by the siblings growing up together are not chosen by themselves, but they are born into it. Thus, this measure is consistent with the inequality of opportunity foundation of distributive justice *a'la* Roemer (1998).

As discussed by Emran and Shilpi (2015), the basic insight of the Becker and Tomes (1986) model that imperfections in the credit market lead to lower mobility also holds for sibling correlation. When the credit market is perfect and parents can borrow at a given interest rate r to finance children's education, optimal investment is independent of family background and depends only on the ability of a child. Under the assumption that the distribution of innate ability does not depend on family background, the variance in the average education of children across families captured by σ_a^2 would be approximately zero. Now, consider the credit market imperfections model of Becker et al. (2018) where the poor (less educated) parents has access to credit market for children's education, but have to pay a higher rate, and the rich (and highly educated) pay low interest rate: $r_l > r > r_h$ with subscripts l and h referring to low educated and high educated parents. In this case, r represents the interest rate faced by the families in the middle of the distribution. Parents in the low educated families optimally invest less in children's education at a given ability level, and the average education of siblings increases with the level of parental education. This increases the variance in children's schooling across families, thus making σ_a^2 and sibling correlation positive. Note that the strength of sibling correlation increases with the degree of credit market imperfections as captured by differences in the interest rates faced by different households. The important point here is that sibling correlation as measure of mobility is grounded in the political philosophy foundation of theory of justice developed by Roemer (1998), and also consistent with the main insights of the Becker-Tomes model.

An important advantage of sibling correlation is that it captures all the observed and unobserved (including the unobservables) family and neighborhood factors shared by siblings while growing up together. This, however, does not mean that sibling correlation provides an upper bound for the effects of family and neighborhood factors on educational opportunities of children. As noted by Bjorklund et al. (2010), while sibling correlation is a broader measure, it is in fact a *lower bound estimate of the effects of family and neighborhood background, because* *it does not include the factors not shared by siblings.*[23]

(3.2) Intergenerational Correlation vs. Sibling Correlation

Given that there is a large literature on intergenerational persistence in education, a natural question to ask is how much of the sibling correlation can be accounted for by the intergenerational transmission from parent's (usually father) to children. If the widely used measures of intergenerational educational mobility such as intergenerational regression coefficient (IGRC) or intergenerational correlation (IGC) can explain most of the sibling correlation, then it would suggest primacy of the family and parents in shaping the educational opportunities of children, and policy should focus on the family.[24] The link between intergenerational transmission and sibling correlation has been a focus of the literature since the early contributions on sibling correlation in income in the United States by Solon et al. (1988) and Solon et al. (1991).

A simple approach to understanding the role of the intergenerational transmission is to estimate sibling correlation with and without conditioning on parental education. Mazumder (2008) uses this approach to estimate the share of parental influences in sibling correlation in income in the United States, but does not estimate the share of intergenerational correlation in sibling correlation in education of children. Emran and Shilpi (2015) adopts this approach to estimate the share of father's influences in sibling correlation in post-reform India.

A second and more widely used approach was developed earlier by Solon (1999). Following Solon (1999) and Bjorklund et al. (2010), we can derive the relation between sibling correlation and intergenerational correlation. We can decompose the family and neighborhood component a_f into two orthogonal parts:[25]

$$a_f = \beta S_f^p + \lambda_f^R \tag{5}$$

[23]Bjorklund and Jantti (2020) note that, for most of the data sets, sibling correlation is a broader measure than inequality of opportunity even though one can include some of the non-shared factors (e.g., birth order) as part of the vector of circumstances in an IOP approach. As noted earlier, in a cross-country analysis only a few indicators of circumstances are included because the feasible set is determined by the lowest common denominator.

[24]IGRC is the slope estimate from a regression of children's schooling on parent's schooling. In contrast, IGC is the slope from a regression model where both children's and parent's schooling are normalized by their respective generation-specific standard deviation in schooling. IGC thus estimates Pearson Correlation.

[25]As noted by Bjorklund and Jantti (2020), this decomposition of sibling correlation was first derived by Solon (1999). But the sociology literature on sibling correlation contains informal discussion on this before the formal derivation by Solon (1999).

where βS_f^p is the part due to parental education and λ_f^R is the residual sibling effect. Taking variance of equation (5), we have:

$$\sigma_a^2 = \beta^2\sigma_p^2 + \sigma_{\lambda R}^2 \tag{6}$$

Dividing through by σ_s^2 we get:

$$\rho_s = \frac{\beta^2\sigma_p^2}{\sigma_s^2} + \frac{\sigma_{\lambda R}^2}{\sigma_s^2} = IGC^2 + Residual\ Sibling\ Correlation \tag{7}$$

If one assumes stationary distributions across generations, then $\sigma_p^2 = \sigma_s^2$ and we have

$$\rho_s = \beta^2 + Residual\ Sibling\ Correlation \tag{8}$$

In fact, Solon (1999) derived the decomposition under the assumption of stationary distributions as in equation (8), while Bjorklund et al. (2010) used equation (7).

Residual sibling correlation represents all other factors shared by siblings but uncorrelated with parental education. Many studies on intergenerational mobility in developed countries used equations (7) and (8), and the conclusion from this literature is that only a small part of sibling correlation could be explained by the parental education. According to the estimates for years of schooling reported by Bjorklund and Jantti (2020), the IGC estimate for Sweden is 0.30 and sibling correlation is 0.43. The squared IGC is thus 0.09, only about 20 percent of sibling correlation is explained by IGC.

However, equation (5) is motivated by the workhorse linear mobility equation for estimating IGC which imposes a number of assumptions that are likely to be rejected on both theoretical and empirical grounds. Recent theoretical advances suggest that the assumption of linearity is likely to be violated in many cases. Becker et al. (2015) develop a model of intergenerational educational persistence between parents and children where the mobility equation can be concave (due to diminishing returns) or convex (due to complementarities).[26] A concave or convex intergenerational persistence equation has two important implications:

[26] Recent evidence suggests that the mobility equation is not linear in most of the cases. Emran et al. (2020) finds that the mobility equation in India is concave irrespective of gender. Ahsan et al. (2022) provides evidence suggesting concave or convex mobility equations for years of schooling in China, India, and Indonesia.

(i) the effects of parents on children as captured by IGC (β) are heterogeneous across families; and (ii) the parameter β can be positively (for convex mobility function) or negatively (concave mobility function) correlated with parental education. Bingley and Cappellari (2019) develop a decomposition method that allows for heterogeneous β and arbitrary correlation between β and S_f^p. They show that, for sibling correlation in income, relaxation of the implicit assumptions in equation (5) makes a big difference.

To the best of our knowledge, we are the first to implement the Bingley and Cappellari (2019) approach for estimating the intergenerational share in sibling correlation in education, and we do it for a large number of developing countries (53 countries) using comparable data from the DHS surveys. We provide a brief discussion of the Bingley and Cappellari (2019) approach below, and refer the reader to the original paper for details.

(3.3) Decomposition of Sibling Correlation: Bingley and Cappellari (2019) Approach

In the context of our set-up, Bingley and Cappellari (2019) replace equation (5) by the following random coefficient specification:

$$a_f = \left(\bar{\beta} + \beta_f\right) S_f^p + \lambda_f^R \tag{9}$$

where $\bar{\beta}$ is average effect of parental education and β_f is deviation of family f from the mean. This specification thus incorporates heterogeneity in the effects of parental education captured by the parameter β_f. If we relax only the heterogeneity assumption but retain the assumption that the magnitudes of the parental effect is independent of the level of parental education, we have the following decomposition:

$$\rho_s = \frac{\left(\bar{\beta}^2 + \sigma_\beta^2\right)\sigma_p^2}{\sigma_s^2} + \frac{\sigma_{\lambda R}^2}{\sigma_s^2} \tag{10}$$

But as we discussed above, there are plausible theoretical models that suggest that β_f is correlated with S_f^p. Using a result on the exact variance of the product of two random variables due to Bohrnstedt and Goldberger (1969), Bingley and Cappellari (2019) derive the following decomposition (under normality):

$$\rho_s = \frac{\left(\bar{\beta}^2 + \sigma_\beta^2\right)\sigma_p^2 + cov\left(\beta_f S_f^p\right)^2}{\sigma_s^2} + \frac{\sigma_{\lambda R}^2}{\sigma_s^2} \tag{11}$$

Since $cov\left(\beta_f S_f^p\right)^2 \geq 0$, assuming independence in equation (10) will in general underestimate the role of the intergenerational transmission of schooling. The evidence on intergenerational income mobility in Denmark reported by Bingley and Cappellari (2019) suggests that the relaxation of the independence assumption is especially important; the estimated share of the intergenerational component (father's income) increases substantially as a result.

The decomposition in equation (11) relaxes two important restrictive assumptions in the standard specification (5): homogeneity in β_f and independence between β_f and S_f^p, but it relies on the normality assumption which is rejected by the data in most of the cases. Bingley and Cappellari (2019) find that imposing normality tends to underestimate the share of intergenerational influences in sibling correlation. They relax the normality assumption by using an unrestricted form of the intergenerational correlation between the children and parents. In our empirical analysis, we will report estimates from both the classic methods (Bjorklund et al. (2010), Mazumder (2008), Solon (1999)), and the method due to Bingley and Cappellari (2019).

(4) Data and Estimation Methods

A major hurdle for credible cross-country ranking of inequality of opportunity and intergenerational mobility is that data from different surveys may not be comparable. As noted earlier, the survey instruments used for education information by DHS are standardized across countries which makes the data much more comparable.[27] Even when trying to elicit the same information (say education of parents and children), different household surveys may contain different kinds of data. In the context of intergenerational educational mobility, there are two issues relevant here. First, whether data on educational attainment refer to years of schooling or education categories (primary, secondary etc.). The DHS data we use have information on years of schooling for both the parents (father) and children. Second, in many household surveys used for intergenerational mobility analysis in developing countries, data on parental

[27] We take the information from the household roster which is the same in all DHS surveys.

education are based on children's recalled information, and thus may contain non-negligible measurement error (Emran and Shilpi (2021), Torche (2019)). This would tend to bias downward the estimated parents-children persistence in education and, in turn, lead to downward biased estimate of the share of the intergenerational transmission in sibling correlation. The DHS data on parents are not based on recall, and thus are much more reliable.

There are 53 countries in our sample. We use 230 waves of DHS surveys to build our data base. We exclude 42 countries where at least one DHS survey is available but the sample size is small. The cut-off for inclusion is a minimum of 1000 observations in the sample. The trade-off between country coverage and sample size is well-appreciated in the literature. For a recent analysis of intergenerational educational mobility covering a large number of countries (153), see Narayan et al. (2018), but, as noted earlier, they do not provide estimates of sibling correlation.[28]

In each wave of DHS, our sample includes children of age 16-28 in the survey year. The exclusion of relatively older age cohorts in each wave is motivated by two considerations. First, it reduces the possibility of sample truncation due to grown-up children leaving the household for work or to start a family. Second, as noted earlier, we would like to exclude children who are born far apart as they are likely to face different family, neighborhood, and school environments. Among our 53 countries, there are 6 countries with fewer than 2000 observations, and 22 countries with sample size more than 5000. The total number of observations in our data set is 544624. The country level estimation samples include children from the 1960s to 1990s birth cohorts. But in many countries, the number of observations for the 1960s birth cohort is small because only a limited number of DHS surveys were administered in these countries in the 1990s and earlier. For the analysis of the evolution of educational mobility across cohorts we thus do not include 1960s observations, and focus on the three decade wise birth cohorts from 1970s to 1990s.

The estimation method adopted by Bingley and Cappellari (2019) is the method of moments. The data requirements for the analysis are more demanding because the Bingley and

[28]The price of the extensive country coverage in Narayan et al. (2018) is that in 57 countries, the sample size is less than 1000 observations, and in 19 countries less than 500 observations. There are 25 countries with more than 5000 observations. The authors are very much aware of this trade-off and report the number of observations for each estimate so that a reader can make an informed judgment. The study by Hertz et al. (2008) include 42 countries with a minimum sample size of 1047 observations (Philippines).

Cappellari (2019) approach is based on family triads with father and two children in a family. We take the oldest two children from those families where the number of children is more than 2. To ensure that the siblings are not too far apart, we follow Bingley and Cappellari (2019) and restrict their age gap to a maximum of 12 years. The intergenerational transmission is estimated as the average of the persistence between father and the first child, and between father and the second child in the sample. The birth cohorts are defined based on the birth year of the older sibling in a household. For the estimation of the share of intergenerational component, we do not impose the stationary distributions assumption across generations used by Bingley and Cappellari (2019) as this assumption is rejected by our data.[29] We also find that the estimated share can be more than 100 percent if we incorrectly impose stationary distribution assumption within the Bingley and Cappellari (2019) approach. The estimates from the Mazumder (2008) method for the share of intergenerational component are implemented using Restricted Maximum Likelihood (REML) in a mixed effects model.

(5) Evidence: Geography of Sibling Correlation and Evolution over Three Birth Cohorts

(5.1) Geography of Sibling Correlation across Regions and Nations

The estimates suggest that there are substantial regional variations in intergenerational educational mobility as measured by sibling correlation. Figure 1 presents the average sibling correlation estimates for six regions of the world. The country specific estimates are reported in Table 1. The estimates suggest that intergenerational educational mobility for the 1960s-1990s birth cohorts is the lowest in the Latin American and Caribbean countries with an average sibling correlation of 0.65.[30] Compare this with an average of 0.41 for developed countries noted before. This evidence on Latin America and Caribbean is interesting as the countries in this region also experienced some of the highest income inequality during this period (De Ferranti (2004)). Thus, high cross-sectional inequality was coupled with low intergenerational mobility, a doubly undesirable distributional outcome. Among the countries in

[29] In the income data used by Bingley and Cappellari (2019), the null hypothesis of stationary distributions is not rejected. Stationary distributions are also assumed by Solon (1999).

[30] The only other study that reports estimates of sibling correlation in schooling for Latin American countries is Dahan and Gaviria (2001). However, as discussed earlier, their estimates are not comparable to our estimates or other estimates in the literature.

this region, Guatemala has the unfortunate distinction of having the highest sibling correlation in schooling: 0.71, and the country with the lowest sibling correlation is the Dominican Republic with an estimate of 0.57.

Intergenerational educational mobility is also low (comparable to Latin America) in East Asia (average estimate 0.64) and South Asia (average estimate 0.62).[31] Among the East Asian countries, Cambodia and Vietnam have the lowest intergenerational educational mobility, with a sibling correlation estimate of 0.66 in both countries. The sibling correlation estimate for Philippines is the lowest in this region (0.60). In South Asia, the estimates are very close in four out of five countries, ranging from 0.62 (Nepal) to 0.64 (Bangladesh). Afghanistan enjoys the highest intergenerational educational mobility with an estimate of 0.56.

We have two countries from the Middle-East and North Africa region for which the required DHS data were available: the Arab Republic of Egypt and Jordan.[32] The estimate suggests that sibling correlation is much lower compared to the three regions discussed above (Latin America and Caribbean, South Asia, and East Asia and Pacific). Sibling correlation in schooling is 0.48 in Jordan and 0.54 in Egypt which are smaller than, for example, the estimate for the most mobile country in South Asia, Afghanistan (0.56).

For Sub-Saharan Africa, we have 30 countries (please see Table 1 for the list of the countries), with an average sibling correlation of 0.59. On average, Sub-Saharan Africa is more mobile than Latin America, South Asia and East Asia, but the mean estimate hides substantial heterogeneity across countries. The highest estimate is 0.77 for Madagascar which is also the highest among our 53 countries. There are three other countries with estimates of 0.70 or higher: Chad (0.74), Nigeria (0.70), and Ethiopia (0.70). The lowest estimate is 0.49, for South Africa.

The region with the highest intergenerational educational mobility is Europe and central Asia; the average sibling correlation is 0.48.[33] Among the 5 countries from this region in

[31] Our East Asia sample does not include Japan because DHS surveys do not cover Japan. The countries included are: Cambodia, Indonesia, Philippines, and Vietnam. The South Asia sample includes Afghanistan, Bangladesh, India, Nepal, and Pakistan.

[32] As discussed in the data section, we excluded the countries with DHS survey if the sample size is less than 1000 observations.

[33] This average estimate is slightly higher than the average of 0.44 for 6 developed countries reported by Grtz et al. (2021). The developed countries are: Finland, Norway, Germany, United States, United Kingdom, and Sweden. The countries in our sample are: Albania, Armenia, Kyrgyz Republic, Tajikistan, and Turkiye.

our sample, Kyrgyz Republic comes out at the top with an estimated sibling correlation of 0.38 which is also the lowest among the 53 countries in Table 1. Turkiye (previously known as Turkey) and Armenia share the unfortunate distinction of the lowest intergenerational mobility in this region with an estimated sibling correlation of 0.55.

(5.2) Evolution Over Time: Estimates from Decade Wise Birth Cohorts

As noted earlier, in many countries, the sample size for the 1960s birth cohort is too small for credible estimation of sibling correlation. We thus focus on the three decade wise birth cohorts, from the 1970s to the 1990s. The children born in the 1970s are likely to face significantly different economic and educational policies when compared to the children born in the later decades. There were two major policy developments in the 1980s and the following decades that might have affected the educational opportunities of children. First, there were economic liberalization and reform across many developing countries including trade liberalization, privatization and deregulation. The reform yielded impressive economic growth and substantial reductions in poverty in many countries, but at the same time increased income inequality (World Bank (2006)). Second, there were dramatic expansion of schools in many developing countries.[34] Many countries also implemented compulsory primary and secondary schooling in the decades of 1980s-2000s. As a result, significant gains in school enrollment and schooling attainment were achieved over these decades (World Bank (2018)). Did the poverty reduction and the expansion of schooling and other educational policies outweigh the countervailing effects of inequitable growth and might have actually expanded the educational opportunities for the children in the later decades? Are there important regional differences in the evolution of inequality of educational opportunities over these decades? We make some progress on these questions in this section.

Figure 2 presents the estimates of sibling correlation for the six regions dis-aggregated by the decade of birth (1970s-1990s birth cohorts).[35] The first impression that jumps out of

[34]For a comprehensive discussion on the school expansion in developing countries, see chapter 2 titled "The great school expansion- and those it has left behind" in World Bank (2018). For recent analysis of the effects of school construction on intergenerational mobility, see Mazumder et al. (2019) and Ahsan et al. (2023).

[35]The countries in a region in Figure 2 may vary from Figure 1, as we included only those countries for which estimates for all three decades are available. For example, Figure 2 does not include Brazil where the last DHS survey was done in 1986, and as a result, we do not have enough observations for the 1980s and 1990s birth cohorts. A cohort-wise graph including all countries can be found in the online appendix.

Figure 2 is that there are substantial regional heterogeneity in the evolution of inequality of educational opportunity. Of the 6 regions, 2 show monotonic improvements over the three decades, they are Latin America and Caribbean, and East Asia and Pacific. The largest decline in sibling correlation is experienced in the Latin America and Caribbean region (14.71 percent reduction, from 0.68 in the 1970s to 0.58 in the 1990s), with East Asia and Pacific also achieving a substantial decline (9.23 percent reduction, from 0.65 (1970s) to 0.59 (1990s)). The substantial improvements in intergenerational educational mobility in Latin American countries is a welcome development because of its historically high income inequality levels (De Ferranti (2004)). In fact, all 5 countries in our Latin America and Caribbean sample registered better intergenerational mobility for the 1990s birth cohort compared to that for the 1970s birth cohort. However, even after substantial decline over three decades, the estimated sibling correlation for the 1990s birth cohort remains much larger in Latin America compared to the estimate for Europe and Central Asian countries in our sample (0.42 in the 1990s).

Middle East and North Africa stands out as the only region where we observe a monotonically increasing average sibling correlation from the 1970s birth cohort to the 1990s birth cohort.[36] Although sibling correlation was low for the 1970s cohort in these countries (0.50), it increased by 14 percent to 0.57 in the 1990s cohort which is close to the estimate of 0.58 for the Latin America and Caribbean region for the same birth cohort.

In contrast, the changes in sibling correlation in South Asia and Sub-Saharan Africa are not monotonic across different birth cohorts. More important, the magnitudes of changes are rather small: a less than 2 percent decline in the sibling correlation estimate from the 1970's cohort to the 1990's cohort in both regions. In South Asia, the estimated sibling correlation declined marginally from 0.64 in the 1970s to 0.63 in the 1990s, although the cohorts born in the 1980s experienced a slightly better outcome (sibling correlation 0.62). This picture of stagnation in South Asia, however, conceals important heterogeneity; for example, the trajectories of change over time are opposite in Bangladesh vs. Pakistan. Sibling correlation declined substantially in Bangladesh from 0.67 in the 1970s cohort to 0.61 in the 1990s cohort, while Pakistan experienced a substantial increase from 0.60 in the 1970s to 0.68

[36]A caveat here is that we have two countries from this region in our sample so the average estimate may not be representative of other countries of this region. But Egypt is by far the largest country in the region. These countries have about 20 percent of the region's population.

in the 1990s cohort. Evidence on India, by far the largest country in the region, suggests that intergenerational mobility remained largely unchanged over the three birth cohorts. This is striking because following extensive economic reforms including dramatic trade liberalization and domestic deregulation in 1991, India reaped impressive economic growth and poverty reduction in the decades of 1990s and 2000s during which the children of the 1980s and 1990s birth cohorts went to school.[37] The evidence thus suggests that the gains in growth and poverty reduction failed to translate into better educational opportunities for the children of liberalization in India.[38]

As noted earlier, Sub-Saharan Africa as a region also did not experience any substantial improvements over the three decades. Again, the average estimates conceal substantial country level diversity in mobility experiences. We observe some of the most dramatic declines in intergenerational educational mobility in this region. For example, sibling correlation in Mozambique increased from 0.52 in the 1970s cohort to 0.68 in the 1990s cohort, and in Nigeria from 0.64 (1970s) to 0.74 (1990s). There are also a number of countries in this region that experienced substantial improvements. For example, sibling correlation in Uganda declined from 0.65 (1970s) to 0.55 (1990s), and in Tanzania, from 0.56 (1970s) to 0.48 (1990s). Out of 27 countries in this region for which we have estimates for both the 1970s and the 1990s cohorts, 16 countries registered improvements, while 10 experienced a setback in intergenerational educational mobility.

[37]Based on Indian government official poverty line, the proportion of poor people in rural areas declined from 47 percent in 1983 to 28 percent in 2004-2005. The corresponding decline in urban India is from 42 percent in 1983 to 26 percent in 2004=2005. See Bank (2011).

[38]This evidence of no significant improvements in educational opportunities in India is in contrast to the evidence of substantial improvements based on the estimates of intergenerational regression coefficient (IGRC) in educational attainment reported by Azam and Bhatt (2015), Jalan and Murgai (2008), Kishan (2018), and Maitra and Sharma (2010). However, this conclusion is consistent with the analysis of Emran and Shilpi (2015) which uses two rounds of DHS (called NFHS in India) surveys (1992/93 and 2006) and focuses on the 16-27 year old children in the survey year. Using sibling correlation and intergenerational correlation (IGC), they show that there has been almost no change in educational opportunities from 1992/93 to 2006. Since IGRC and IGC are partial measures and cannot take into account many factors shared by the siblings, one can make a plausible argument in favor of the conclusions based on sibling correlation estimates. We will discuss later the changes in the share of intergenerational correlation (IGC) over time in our sibling correlation estimates. Please see section 6 below.

(6) How Long is the Father's Shadow? Estimating the Intergenerational Share in Sibling Correlation

To understand the importance of intergenerational transmission between the father and children, we primarily rely on the Bilgley and Cappellari (2019) approach. For comparison, we also report estimates from the three "traditional" methods used in the literature: Solon (1999), Bjorklund et al. (2010), and Mazumder (2008). A comparison of the three traditional methods shows that the estimates from the Mazumder (2008) approach are the lowest in magnitude, while the estimates from the Bjorklund et al. (2010) approach are the largest. Recall that Bjorklund et al. (2010) do not impose the stationary distribution assumption, unlike Solon (1999). As discussed earlier, our estimates using the Bingley and Cappellari (2019) approach also do not impose the stationarity assumption. We focus on a comparison of the estimates from Bjorklund et al. (2010) and Bingley and Cappellari (2019) methods in our discussion below.

(6.1) Geography of Intergenerational Share

Figure 3A presents the estimated share of the intergenerational transmission for our six regions based on the Bingley and Cappellari (2019) method. The corresponding shares from the Bjorklund et al. (2010) method are in Figure 3B. A comparison of these two methods suggests three major conclusions. First, the estimates from the Bingley and Cappellari (2019) approach are much larger: the lowest estimate is 0.70 (MENA region), while the highest estimate from the Bjorklund et al. (2010) approach is only 0.40 (East Asia and Pacific). The average intergenerational share for the 53 countries is 74 percent according to the Bingley and Cappellari (2019) approach, while it is only 34 percent according to the Bjorklund et al. (2010) approach.[39] This is consistent with the evidence on income mobility in Denmark reported by Bingley and Cappellari (2019), and vindicates, in a much wider context, their argument that the low estimates in the existing literature are driven by restrictive homogeneity and independence assumptions. Second, the ranking of regions may change depending on the method of decomposition used. For example, according to the Bjorklund et al. (2010) method, the share of the intergenerational transmission is larger in East Asia and Pacific (40 percent)

[39] The estimates from the other two traditional methods are even lower, and in particular, the method due to Mazumder (2008) seems to yield very low estimates.

than that in South Asia (0.34). But the share is identical in these two regions (76 percent) according to the Bingley and Cappellari (2019) estimates.

The disaggregated country level estimates of the intergenerational share are reported in Table 3 using the Bingley and Cappellari (2019) method. The estimated share is high in most of the countries (more than 60 percent in every case), and there are some countries where 80 percent or more of the sibling correlation can be attributed to the intergenerational transmission between father and the children. They are Philippines (0.81) and Vietnam (0.82) in South East Asia, Bangladesh (0.82) and Pakistan (0.81) in South Asia, and Benin (0.82), cameroon (0.85), the Republic of Congo (0.85), Madagascar (0.85), Mozambique (0.82), Senegal (0.80), and Togo (0.82) in Sub-Saharan Africa. Interestingly, none of the countries in the Latin America and Caribbean region have such a high share of the intergenerational transmission even though some of these countries have very high sibling correlation.

(6.2) Evolution of Intergenerational Share

We next look at the evolution of the intergenerational share across the three birth cohorts in the six regions. Figure 4 presents the results based on the Bingley and Cappellari (2019) method. It is striking that in every region, the share of intergenerational transmission declined from the 1970s cohort to the 1990s cohort, even though in some cases the magnitude is negligible (for example, Latin America and Caribbean where the share was 74 percent in the 1970s and 73 percent in the 1990s) . This can be interpreted as a declining role of parents in shaping the educational opportunities of children over time. The evolution of the share over time offers some contrasting patterns when compared to the evidence on sibling correlation across cohorts discussed earlier. The share of the intergenerational transmission remained virtually unchanged in the Latin America and Caribbean despite the substantial decline in the sibling correlation we discussed above. In the Middle East and North Africa region, the intergenerational share declined substantially across the cohorts which stands in sharp contrast to the monotonically increasing magnitudes of sibling correlation.

The individual country level estimates of the intergenerational share across cohorts show a variety of mobility experiences. Although the share of intergenerational transmission declined in most of the cases, there are some countries, especially in the Sub-Saharan Africa, which experienced a higher share in the 1990s (out of the 13 countries with a higher share, 11 are in Sub-Saharan Africa). In South Asia, all countries experienced a decline in the intergenerational share, with India registering the largest reduction. The evidence suggests that the evolution of intergenerational share does not depend systematically on the level or evolution of sibling correlation in a country. The share of intergenerational transmission declined in both Pakistan and Bangladesh even though their trajectories for sibling correlation were opposite (an increasing sibling correlation in Pakistan, and a declining one in Bangladesh).

(7) Conclusions

We provide comparable estimates of intergenerational educational mobility for 53 developing countries using sibling correlation as an omnibus measure, and data from 230 waves of Demographic and Health Surveys. Sibling correlation is an omnibus measure because it captures all the observed and unobserved family and neighborhood factors shared by the siblings when growing up together. Sibling correlation is thus a much broader measure compared to the other widely used measures in the literature such as intergenerational regression coefficient, intergenerational correlation (Pearson correlation), and intergenerational rank-rank slope. Another important advantage is that sibling correlation is less susceptible to the biases caused by coresidency restrictions in the surveys as missing older children who grew up many years ago in different family and neighborhood environment does not bias the estimates. The Demographic and Health Surveys provide high quality data on schooling of children and father (years of schooling, not categorical), and the data on father's schooling are not based on children's recall. To the best of our knowledge, this is the first paper to provide estimates of sibling correlation in schooling for a large number of developing countries using high quality data standardized across countries.

The estimates suggest that sibling correlation in schooling in developing countries is much higher (average 0.59) than that in developed countries (average 0.41). We find substantial spatial heterogeneity across regions, Latin America and Caribbean with the highest (0.65) and Europe and Central Asia with the lowest (0.48) estimates. Country level heterogeneity within a region is more pronounced. The evolution of sibling correlation suggests a variety of mobility experiences, with some regions registering a monotonically declining trend from the 1970s birth cohort to the 1990s birth cohort (Latin America and Caribbean and East Asia and Pacific), while others remained trapped in stagnancy (South Asia and Sub-Saharan Africa).

The only region that experienced monotonically increasing sibling correlation is Middle East and North Africa implying consistently declining educational mobility across cohorts. We take advantage of the recent approach of Bingley and Cappellari (2019) to estimate the share of sibling correlation due to intergenerational transmission of schooling from parents to children. We find that relaxing the homogeneity and independence assumptions implicit in the standard methods of decomposition makes the estimated share much larger. In our sample, at least 60 percent of sibling correlation can be attributed to the intergenerational component, while there are some countries where the share is more than 80 percent (most in Sub-Saharan Africa). The average intergenerational share for the 53 countries is 74 percent. This suggests a dominant role for the parents in shaping educational opportunities of children, providing an argument in favor of causal analysis and economic policies focusing on the family. Evidence on the evolution of the intergenerational share, however, suggests a declining importance of the intergenerational transmission component in many countries, but the pattern is very diverse. In some cases, the trend in the intergenerational share is opposite to the trend in sibling correlation.

References

Ahsan, M. N., Shilpi, F., and Emran, S. (2023). Public Primary School Expansion, Gender-Based Crowding Out, and intergenerational educational mobility. Working paper, DECRG, The World Bank.

Ahsan, N., Emran, M. S., and Shilpi, F. (2021). Complementarities and Intergenerational Educational Mobility: Theory and Evidence from Indonesia. MPRA Paper 111125, University Library of Munich, Germany.

Ahsan, N. M., Emran, M. S., Jiang, H., Murphy, O., and Shilpi, F. J. (2022). When Measures Conflict: Towards a Better Understanding of Intergenerational Educational Mobility. Working paper, The World Bank.

Alesina, A., Hohmann, S., Michalopoulos, S., and Papaioannou, E. (2021). Intergenerational Mobility in Africa. *Econometrica*, 89(1):1–35.

Assaad, R. and Saleh, M. (2018). Does Improved Local Supply of Schooling Enhance Intergenerational Mobility in Education? Evidence from Jordan. *World Bank Economic Review*, 32(3):633–655.

Azam, M. (2016). Intergenerational Educational Persistence among Daughters: Evidence from India. *The B.E. Journal of Economic Analysis & Policy*, 16(4):1–16.

Azam, M. and Bhatt, V. (2015). Like Father, Like Son? Intergenerational Educational Mobility in India. *Demography*, 52(6):1929–1959.

Azomahou, T. T. and Yitbarek, E. (2021). Intergenerational Mobility in Education: Is Africa Different? *Contemporary Economic Policy*, Forthcoming.

Bank, W. (2011). *Perspectives on poverty in India: Stylized facts from survey data*. The World Bank.

Becker, G., Kominers, S. D., Murphy, K., and Spenkuch, J. (2015). A Theory of Intergenerational Mobility. MPRA Paper 66334, University Library of Munich, Germany.

Becker, G. and Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4(3):1–39.

Becker, G. S., Kominers, S. D., Murphy, K. M., and Spenkuch, J. L. (2018). A Theory of Intergenerational Mobility. *Journal of Political Economy*, 126(S1):7–25.

Bhalotra, S. and Rawlings, S. (2013). Gradients of the Intergenerational Transmission of Health in Developing Countries. *The Review of Economics and Statistics*, 95(2):660–672.

Bingley, P. and Cappellari, L. (2019). Correlation of Brothers' Earnings and Intergenerational Transmission. *The Review of Economics and Statistics*, 101(2):370–383.

Bjorklund, A. and Jantti, M. (2020). Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches. *Research in Social Stratification and Mobility*, 70:100455.

Bjorklund, A., Lindahl, L., and Lindquist, M. J. (2010). What More Than Parental Income, Education and Occupation? An Exploration of What Swedish Siblings Get from Their Parents. *The B.E. Journal of Economic Analysis & Policy*, 10(1):1–40.

Bjorklund, A. and Salvanes, K. G. (2011). Education and Family Background: Mechanisms and Policies. In Hanushek, E., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 3, pages 201–247. Elsevier.

Black, S. E. and Devereux, P. J. (2011). Recent Developments in Intergenerational Mobility.

In *Handbook of Labor Economics*, volume 4, chapter 16, pages 1487–1541. Elsevier.

Bohrnstedt, G. and Goldberger, A. (1969). On the Exact Covariance of Products of Random Variables. *Journal of the American Statistical Association*, 64(328):1439–1442.

Breen, R. (2010). Educational expansion and social mobility in the 20th century. *Social Forces*, 89(2):365–388.

Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2022). *World inequality report 2022*. Harvard University Press.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., and York, R. L. (1966). *Equality of Educational Opportunity*. Government Printing Office, Washington, DC.

Dahan, M. and Gaviria, A. (2001). Sibling correlations and intergenerational mobility in latin america. *Economic Development and Cultural Change*, 49(3):537–54.

De Ferranti, D. M. (2004). *Inequality in Latin America: breaking with history?* world bank publications.

Deutscher, N. and Mazumder, B. (2022). Measuring Intergenerational Income Mobility: A Synthesis of Approaches. *Journal of Economic Literature*.

Emran, M. S., Jiang, H., and Shilpi, F. J. (2020). Gender Bias and Intergenerational Educational Mobility : Theory and Evidence from China and India. Policy Research Working Paper Series 9250, The World Bank.

Emran, M. S. and Shilpi, F. (2015). Gender, Geography, and Generations: Intergenerational Educational Mobility in Post-Reform India. *World Development*, 72:362–380.

Emran, M. S. and Shilpi, F. (2021). Economic approach to intergenerational mobility: Measures, methods, and challenges in developing countries. In Iversen, V., Krishna, A., and Sen, K., editors, *Social mobility in developing countries: Concepts, methods, and determinants*. Oxford University Press.

Emran, M. S. and Sun, Y. (2015). Magical transition? intergenerational educational and occupational mobility in rural China : 1988-2002. Policy Research Working Paper Series

7316, The World Bank.

Fan, Y., Yi, J., and Zhang, J. (2021). Rising intergenerational income persistence in china. *American Economic Journal: Economic Policy*, 13(1):202–30.

Grätz, M., Barclay, K. J., Wiborg, O. N., Lyngstad, T. H., Karhula, A., Erola, J., Präg, P., Laidley, T., and Conley, D. (2021). Sibling Similarity in Education Across and Within Societies. *Demography*, 58(3):1011–1037.

Heckman, J. and Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics*, 6(1):689–733.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2008). The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. *The B.E. Journal of Economic Analysis & Policy*, 7(2):1–48.

Iversen, V., Krishna, A., and Sen, K. (2019). Beyond poverty escapes – social mobility in developing countries: A review article. *World Bank Research Observer*, 34(2):239–273.

Jalan, J. and Murgai, R. (2008). Intergenerational Mobility in Education in India. Technical report.

Kishan, P. K. V. (2018). Is the Past Still Holding Us Back? A Study on Intergenerational Education Mobility in India. IIMA Working Papers WP 2018-01-03, Indian Institute of Management Ahmedabad, Research and Publication Department.

Kremer, M. and Maskin, E. (1996). Wage inequality and segregation by skill. Technical report, National bureau of economic research Cambridge, Mass., USA.

Kundu, A. and Sen, K. (2022). Multigenerational Mobility Among Males in India. *Review of Income and Wealth*.

Lleras-Muney, A., Tarozzi, A., Aurino, E., and Tinoco, B. (2022). The Rise and Fall of SES Gradients in Heights Around the World. *Available at SSRN 4127021*.

Lubotsky, D. and Wittenberg, M. (2006). Interpretation of Regressions with Multiple Proxies. *The Review of Economics and Statistics*, 88(3):549–562.

Maitra, P. and Sharma, A. (2010). Parents and children: Education across generations in India. Technical report.

Mazumder, B. (2008). Sibling similarities and economic inequality in the US. *Journal of Population Economics*, 21(3):685–701.

Mazumder, B., Rosales-Rueda, M., and Triyana, M. (2019). Intergenerational Human Capital Spillovers: Indonesia's School Construction and Its Effects on the Next Generation. *AEA Papers and Proceedings*, 109:243–249.

Mogstad, M. and Torsvik, G. (2021). Family Background, Neighborhoods and Intergenerational Mobility. NBER Working Papers 28874, National Bureau of Economic Research, Inc.

Narayan, A., der Weide, R. V., Cojocaru, A., Lakner, C., Redaelli, S., Mahler, D. G., Ramasubbaiah, R. G. N., and Thewissen, S. (2018). *Fair Progress? Economic Mobility across Generations Around the World*. Number 28428 in World Bank Publications. The World Bank.

Neidhofer, G., Serrano, J., and Gasparini, L. (2018). Educational inequality and intergenerational mobility in Latin America: A new database. *Journal of Development Economics*, 134(C):329–349.

Nicoletti, C. and Rabe, B. (2013). Inequality in pupils' test scores: How much do family, sibling type and neighbourhood matter? *Economica*, 80(318):197–218.

Prag, P., Akimova, E., and Monden, C. (2019). Intergenerational Transmission of Education: A Meta-Analysis of Sibling Correlations Published Between 1972-2018. Technical report, University of Oxford.

Razzu, G. and Wambile, A. (2022). Four Decades of Intergenerational Educational Mobility in Sub-Saharan Africa. *The Journal of Development Studies*, 58(5):931–950.

Roemer, J. and Trannoy, A. (2016). Equality of Opportunity: Theory and Measurement. *Journal of Economic Literature*, 54(4):1288–1332.

Roemer, J. E. (1998). *Equality of Opportunity*. Harvard University Press.

Solon, G. (1999). Intergenerational mobility in the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics*, pages 1761–1800. Elsevier.

Solon, G., Corcoran, M., Gordon, R., and Laren, D. (1988). Sibling and Intergenerational Correlations in Welfare Program Participation. *Journal of Human Resources*, 23(3):388–396.

Solon, G., Corcoran, M., Gordon, R., and Laren, D. (1991). A Longitudinal Analysis of Sibling

Correlations in Economic Status. *Journal of Human Resources*, 26(3):509–534.

Torche, F. (2015a). Intergenerational Mobility and Equality of Opportunity. *European Journal of Sociology*, 56(3):343–371.

Torche, F. (2015b). Intergenerational Mobility and Gender in Mexico. *Social Forces*, 94:5633–587.

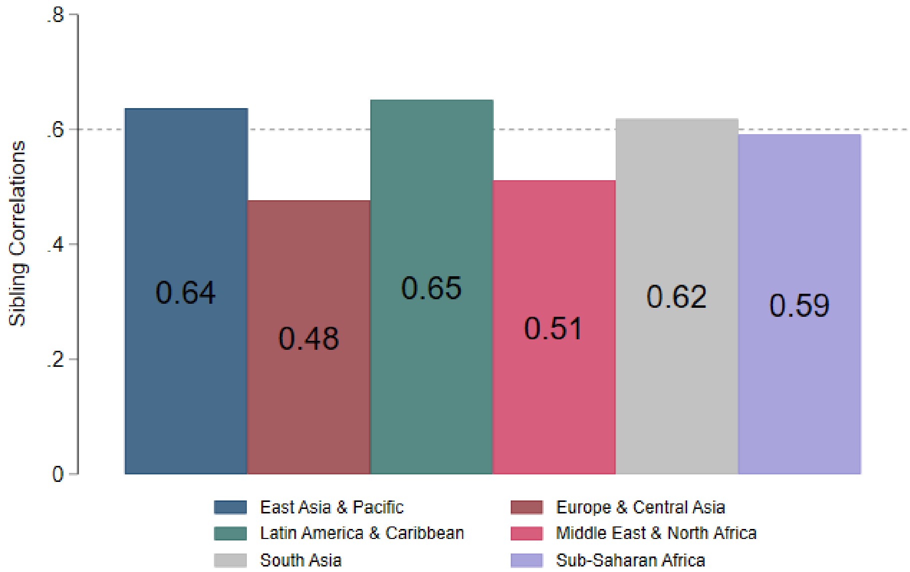
Torche, F. (2019). Educational mobility in developing countries. WIDER Working Paper Series 2019-88.

Vosters, K. and Nybom, M. (2017). Intergenerational Persistence in Latent Socioeconomic Status: Evidence from Sweden and the United States. *Journal of Labor Economics*, 35(3):869–901.

World Bank (2006). *World development report 2006: Equity and development*. The World Bank.

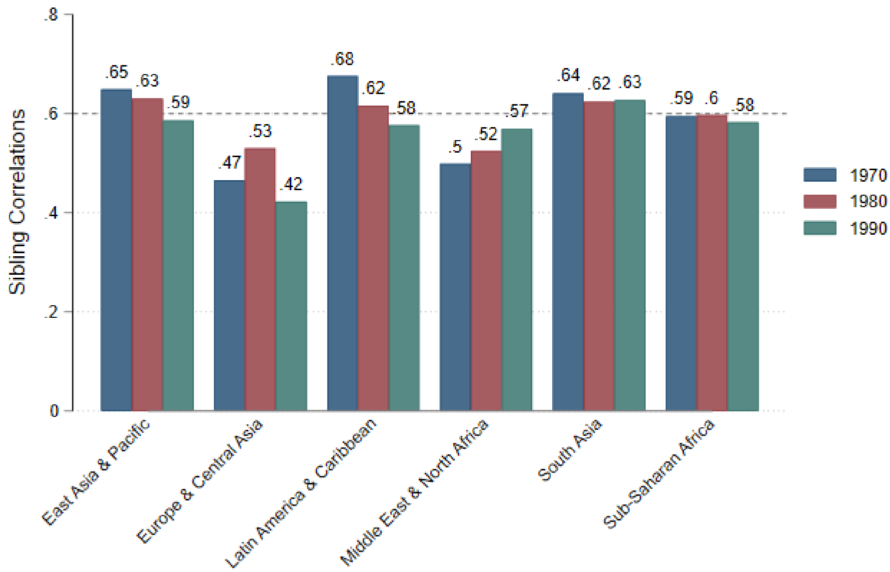
World Bank (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Number 28340 in World Bank Publications - Books. The World Bank Group.

Figure 1: Sibling Correlations by Regions



Notes: This figure presents the average sibling correlation estimates for six regions of the world using the full sample. Data come from 53 developing countries in the Demographic and Health Surveys (DHS). Specific countries included in each region are reported in Table 1. The dashed line represents the average sibling correlation estimates for all countries in our sample (0.60). For comparison purposes, the average sibling correlation for developed countries in the current literature is 0.41.

Figure 2: Sibling Correlations by Regions and Cohorts



Notes: This figure presents the average sibling correlation estimates for six regions of the world dis-aggregated by different decades of birth cohorts (the 1970s, 1980s, and 1990s). Data come from 53 developing countries in the Demographic and Health Surveys (DHS). Specific countries included in each region are reported in Table 1. The dashed line represents the average sibling correlation estimates for all countries in our sample (0.60). For comparison purposes, the average sibling correlation for developed countries in the current literature is 0.41.

Figure 3A: Proportion of Sibling Correlations Explained by Intergenerational Transmission by Regions (Bingley and Cappellari 2019 Method)

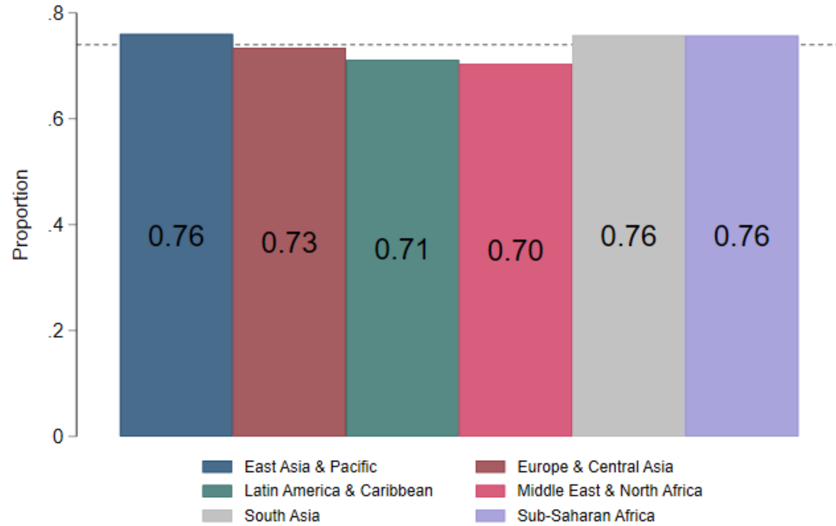
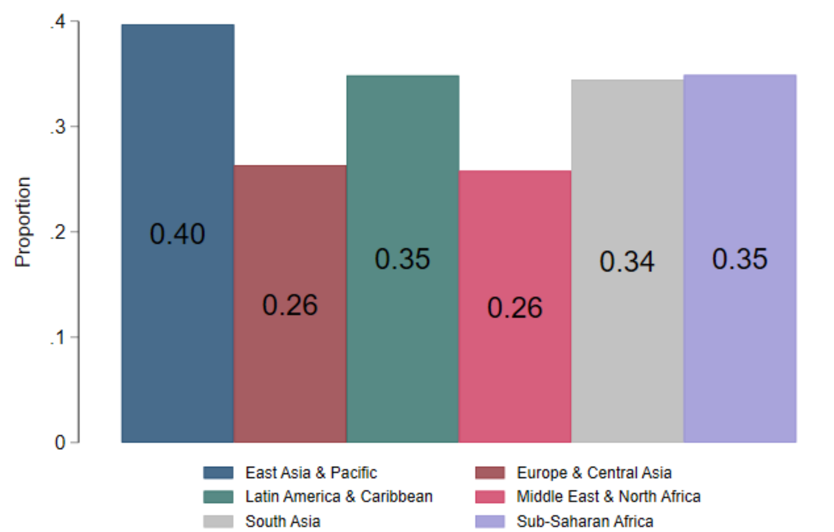
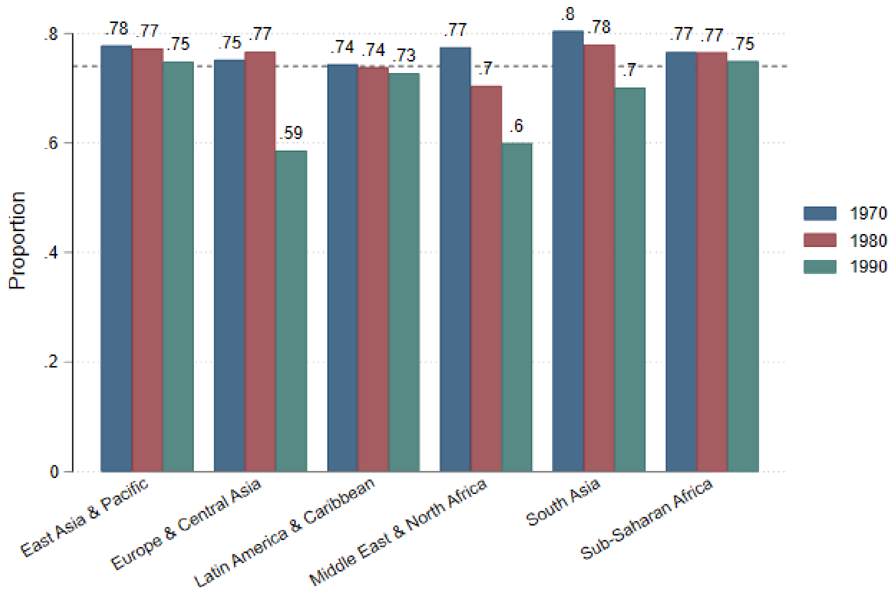


Figure 3B: Proportion of Sibling Correlations Explained by Intergenerational Transmission by Regions (Bjorklund et al. 2010 Method)



Notes: This figure presents the average estimated share of the intergenerational component for six regions of the world using the full sample. Data come from 53 developing countries in the Demographic and Health Surveys (DHS). Specific countries included in each region are reported in Table 1. Panel A uses the Bingley and Cappellari (2019) method, while Panel B uses the Bjorklund et al. (2010) method. The dashed line in Panel A plots the average estimated share of the intergenerational component for all countries using the full sample (0.74).

Figure 4: Proportion of Sibling Correlations Explained by Intergenerational Transmission by Regions and Cohorts (Bingley and Cappellari 2019 Method)



Notes: This figure presents the average estimated share of the intergenerational component using the Bingley and Cappellari (2019) method for six regions of the world dis-aggregated by different decades of birth cohorts (the 1970s, 1980s, and 1990s). Data come from 53 developing countries in the Demographic and Health Surveys (DHS). Specific countries included in each region are reported in Table 1. The dashed line plots the average estimated share of the intergenerational component for all countries using the full sample (0.74).

Table 1: Country-Specific Sibling Correlation Estimates (Full Sample)

Country	Sibling Corr.	S.E.	N	Country	Sibling Corr.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.662	0.006	10521	Benin	0.600	0.010	5480
Indonesia	0.625	0.004	33209	Burkina Faso	0.620	0.013	3498
Philippines	0.595	0.007	15064	Burundi	0.539	0.013	3897
Vietnam	0.664	0.011	3692	Cameroon	0.627	0.012	3663
				Chad	0.735	0.013	2710
Europe & Central Asia				Congo, Rep.	0.532	0.018	1781
Albania	0.471	0.020	2311	Cote d'Ivoire	0.584	0.018	1565
Armenia	0.546	0.018	3215	Ethiopia	0.696	0.007	7047
Kyrgyz Republic	0.379	0.031	1455	Gabon	0.549	0.023	1212
Tajikistan	0.439	0.020	3238	Ghana	0.588	0.014	3313
Turkey	0.547	0.009	8292	Guinea	0.552	0.016	2589
				Kenya	0.541	0.010	7559
Latin America & Caribbean				Lesotho	0.581	0.011	4382
Bolivia	0.678	0.009	6971	Liberia	0.532	0.019	1771
Brazil	0.698	0.012	2967	Madagascar	0.769	0.010	4376
Colombia	0.606	0.006	17607	Malawi	0.598	0.010	6164
Dominican Republic	0.566	0.009	8190	Mali	0.604	0.013	3910
Guatemala	0.713	0.007	6553	Mozambique	0.540	0.016	2910
Haiti	0.680	0.008	6022	Namibia	0.524	0.019	2581
Peru	0.617	0.004	34974	Niger	0.675	0.015	1505
				Nigeria	0.699	0.007	11380
Middle East & North Africa				Rwanda	0.536	0.010	6848
Egypt, Arab Rep.	0.542	0.005	27042	Senegal	0.598	0.012	3850
Jordan	0.481	0.008	17023	Sierra Leone	0.517	0.019	2209
				South Africa	0.490	0.020	2407
South Asia				Tanzania	0.528	0.012	5619
Afghanistan	0.563	0.009	7585	Togo	0.521	0.020	1615
Bangladesh	0.641	0.007	12031	Uganda	0.620	0.012	4176
India	0.629	0.002	151142	Zambia	0.643	0.010	5595
Nepal	0.623	0.010	5574	Zimbabwe	0.589	0.015	3369
Pakistan	0.633	0.003	40964				

Notes: This table presents the sibling correlation estimates for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the full sample.

Table 2: Country-Specific Sibling Correlation Estimates (By Cohorts)

Country	1970	1980	1990	Country	1970	1980	1990
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.674	0.657	0.685	Benin	0.653	0.595	0.579
Indonesia	0.659	0.616	0.523	Burkina Faso	0.584	0.638	0.506
Philippines	0.614	0.616	0.550	Burundi	n.a	0.584	0.500
Vietnam	0.682	0.616	n.a	Cameroon	0.608	0.629	0.629
				Chad	0.743	0.727	0.760
Europe & Central Asia				Congo, Rep.	0.466	0.532	0.625
Albania	n.a	0.502	0.389	Cote d'Ivoire	0.633	0.592	0.640
Armenia	0.468	0.592	0.492	Ethiopia	0.793	0.656	0.630
Kyrgyz Republic	0.379	0.438	0.342	Gabon	0.475	0.614	0.576
Tajikistan	n.a	0.461	0.424	Ghana	0.592	0.617	0.500
Turkey	0.550	0.561	0.433	Guinea	0.597	0.515	0.559
				Kenya	0.560	0.584	0.503
Latin America & Caribbean				Lesotho	0.592	0.592	0.483
Bolivia	0.692	0.653	n.a	Liberia	n.a	0.535	0.522
Brazil	0.709	n.a	n.a	Madagascar	0.824	0.790	n.a
Colombia	0.666	0.582	0.546	Malawi	0.609	0.602	0.585
Dominican Republic	0.595	0.538	0.435	Mali	0.626	0.619	0.573
Guatemala	0.750	0.697	0.680	Mozambique	0.526	0.522	0.681
Haiti	0.697	0.674	0.678	Namibia	0.554	0.535	0.516
Peru	0.668	0.591	0.546	Niger	0.670	0.684	0.645
				Nigeria	0.643	0.666	0.741
Middle East & North Africa				Rwanda	0.546	0.550	0.522
Egypt, Arab Rep.	0.523	0.581	0.588	Senegal	0.588	0.584	0.588
Jordan	0.474	0.468	0.551	Sierra Leone	n.a	0.517	0.515
				South Africa	0.521	0.383	0.501
South Asia				Tanzania	0.559	0.563	0.484
Afghanistan	n.a	0.523	0.591	Togo	0.446	0.570	0.519
Bangladesh	0.673	0.614	0.609	Uganda	0.648	0.666	0.552
India	0.658	0.631	0.631	Zambia	0.664	0.681	0.620
Nepal	0.638	0.618	0.588	Zimbabwe	0.569	0.609	0.625
Pakistan	0.595	0.633	0.679				

Notes: This table presents the sibling correlation estimates for each of the 53 developing countries in the Demographic and Health Surveys (DHS) dis-aggregated by different decades of birth cohorts (the 1970s, 1980s, and 1990s).

Table 3: Country-Specific Proportion of Sibling Correlations Explained by Intergenerational Transmission Estimates (Full Sample)

Country	Prop.	S.E.	N	Country	Prop.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.688	0.010	10521	Benin	0.824	0.018	5480
Indonesia	0.780	0.006	33209	Burkina Faso	0.720	0.021	3498
Philippines	0.812	0.011	15064	Burundi	0.660	0.024	3897
Vietnam	0.824	0.018	3692	Cameroon	0.848	0.021	3663
				Chad	0.721	0.018	2710
Europe & Central Asia				Congo, Rep.	0.853	0.039	1781
Albania	0.757	0.043	2311	Cote d'Ivoire	0.791	0.033	1565
Armenia	0.767	0.030	3215	Ethiopia	0.713	0.011	7047
Kyrgyz Republic	0.661	0.066	1455	Gabon	0.771	0.045	1212
Tajikistan	0.641	0.037	3238	Ghana	0.786	0.023	3313
Turkey	0.774	0.016	8292	Guinea	0.704	0.034	2589
				Kenya	0.768	0.020	7559
Latin America & Caribbean				Lesotho	0.717	0.021	4382
Bolivia	0.720	0.012	6971	Liberia	0.690	0.033	1771
Brazil	0.752	0.018	2967	Madagascar	0.845	0.014	4376
Colombia	0.729	0.009	17607	Malawi	0.752	0.016	6164
Dominican Republic	0.691	0.015	8190	Mali	0.745	0.021	3910
Guatemala	0.740	0.011	6553	Mozambique	0.823	0.032	2910
Haiti	0.640	0.014	6022	Namibia	0.749	0.034	2581
Peru	0.755	0.006	34974	Niger	0.704	0.027	1505
				Nigeria	0.704	0.009	11380
Middle East & North Africa				Rwanda	0.735	0.021	6848
Egypt, Arab Rep.	0.740	0.010	27042	Senegal	0.802	0.024	3850
Jordan	0.667	0.015	17023	Sierra Leone	0.668	0.030	2209
				South Africa	0.786	0.040	2407
South Asia				Tanzania	0.659	0.021	5619
Afghanistan	0.618	0.016	7585	Togo	0.825	0.043	1615
Bangladesh	0.820	0.010	12031	Uganda	0.735	0.020	4176
India	0.744	0.003	151142	Zambia	0.746	0.016	5595
Nepal	0.658	0.015	5574	Zimbabwe	0.707	0.024	3369
Pakistan	0.808	0.006	40964				

Notes: This table presents the estimated share of the intergenerational component using the Bingley and Cappellari (2019) method for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the full sample.

Table 4: Country-Specific Proportion of Sibling Correlations Explained by Intergenerational Transmission Estimates (By Cohorts)

Country	1970	1980	1990	Country	1970	1980	1990
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.731	0.696	0.659	Benin	0.867	0.840	0.809
Indonesia	0.787	0.791	0.774	Burkina Faso	0.844	0.688	0.727
Philippines	0.816	0.830	0.810	Burundi	n.a	0.740	0.639
Vietnam	0.829	0.867	n.a	Cameroon	0.742	0.841	0.871
				Chad	0.735	0.753	0.677
Europe & Central Asia				Congo, Rep.	0.926	0.863	0.728
Albania	n.a	0.858	0.762	Cote d'Ivoire	0.737	0.845	0.732
Armenia	0.940	0.746	0.555	Ethiopia	0.749	0.781	0.679
Kyrgyz Republic	0.496	0.786	0.501	Gabon	0.890	0.672	0.775
Tajikistan	n.a	0.700	0.587	Ghana	0.853	0.781	0.796
Turkey	0.819	0.767	0.701	Guinea	0.452	0.787	0.753
				Kenya	0.656	0.746	0.877
Latin America & Caribbean				Lesotho	0.716	0.727	0.817
Bolivia	0.761	0.709	n.a	Liberia	n.a	0.680	0.699
Brazil	0.758	n.a	n.a	Madagascar	0.859	0.816	n.a
Colombia	0.778	0.755	0.706	Malawi	0.764	0.767	0.745
Dominican Republic	0.701	0.727	0.838	Mali	0.855	0.745	0.699
Guatemala	0.759	0.796	0.722	Mozambique	0.914	0.765	0.759
Haiti	0.709	0.644	0.629	Namibia	0.697	0.760	0.823
Peru	0.771	0.771	0.738	Niger	0.767	0.704	0.705
				Nigeria	0.707	0.700	0.693
Middle East & North Africa				Rwanda	0.727	0.753	0.715
Egypt, Arab Rep.	0.781	0.746	0.656	Senegal	0.820	0.810	0.792
Jordan	0.769	0.662	0.543	Sierra Leone	n.a	0.764	0.597
				South Africa	0.877	0.836	0.592
South Asia				Tanzania	0.581	0.672	0.711
Afghanistan	n.a	0.725	0.561	Togo	0.832	0.839	0.837
Bangladesh	0.854	0.861	0.718	Uganda	0.784	0.755	0.704
India	0.860	0.759	0.686	Zambia	0.726	0.777	0.741
Nepal	0.680	0.684	0.640	Zimbabwe	0.688	0.687	0.721
Pakistan	0.823	0.814	0.762				

Notes: This table presents the estimated share of the intergenerational component using the Bingley and Cappellari (2019) method for each of the 53 developing countries in the Demographic and Health Surveys (DHS) disaggregated by different decades of birth cohorts (the 1970s, 1980s, and 1990s).

Table A1: Countries from the Demographic and Health Surveys (DHS)

Region	Country	Used	Region	Country	Used
East Asia & Pacific	Cambodia	Yes	Sub-Saharan Africa	Angola	No
	Indonesia	Yes		Benin	Yes
	Lao PDR	No		Botswana	No
	Myanmar	No		Burkina Faso	Yes
	Papua New Guinea	No		Burundi	Yes
	Philippines	Yes		Cameroon	Yes
	Samoa	No		Cape Verde	No
	Thailand	No		Central African Republic	No
	Timor-Leste	No		Chad	Yes
	Vietnam	Yes		Comoros	No
Europe & Central Asia	Albania	Yes		Congo	Yes
	Armenia	Yes		Congo Democratic Republic	No
	Azerbaijan	No		Cote d'Ivoire	Yes
	Georgia	No		Equatorial Guinea	No
	Kazakhstan	No		Eritrea	No
	Kyrgyz Republic	Yes		Eswatini	No
	Moldova	No		Ethiopia	Yes
	Romania	No		Gabon	Yes
	Tajikistan	Yes		Gambia	No
	Turkey	Yes		Ghana	Yes
	Turkmenistan	No		Guinea	Yes
	Ukraine	No		Kenya	Yes
	Uzbekistan	No		Lesotho	Yes
Latin America & Caribbean	Bolivia	Yes		Liberia	Yes
	Brazil	Yes		Madagascar	Yes
	Colombia	Yes		Malawi	Yes
	Dominican Republic	Yes		Mali	Yes
	Ecuador	No		Mauritania	No
	El Salvador	No		Mozambique	Yes
	Guatemala	Yes		Namibia	Yes
	Guyana	No		Niger	Yes
	Haiti	Yes		Nigeria	Yes
	Honduras	No		Nigeria (Ondo State)	No
	Jamaica	No		Rwanda	Yes
	Mexico	No		Sao Tome and Principe	No
	Nicaragua	No		Senegal	Yes
	Paraguay	No		Sierra Leone	Yes
	Peru	Yes		South Africa	Yes
	Trinidad and Tobago	No		Sudan	No
Middle East & North Africa	Egypt	Yes		Tanzania	Yes
	Jordan	Yes		Togo	Yes
	Morocco	No		Uganda	Yes
	Tunisia	No		Zambia	Yes
	Yemen	No		Zimbabwe	Yes
South Asia	Afghanistan	Yes			
	Bangladesh	Yes			
	India	Yes			
	Maldives	No			
	Nepal	Yes			
	Pakistan	Yes			
	Sri Lanka	No			

Notes: Data come from the Demographic and Health Surveys (DHS). 53 countries are used and accessed between April 2021 and July 2021. 42 countries are not used in the analytic sample where at least one DHS survey is available but the sample size is small. The total number of observations used in the analytic sample is 544624.

Table A2: Country-Specific Sibling Correlation Estimates (1970 Cohort)

Country	Sibling Corr.	S.E.	N	Country	Sibling Corr.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.674	0.013	2271	Benin	0.653	0.020	1028
Indonesia	0.659	0.006	12924	Burkina Faso	0.584	0.022	1314
Philippines	0.614	0.012	4435	Burundi	n.a	n.a	n.a
Vietnam	0.682	0.013	2286	Cameroon	0.608	0.037	399
				Chad	0.743	0.024	658
Europe & Central Asia				Congo, Rep.	0.466	0.050	275
Albania	n.a	n.a	n.a	Cote d'Ivoire	0.633	0.030	507
Armenia	0.468	0.034	1011	Ethiopia	0.793	0.010	2106
Kyrgyz Republic	0.379	0.057	468	Gabon	0.475	0.037	478
Tajikistan	n.a	n.a	n.a	Ghana	0.592	0.025	980
Turkey	0.550	0.013	3375	Guinea	0.597	0.033	635
				Kenya	0.560	0.022	1904
Latin America & Caribbean				Lesotho	0.592	0.024	787
Bolivia	0.692	0.011	3329	Liberia	n.a	n.a	n.a
Brazil	0.709	0.014	1843	Madagascar	0.824	0.011	1658
Colombia	0.666	0.011	4009	Malawi	0.609	0.020	1285
Dominican Republic	0.595	0.013	3097	Mali	0.626	0.027	842
Guatemala	0.750	0.011	2117	Mozambique	0.526	0.025	1347
Haiti	0.697	0.015	1272	Namibia	0.554	0.028	846
Peru	0.668	0.007	10144	Niger	0.670	0.053	158
				Nigeria	0.643	0.040	404
Middle East & North Africa				Rwanda	0.546	0.021	1521
Egypt, Arab Rep.	0.523	0.009	9541	Senegal	0.588	0.026	838
Jordan	0.474	0.014	4118	Sierra Leone	n.a	n.a	n.a
				South Africa	0.521	0.025	1325
South Asia				Tanzania	0.559	0.023	1617
Afghanistan	n.a	n.a	n.a	Togo	0.446	0.029	1130
Bangladesh	0.673	0.009	4275	Uganda	0.648	0.023	876
India	0.658	0.006	13288	Zambia	0.664	0.016	1710
Nepal	0.638	0.016	1816	Zimbabwe	0.569	0.026	1150
Pakistan	0.595	0.009	6202				

Notes: This table presents the sibling correlation estimates for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the 1970s birth cohort sample.

Table A3: Country-Specific Sibling Correlation Estimates (1980 Cohort)

Country	Sibling Corr.	S.E.	N	Country	Sibling Corr.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.657	0.009	6402	Benin	0.595	0.016	2439
Indonesia	0.616	0.007	10148	Burkina Faso	0.638	0.017	1636
Philippines	0.616	0.011	4746	Burundi	0.584	0.021	1396
Vietnam	0.616	0.021	1232	Cameroon	0.629	0.017	1755
				Chad	0.727	0.020	887
Europe & Central Asia				Congo, Rep.	0.532	0.022	1212
Albania	0.502	0.032	1106	Cote d'Ivoire	0.592	0.029	620
Armenia	0.592	0.022	1735	Ethiopia	0.656	0.011	3006
Kyrgyz Republic	0.438	0.036	575	Gabon	0.614	0.034	485
Tajikistan	0.461	0.023	1469	Ghana	0.617	0.020	1588
Turkey	0.561	0.015	3265	Guinea	0.515	0.028	983
				Kenya	0.584	0.017	2587
Latin America & Caribbean				Lesotho	0.592	0.013	3001
Bolivia	0.653	0.015	3118	Liberia	0.535	0.027	787
Brazil	n.a	n.a	n.a	Madagascar	0.790	0.010	2084
Colombia	0.582	0.009	9258	Malawi	0.602	0.015	2470
Dominican Republic	0.538	0.014	4279	Mali	0.619	0.019	1595
Guatemala	0.697	0.016	1281	Mozambique	0.522	0.024	1176
Haiti	0.674	0.014	2594	Namibia	0.535	0.029	1072
Peru	0.591	0.006	19330	Niger	0.684	0.019	1031
				Nigeria	0.666	0.011	4833
Middle East & North Africa				Rwanda	0.550	0.015	3381
Egypt, Arab Rep.	0.581	0.008	11306	Senegal	0.584	0.016	2157
Jordan	0.468	0.012	7316	Sierra Leone	0.517	0.026	1029
				South Africa	0.383	0.050	343
South Asia				Tanzania	0.563	0.018	1952
Afghanistan	0.523	0.015	2727	Togo	0.570	0.039	430
Bangladesh	0.614	0.010	5228	Uganda	0.666	0.017	1489
India	0.631	0.003	53132	Zambia	0.681	0.018	1325
Nepal	0.618	0.016	2171	Zimbabwe	0.609	0.021	1481
Pakistan	0.633	0.004	27858				

Notes: This table presents the sibling correlation estimates for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the 1980s birth cohort sample.

Table A4: Country-Specific Sibling Correlation Estimates (1990 Cohort)

Country	Sibling Corr.	S.E.	N	Country	Sibling Corr.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.685	0.016	1751	Benin	0.579	0.019	1915
Indonesia	0.523	0.014	4540	Burkina Faso	0.506	0.055	241
Philippines	0.550	0.014	4301	Burundi	0.500	0.018	2484
Vietnam	n.a	n.a	n.a	Cameroon	0.629	0.021	1502
				Chad	0.760	0.018	1058
Europe & Central Asia				Congo, Rep.	0.625	0.043	291
Albania	0.389	0.036	1115	Cote d'Ivoire	0.640	0.057	154
Armenia	0.492	0.065	428	Ethiopia	0.630	0.016	1935
Kyrgyz Republic	0.342	0.094	304	Gabon	0.576	0.056	249
Tajikistan	0.424	0.030	1744	Ghana	0.500	0.036	577
Turkey	0.433	0.038	742	Guinea	0.559	0.026	949
				Kenya	0.503	0.019	2568
Latin America & Caribbean				Lesotho	0.483	0.036	593
Bolivia	n.a	n.a	n.a	Liberia	0.522	0.026	909
Brazil	n.a	n.a	n.a	Madagascar	n.a	n.a	n.a
Colombia	0.546	0.014	3729	Malawi	0.585	0.017	2178
Dominican Republic	0.435	0.044	540	Mali	0.573	0.022	1332
Guatemala	0.680	0.012	2690	Mozambique	0.681	0.041	279
Haiti	0.678	0.013	2144	Namibia	0.516	0.063	279
Peru	0.546	0.019	2645	Niger	0.645	0.041	308
				Nigeria	0.741	0.010	5396
Middle East & North Africa				Rwanda	0.522	0.023	1362
Egypt, Arab Rep.	0.588	0.018	2934	Senegal	0.588	0.040	354
Jordan	0.551	0.015	5047	Sierra Leone	0.515	0.026	1139
				South Africa	0.501	0.043	578
South Asia				Tanzania	0.484	0.030	1213
Afghanistan	0.591	0.011	4857	Togo	0.519	0.038	410
Bangladesh	0.609	0.018	1712	Uganda	0.552	0.021	1710
India	0.631	0.003	72192	Zambia	0.620	0.018	2011
Nepal	0.588	0.024	1317	Zimbabwe	0.625	0.037	518
Pakistan	0.679	0.009	5583				

Notes: This table presents the sibling correlation estimates for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the 1990s birth cohort sample.

Table A5: Country-Specific Proportion of Sibling Correlations Explained by Intergenerational Transmission Estimates (1970 Cohort)

Country	Prop.	S.E.	N	Country	Prop.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.731	0.024	2271	Benin	0.867	0.037	1028
Indonesia	0.787	0.009	12924	Burkina Faso	0.844	0.043	1314
Philippines	0.816	0.021	4435	Burundi	n.a	n.a	n.a
Vietnam	0.829	0.022	2286	Cameroon	0.742	0.073	399
				Chad	0.735	0.039	658
Europe & Central Asia				Congo, Rep.	0.926	0.134	275
Albania	n.a	n.a	n.a	Cote d'Ivoire	0.737	0.050	507
Armenia	0.940	0.076	1011	Ethiopia	0.749	0.017	2106
Kyrgyz Republic	0.496	0.112	468	Gabon	0.890	0.096	478
Tajikistan	n.a	n.a	n.a	Ghana	0.853	0.046	980
Turkey	0.819	0.026	3375	Guinea	0.452	0.060	635
				Kenya	0.656	0.036	1904
Latin America & Caribbean				Lesotho	0.716	0.048	787
Bolivia	0.761	0.017	3329	Liberia	n.a	n.a	n.a
Brazil	0.758	0.022	1843	Madagascar	0.859	0.017	1658
Colombia	0.778	0.017	4009	Malawi	0.764	0.037	1285
Dominican Republic	0.701	0.023	3097	Mali	0.855	0.053	842
Guatemala	0.759	0.017	2117	Mozambique	0.914	0.056	1347
Haiti	0.709	0.028	1272	Namibia	0.697	0.052	846
Peru	0.771	0.011	10144	Niger	0.767	0.093	158
				Nigeria	0.707	0.054	404
Middle East & North Africa				Rwanda	0.727	0.043	1521
Egypt, Arab Rep.	0.781	0.017	9541	Senegal	0.820	0.054	838
Jordan	0.769	0.032	4118	Sierra Leone	n.a	n.a	n.a
				South Africa	0.877	0.054	1325
South Asia				Tanzania	0.581	0.034	1617
Afghanistan	n.a	n.a	n.a	Togo	0.832	0.069	1130
Bangladesh	0.854	0.015	4275	Uganda	0.784	0.041	876
India	0.860	0.009	13288	Zambia	0.726	0.028	1710
Nepal	0.680	0.024	1816	Zimbabwe	0.688	0.046	1150
Pakistan	0.823	0.017	6202				

Notes: This table presents the estimated share of the intergenerational component using the Bingley and Cappellari (2019) method for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the 1970s birth cohort sample.

Table A6: Country-Specific Proportion of Sibling Correlations Explained by Intergenerational Transmission Estimates (1980 Cohort)

Country	Prop.	S.E.	N	Country	Prop.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.696	0.014	6402	Benin	0.840	0.028	2439
Indonesia	0.791	0.012	10148	Burkina Faso	0.688	0.029	1636
Philippines	0.830	0.019	4746	Burundi	0.740	0.041	1396
Vietnam	0.867	0.036	1232	Cameroon	0.841	0.028	1755
				Chad	0.753	0.033	887
Europe & Central Asia				Congo, Rep.	0.863	0.048	1212
Albania	0.858	0.065	1106	Cote d'Ivoire	0.845	0.055	620
Armenia	0.746	0.035	1735	Ethiopia	0.781	0.021	3006
Kyrgyz Republic	0.786	0.086	575	Gabon	0.672	0.057	485
Tajikistan	0.700	0.055	1469	Ghana	0.781	0.032	1588
Turkey	0.767	0.024	3265	Guinea	0.787	0.067	983
				Kenya	0.746	0.030	2587
Latin America & Caribbean				Lesotho	0.727	0.025	3001
Bolivia	0.709	0.020	3118	Liberia	0.680	0.050	787
Brazil	n.a	n.a	n.a	Madagascar	0.816	0.015	2084
Colombia	0.755	0.015	9258	Malawi	0.767	0.026	2470
Dominican Republic	0.727	0.023	4279	Mali	0.745	0.032	1595
Guatemala	0.796	0.026	1281	Mozambique	0.765	0.047	1176
Haiti	0.644	0.021	2594	Namibia	0.760	0.053	1072
Peru	0.771	0.010	19330	Niger	0.704	0.032	1031
				Nigeria	0.700	0.016	4833
Middle East & North Africa				Rwanda	0.753	0.032	3381
Egypt, Arab Rep.	0.746	0.013	11306	Senegal	0.810	0.033	2157
Jordan	0.662	0.025	7316	Sierra Leone	0.764	0.049	1029
				South Africa	0.836	0.136	343
South Asia				Tanzania	0.672	0.032	1952
Afghanistan	0.725	0.030	2727	Togo	0.839	0.070	430
Bangladesh	0.861	0.017	5228	Uganda	0.755	0.029	1489
India	0.759	0.005	53132	Zambia	0.777	0.029	1325
Nepal	0.684	0.023	2171	Zimbabwe	0.687	0.034	1481
Pakistan	0.814	0.007	27858				

Notes: This table presents the estimated share of the intergenerational component using the Bingley and Cappellari (2019) method for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the 1980s birth cohort sample.

Table A7: Country-Specific Proportion of Sibling Correlations Explained Explained by Intergenerational Transmission Estimates (1990 Cohort)

Country	Prop.	S.E.	N	Country	Prop.	S.E.	N
East Asia & Pacific				Sub-Saharan Africa			
Cambodia	0.659	0.020	1751	Benin	0.809	0.031	1915
Indonesia	0.774	0.023	4540	Burkina Faso	0.727	0.095	241
Philippines	0.810	0.026	4301	Burundi	0.639	0.033	2484
Vietnam	n.a	n.a	n.a	Cameroon	0.871	0.032	1502
				Chad	0.677	0.024	1058
Europe & Central Asia				Congo, Rep.	0.728	0.061	291
Albania	0.762	0.080	1115	Cote d'Ivoire	0.732	0.078	154
Armenia	0.555	0.084	428	Ethiopia	0.679	0.023	1935
Kyrgyz Republic	0.501	0.159	304	Gabon	0.775	0.095	249
Tajikistan	0.587	0.053	1744	Ghana	0.796	0.066	577
Turkey	0.701	0.075	742	Guinea	0.753	0.062	949
				Kenya	0.877	0.041	2568
Latin America & Caribbean				Lesotho	0.817	0.070	593
Bolivia	n.a	n.a	n.a	Liberia	0.699	0.048	909
Brazil	n.a	n.a	n.a	Madagascar	n.a	n.a	n.a
Colombia	0.706	0.022	3729	Malawi	0.745	0.027	2178
Dominican Republic	0.838	0.093	540	Mali	0.699	0.034	1332
Guatemala	0.722	0.019	2690	Mozambique	0.759	0.055	279
Haiti	0.629	0.021	2144	Namibia	0.823	0.115	279
Peru	0.738	0.028	2645	Niger	0.705	0.056	308
				Nigeria	0.693	0.011	5396
Middle East & North Africa				Rwanda	0.715	0.043	1362
Egypt, Arab Rep.	0.656	0.025	2934	Senegal	0.792	0.069	354
Jordan	0.543	0.023	5047	Sierra Leone	0.597	0.039	1139
				South Africa	0.592	0.072	578
South Asia				Tanzania	0.711	0.052	1213
Afghanistan	0.561	0.017	4857	Togo	0.837	0.073	410
Bangladesh	0.718	0.025	1712	Uganda	0.704	0.037	1710
India	0.686	0.004	72192	Zambia	0.741	0.027	2011
Nepal	0.640	0.032	1317	Zimbabwe	0.721	0.054	518
Pakistan	0.762	0.013	5583				

Notes: This table presents the estimated share of the intergenerational component using the Bingley and Cappellari (2019) method for each of the 53 developing countries in the Demographic and Health Surveys (DHS) using the 1990s birth cohort sample.