

# Inequality of Opportunity in Education

Accounting for the Contributions of Sibs,  
Schools and Sorting across East Africa

*Paul Anand*

*Jere R. Behrman*

*Hai-Anh H. Dang*

*Sam Jones*



**WORLD BANK GROUP**

Development Economics  
Development Research Group  
October 2018

## Abstract

Inequalities in the opportunity to obtain a good education in low-income countries are widely understood to be related to household resources and schooling quality. Yet, to date, most researchers have investigated the contributions of these two factors separately. This paper considers them jointly, paying special attention to their covariation, which indicates whether schools exacerbate or compensate for existing household-based inequalities. The paper develops a new variance decomposition framework and applies it to data on more than one million children in three low-income East African countries. The empirical

results show that although household factors account for a significant share of total test score variation, variation in school quality and positive sorting between households and schools are, together, no less important. The analysis also finds evidence of substantial geographical heterogeneity in schooling quality. The paper concludes that promoting equity in education in East Africa requires policies that go beyond raising average school quality and should attend to the distribution of school quality as well as assortative matching between households and schools.

---

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/research>. The authors may be contacted at [hdang@worldbank.org](mailto:hdang@worldbank.org) and [jones@wider.unu.edu](mailto:jones@wider.unu.edu).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# Inequality of Opportunity in Education: Accounting for the Contributions of Sibs, Schools and Sorting across East Africa

Paul Anand      Jere R. Behrman      Hai-Anh H. Dang      Sam Jones\*

**JEL:** D6, H0, I2, O1

**Key words:** inequality of opportunity, education achievement, decomposition, household, school, sorting, Africa

---

\*Jones ([jones@wider.unu.edu](mailto:jones@wider.unu.edu); corresponding author) is a Research Fellow with UNU-WIDER; Dang ([hdang@worldbank.org](mailto:hdang@worldbank.org); corresponding author) is an economist in the Survey Unit, Development Data Group, World Bank, and a non-resident senior research fellow with Vietnam's Academy of Social Sciences; Behrman is the William R. Kenan, Jr. Professor of Economics at the University of Pennsylvania; Paul Anand is a Professor at the Open University and Research Associate in the Department of Social Policy and Intervention in Oxford University and the London School of Economics. We would like to thank Jed Friedman, Ha Nguyen, and Ron Smith for helpful comments on earlier versions. We are grateful to the UK Department of International Development for funding assistance through its Strategic Research Program (SRP) program.

## 1 Introduction

The inability of some children in low-income countries to access quality schooling is a matter of concern, both for economic efficiency and social justice. If able children do not achieve their educational potential, countries face potentially significant losses against the counterfactual where all have an equal opportunity to develop their talents and skills. Likewise, there are good reasons based on social justice to ensure that development is as equitable a process as it can be (Sen, 2002). In a widely cited overview, Corak (2013) shows that interactions between family circumstances, labor markets and public policies ‘all structure a child’s opportunities’ and concludes there is a need to promote policies that promote children’s human capital in a way that offers relatively greater benefits to the relatively disadvantaged. Though focused mainly on higher-income countries, there is little reason to think these conclusions might not apply to lower-income countries. Indeed, international comparisons of educational achievement highlight large gaps between students from richer and poorer countries as well as substantial within-country gaps in both grade attainment and learning outcomes (e.g., Dabalen, 2015; Sandefur, 2018).

Following Roemer (1996), promotion of equity in education has typically focused on tackling inequalities that can be traced to differences in opportunities, defined as the circumstances that lie beyond the control of individual children, rather than those due to effort or personal choice.<sup>1</sup> Existing research in this domain has primarily operationalized inequality of opportunity in education (IOE) as the (proportional) contribution of the home environment to inequalities in educational outcomes. Studies from a range of higher income countries suggest that IOE is surprisingly high, with upwards of 40% of variation in schooling outcomes being associated with given household circumstances (Björklund and Salvanes, 2011). A more limited number of studies for developing countries also indicate that differences in family circumstances account for a material share of differences in both grade attainment and achievement (Ferreira et al., 2008; Ferreira and Gignoux, 2014).

An important limitation of previous studies is that they mostly focus on capturing IOE via a single factor, such as home circumstances. However, family circumstances may be only one of several factors that can be considered part of the ‘given’ circumstances that children face. Access to quality schooling is a no less important factor that determines children’s educational achievement and yet, in general, this factor is also beyond the control of individual children and many (disadvantaged) households. Schools in developing countries

---

<sup>1</sup> See Roemer (1996, 2002), who notes that equality of opportunity is the most universally supported conception of justice in advanced societies. Within Sen’s (1985) framework it is possible to view household factors as not just having a direct effect on educational achievement but also as both helping a child access outside educational resources as well as conditioning the child’s ability to convert school quality into scholastic outcomes.

are often ill-equipped and some do not even meet their basic functionalities, with teachers sometimes not coming to school to teach (Glewwe and Muralidharan, 2016). In addition, even in richer countries, other characteristics related to the oftentimes exogenous organization and financing of school systems (e.g., use of ability grouping) have been shown to influence the magnitude of inequalities in final achievement (Rivkin et al., 2005; Hanushek and Rivkin, 2006).

In this paper, therefore, we seek to provide a more comprehensive understanding of the magnitude and sources of inequality of opportunity in education. To do so, we develop a framework that jointly accounts for the contributions of both school and household factors, as well as their covariance, to variation in learning outcomes. The framework is then applied to a rich micro-data set for over 1 million children from three East African countries (Kenya, Tanzania and Uganda). As such, the paper makes three contributions to the literature on inequality of opportunity in education (IOE). Firstly, taking as a point of departure the idea that IOE may not only be attributed to family circumstances in a developing country context, we add to the literature by quantifying the distinct contributions of *both* households and schools to variation in learning outcomes. To our knowledge, this is the first paper to do so in a developing country context. In any case, the existing literature offers very few studies that examine both of these factors, particularly for several countries at the same time as attempted here; and, of the few that do, most focus on richer countries.<sup>2</sup>

Secondly, we advance the field by investigating the interactions (covariance) of school and household effects. In a purely theoretical IOE setting, these two types of effects may exist independently of each other and are exogenously given to the household. But in reality, even in a lower-income country context, some (richer) households may be able to select (better) schools through their (stronger) resources and social connections. Put differently, there is likely to be some sorting (assortative matching) between households and schools. Identifying the magnitude of this sorting effect can help policy makers reduce inequalities, for example, by setting school zoning or mobility policies that ensure similar chances of access to high-quality school for all households, including the most disadvantaged.

Thirdly, we build on the existing variance-decomposition framework in the established IOE literature. An unresolved challenge in the present setting, where more than one unobserved factor is considered, concerns how to specify the between-factor covariance. Indeed, the ‘true’ nature of this covariance typically cannot be identified unambiguously from the underlying data; and different empirical approaches effectively adopt contrasting *a priori* assumptions regarding how the covariance is allocated across the factors.<sup>3</sup> We propose a new

---

<sup>2</sup> We return to a more detailed discussion of the literature in the next section.

<sup>3</sup> See, e.g., Gibbons et al. (2014) for a related study in the field of urban economics.

empirical procedure that not only permits simultaneous estimation of two unobserved fixed latent factors (household and school effects), but also allows alternative assumptions regarding their covariance to be handled in a transparent fashion. This provides (strict) bounds on the variance contributions of interest. Furthermore, in our empirical implementation, we validate (cross-test) our results showing that extreme assumptions about the between-factor covariance can be ruled out.

The rest of this paper consists of four sections. Section 2 sets out a general framework to account for IOE, which also provides a window on findings of previous studies. It also describes the empirical strategy we use to identify the household and school factors, treating them both as unobserved effects, and proposes a simple mechanism that offers a unified variance decomposition strategy, covering the full range of alternative assumptions regarding the between-factor covariance.

Section 3 applies the proposed approach to extensive test score data on over 1,000,000 school-aged children from three East African countries (Kenya, mainland Tanzania and Uganda). It compares variance-decomposition results from three choices of the initialization parameter, each of which corresponds to an intuitive characterization of the between-factor covariance. Using both unconditional and conditional models, we show that household and school circumstances *jointly account* for nearly half of the variance in test scores (normalized by age). However, confirming the limitations of past studies, households cannot be considered the primary or only source of IOE.

More specifically, we find the upper-bound variance share attributable to schools is generally as large as the upper bound attributable to households, which in itself is indicative of a positive association between the latent factors. And under our preferred model specification, we find evidence of substantial positive sorting whereby higher ‘quality’ households tend to be matched to higher quality schools. In Section 4 we validate the main findings using alternative estimation procedures and investigate heterogeneity in the variance decomposition. We find systematic patterns in the level of inequality and the magnitudes of the variance components across different sub-groups and geographical locations, including a larger contribution of sorting in more disadvantaged locations. Section 5 concludes and reflects on our findings.

## 2 Analytical framework

### 2.1 Accounting model

A general framework for the analysis of inequality in education splits the proposed process

generating (differences in) test scores into the effect of given circumstances (opportunities) and the effect of other factors (idiosyncratic effects, effort, preferences, etc.). This suggests an educational production function of the following form:

$$t_{ijk} = f(h_j, s_k) + e_{ijk} \quad (1)$$

Where  $t$  is a measure of educational achievement (e.g., test scores), and indexes  $i = (1, 2, \dots, N)$ ,  $j = (1, 2, \dots, H)$  and  $k = (1, 2, \dots, S)$  refer to individual children, families and schools respectively. Thus,  $f(\cdot)$  captures the contribution of given household and school circumstances, and  $e$  captures remaining individual variation that is treated as orthogonal to the former; i.e.,  $E(e_{ijk} | s_k, h_j) = 0$ .<sup>4</sup> Following our concern to parse out the respective contributions of households and schools to IOE,  $h_j$  is defined as a comprehensive metric of the contribution of all factors shared by children in the same household (hereafter sibs); and  $s_k$  is defined as a comprehensive metric of the contribution of the given school (and grade) to their learning.<sup>5</sup> Note that under the assumption that not all sibs attend the same school and/or grade, the household factor does not nest the school factor; that is, they are crossed.<sup>6</sup>

Equation (1) defines test score levels, from which various metrics of inequality have been proposed. Within the domain of education, the simple variance of outcomes is widely used. As Ferreira and Gignoux (2014) note, the variance is ordinally invariant to standardization procedures often used to express test scores on a comparable scale. Also, the linear additive property of the variance makes it straightforward to isolate the contributions of individual components to the overall variance. However, even if we assume a linear form for  $f$ , the test score variance can be defined in various ways, depending on what assumptions are made about the relations between its constitutive elements. Table 1 describes four main cases of the relationship between  $h$  and  $s$ .<sup>7</sup> Each row sets out an assumed underlying data-generating process (model), which incorporates specific assumptions about the level and variance of  $t$ . The model in the first row assumes the household and school factors make independent,

<sup>4</sup> This last assumption may seem strong, but facilitates our primary interest in identifying the relevant contributions of latent household and school factors. Without additional individual-level controls (see below) individual characteristics (e.g., ability) that are correlated with  $h$  and  $s$  will be absorbed by the latter factors. Even so, there are individual factors that are expected to be orthogonal to the estimated terms. For example, Behrman et al. (1994) use monozygotic (identical) and dizygotic (fraternal) twins in the United States to identify individual-specific factors (orthogonal to the household factors) and find these account for about a quarter of the variance in adult earnings.

<sup>5</sup> Throughout, the dimension of the observed school and household effect vectors is taken to be  $N \times 1$ , where  $N$  is the number of children (observations). However, the number of *unique* household and school effects is strictly less than  $N$ . Indeed, assuming no singletons, at most  $H + S - 1 < N$  fixed effects can be estimated.

<sup>6</sup> This design stands in contrast to the conventional estimation of neighborhood effects (e.g., Solon, 2000), where neighborhoods nest households, meaning the variance contribution of the former cannot be larger than that of the latter (when treated as unobserved effects).

<sup>7</sup> While a number of studies report conditional variances, Table 1 ignores any conditioning variables. However, in our empirical application these are included (see Section 3.2).

additive contributions to outcomes (test scores). In terms of the associated outcome variance, this imposes the assumption:  $E(s'_k h_j) = 0$ , which rules out the possibility of any correlation between the household and school effects.

The zero-covariance assumption embedded in Row 1 appears restrictive. Sorting or clustering of households has been identified in a wide range of domains (Davidoff, 2005; Combes et al., 2008). Indeed, even if school quality were allocated randomly at time zero and held fixed thereafter, households' demand for better schools may bid-up local house prices over time, stimulating residential sorting and generating a positive correlation between household income and school quality (e.g., Hanushek and Yilmaz, 2007). Other behavioral responses of families toward school quality support the possibility of sorting into schools and evidence from a number of countries suggests that certain teachers prefer to teach certain kinds of children or to reside in specific kinds of neighborhoods (Jackson, 2009, 2013; Pop-Eleches and Urquiola, 2013). Thus, pre-existing clustering of households by income or ethnicity may stimulate across-location sorting of teacher quality. In either case, the assumption of a zero covariance between household and school factors in Row 1 becomes untenable, and an unrestricted linear model would apply (Row 2).

One interpretation of the unrestricted linear (sorting) model is that the household and school factors have no direct mutual effects – i.e., while they are separately pre-determined they become correlated through *ex post* processes of sorting. But this is not the only mechanism that could generate a correlation between household and school effects. Some part of the school effect may reflect the (mean) contribution of constituent households, such as when households make direct financial or time commitments to school functioning. This kind of mechanism also is suggested by certain versions of cream-skimming models, where average peer quality in a school (or class) is driven by household characteristics, which in turn directly influences individual achievement (Walsh, 2009). A strong version of this is captured in the third row of Table 1, which assumes that  $s$  can be partitioned into a component that is oblique or parallel to  $h$  and an orthogonal component  $v$ , with own variance  $\sigma_v^2$ :

$$s_k = \frac{\gamma}{N_k} \sum_{\forall j | K=k} h_{jK} + v_k \quad (2)$$

$$E(h'_{jK} v_k) = 0$$

Applying this expression, the second column of Row 3 gives a strict upper bound on the variance contribution due to households.<sup>8</sup> The corollary is given in Row 4, where household

---

<sup>8</sup> To go from equation (2) to the model in Row 3 we have made the simplifying assumption that  $h_j$  is constant within each school/grade  $k$ . Where this is not the case, it can be shown that the variance contribution due to households will be of a somewhat smaller magnitude but remains an upper-bound.

effects are assumed to be (partial) reflections of pre-given school effects, plus an orthogonal component. Note that in both these cases, the observed covariance between household and school effects is attributed wholly to either one of the factors. In other words, the contribution of sorting to the variance of test scores is moot under the assumption of a direct (one-way) causal relationship between the two factors. In this sense, alternative *ex ante* assumptions about the structure of the covariance are used to pin-down the variance contributions of the two factors.

Reflecting on Table 1, previous literature has frequently estimated IOE via some variant of the household upper-bound model (Row 3). Concretely, a number of studies treat family effects as a single fixed unobserved factor and omit any consideration of school effects. Björklund and Salvanes (2011) describe this approach and show how, under this set-up, the relative variance contribution of households is equal to the correlation in outcomes between siblings. The same authors summarize estimates of sibling correlations in various developed countries. Excluding estimates for twins, these range from 0.24 in former East Germany to over 0.60 in the USA. While many of these estimates are based on grades of completed schooling, Mazumder (2011) gives estimates of sibling correlations in various learning domains based on direct tests of children in the USA. His estimates are of the same broad magnitude, ranging from approximately 0.35 to 0.50. For the UK, Nicoletti and Rabe (2013) analyze results from compulsory national tests and find somewhat larger sibling correlations ( $>0.50$ ).

Estimates of sibling correlations in developing countries are scarce, mainly reflecting data constraints. An exception is Behrman et al. (2001), who find the sib correlation in terms of completed grades of schooling lies between around 0.30 and 0.60 across Latin American countries. To get around data constraints, an alternative approach is to identify a number of observed proxies for family effects, estimate their relationship to the outcome variable (using regression techniques), and then derive the variance of their fitted contribution(s). For example, Ferreira and Gignoux (2014) do so using ten variables as proxies for family background including: parental education, father's occupation, access to books at home, and migration status. Similarly, Schütz et al. (2008) use the number of books at home as their main proxy for the effect of background variables. However, since this approach amounts to a partition of the household effect into an observed and unobserved component, the variance attributable to the observed component only can be expected to represent a partial upper bound. This is verified empirically – i.e., the variance contribution of family background estimated via sibling correlations is typically much higher than found in studies using observed proxies. This suggests that these observed factors, such as the level of parental education, are rarely comprehensive (see also Behrman and Rosenzweig, 2004; Freeman and Viarengo, 2014).

School (or teacher) effects are of interest as they point to possible differences in school

quality. A large number of studies seek to assess the magnitude of such effects (e.g., Pritchett and Viarengo, 2015; Sass et al., 2012; Hanushek and Rivkin, 2006), in some cases adjusting for family background. However, studies of this sort are rare for low-income contexts,<sup>9</sup> and even fewer attempt provide a multi-country analysis, as we do here. Furthermore, just a handful explicitly estimate the variance contributions of both schools and households (e.g., Carneiro, 2008), and since most such studies are school-based they generally rely on a relatively limited set of observed proxies for family background.<sup>10</sup> For instance, Freeman and Viarengo (2014) use PISA data to investigate the (sources of) variance in school effects. They report that a regression of test scores on school dummies (only) explains around two-thirds of the variation in the data, while a limited set of observed family background variables accounts for just one-third, after controlling for school effects (also via dummies).

The general point emerging from the above (albeit brief) review is that existing studies have generally focused on estimation of either the contribution of households or of schools to variation in educational outcomes. These are of substantive interest, but such estimates correspond to special cases where the contribution of any covariance (sorting) between these effects is effectively absorbed into the main factor under consideration. Furthermore, even in cases in which such upper-bound estimates are tightened by introducing additional controls, such as proxies for one of the sets of effects, past studies have not reported the full variance decomposition incorporating the implied covariance between households and schools. However, it is precisely this covariance that can be of critical interest: it tells us about the extent to which schools – or educational policies more generally – exacerbate or compensate for differences deriving from given family background. In light of this gap, the next section outlines how a more complete decomposition can be elaborated.

## 2.2 Decomposition methods

Assuming that the household and school factors are not fully observed, decomposing their variance contributions is non-trivial. However, as hinted above, estimation of the upper bounds for each factor is straightforward and can be derived simply by treating each factor (separately) either as a fixed or as a random effect. Variants of this approach have been applied extensively (e.g., Raaum et al., 2006; Lindahl, 2011; Gibbons et al., 2014) and, under additional assumptions, also can be used to identify approximate lower bounds on the variance contribution of the second factor. Similarly, random effects (mixed-linear)

---

<sup>9</sup> See Behrman and Birdsall (1983) and Dang and Glewwe (2018) for two studies that respectively investigate school quality in Brazil and Vietnam.

<sup>10</sup> School-based studies are frequently limited to one or two grades and therefore do not contain data on multiple siblings. However, such studies do tend to provide rich data on school-specific characteristics.

approaches may be used, but these also place specific restrictions on the covariance of the unobserved factors.

In Section 3.3 below we implement such upper-bound and mixed-effects approaches. However, for now, our interest is in a procedure consistent with an unrestricted linear model. This requires we treat both latent factors as fixed effects with a (possible) non-zero covariance – i.e., we are in a two-way fixed-effects setting. Simultaneous estimation of crossed factors raises distinct technical challenges. In most applications, including here, the dimensions of the effects are extremely high, and their design is unbalanced. Consequently, standard approaches, such as inclusion of a full set of dummy variables, are not computationally feasible. Also, some kind of normalization restrictions are required in order for the model to be identifiable due to model over-parameterization (Mittag, 2012). Following the contributions of Abowd et al. (2002), among others, various solutions to these challenges have been proposed. Guimarães and Portugal (2010) show how a partitioned iterative algorithm can be optimized to solve the normal equations of a least-squares problem including multiple high-dimensional fixed-effects. The algorithm avoids the problem of inverting a large sparse matrix and can provide direct estimates of the fixed-effects.<sup>11</sup>

Two more specific challenges arise if the properties of the fixed effects are of standalone interest, as here. First, as the fixed effects are estimated with error, estimates of their variances will be biased upwards. Thus, variance shares calculated directly from the fixed-effects estimates will tend to over-state their importance relative to the residual component (Koedel et al., 2015). Second, Andrews et al. (2008) demonstrate that the covariance of the estimated fixed-effects vectors tends to be biased downwards. This is driven by a quasi-mechanical relation, whereby if one factor (e.g., household effects) is over-estimated then on average the other factor (e.g., schools) will be under-estimated (also Andrews et al., 2012). Intuitively, this reflects the general problem of model over-parameterization; and the magnitude of bias tends to be larger where fewer observations are available to estimate each effect.

Addressing these challenges remains an active area of research. Nonetheless, in Appendix A we set out the details of our proposed solutions. To correct for measurement error, we apply a conventional procedure that shrinks the variance contribution of estimated factors in accordance with the number of observations available to estimate each effect (see also Stanek et al., 1999; Koedel et al., 2015).

To deal with downward bias in the between-factor covariance, we propose a novel approach. Looking ‘under the hood’ of the iterative algorithm reveals that a part of the bias is driven by how starting values for the fixed-effects are calculated. In previous applications, extreme

---

<sup>11</sup> Their procedure is implemented in Stata under the user-written *reghdfe* command (Correia, 2017).

initializations have been employed that apportion the (residual) variation in the outcome to a single (presumed dominant) factor, effectively treating the second factor as orthogonal. Our insight is that these starting assumptions about the fixed effects are important because they become effectively locked-in from one iteration to the next. That is, the starting values represent a crucial identifying assumption for the variance decomposition. Furthermore, we show that the assumed form of the initial between-factor covariance can be explicitly controlled via introduction of a single initialization parameter. Denoted  $\pi \in [0, 1]$ , this parameter controls how the (residual) variation in the outcome is apportioned across the two factors to start the iteration procedure. In keeping with earlier discussion (Table 1), we hypothesize that corner values ( $\pi = 0, \pi = 1$ ) will correspond to upper-bound factor models, such as described in rows 3 and 4 of Table 1, which rule out between-factor sorting. However, an agnostic or midpoint choice ( $\pi = 0.5$ ) initially apportions the variation roughly equally across both factors. So, a corollary hypothesis is that this choice is likely to yield an upper bound on the magnitude of the between-factor covariance (sorting). As such, we do not propose a single correction for the potential bias in the estimates of the between-factor covariance. Instead, we provide a unified approach to the estimation of two-way fixed effects that makes it possible to impose – in a transparent way – alternative (starting) assumptions about the relationship between the two factors, including their covariance.<sup>12</sup> As we show, this serves to bound the estimates of interest.

### 3 Application to East Africa

The previous section set out a general factor model for thinking about IOE and a unified empirical approach to decompose the variance within a two-way fixed-effects context. In the remainder of the paper we implement and compare results from three main choices for the fixed-effects initialization parameter,  $\pi$ , and validate our results by considering both alternative empirical methods and a wider range of choices for  $\pi$ . Based on our preferred results, we also go on to investigate heterogeneity in the variance contributions.

#### 3.1 Data

Our application of these methods uses test score data from East Africa. Since 2010, the Uwezo initiative has undertaken large-scale household surveys in Kenya, mainland Tanzania and Uganda (for further details and comparison to other regional assessments see Uwezo,

---

<sup>12</sup> Note that extensions to more than two variables are possible in theory but add substantially to the number of covariance terms to be estimated, as well as the dimensionality of the choice space for the initialization process.

2012; Jones et al., 2014). The approach adopted by Uwezo has been inspired by exercises carried out in India by the Assessment Survey Evaluation Research Centre (ASER), which has surveyed the literacy and numeracy abilities of over 500,000 children each year since 2005. As with ASER, the target population of the Uwezo surveys has been children (residing in households) aged between the official starting-school age and 16. The surveys have been designed to be representative at both national and district levels, based on the administrative classifications in the most recently available population census.<sup>13</sup> Excluding the initial surveys, five rounds of the Uwezo surveys have been completed (2011-2015) and are used here.

In each assessment, the surveys collected information at the household level covering household characteristics and the demographic and educational details of all resident children (e.g., age, gender, whether or not attending school, etc.). Within each household, the children of school age were individually administered a set of basic oral literacy and numeracy tests. These tests have been based on a common template, but have been tailored to each country and varied by survey round. Specifically, in each round and country, local experts have taken the template and developed item content to reflects competencies stipulated in the national curriculum at the grade 2 level. That is, the tests are anchored to skills that should be achieved by the majority of pupils after two years of completed schooling.

The literacy and numeracy tests (the Uwezo tests) are described in detail in Jones et al. (2014). The literacy tests refer to national languages of instruction in which pupils are tested at the end of primary school – i.e., English and Kiswahili in Tanzania and Kenya; and just English in Uganda. Importantly, the Uwezo tests are not adapted to the children’s ages or their completed level of schooling. Given that they focus on basic competencies, it is thus unsurprising there are strong age- related differences, which affect both the level and variance of scores *between* age cohorts. From the present perspective, this between-cohort variation can be considered unwanted noise (see Mazumder, 2008).<sup>14</sup> As a result, so as to construct an overall metric of achievement, we transform raw integer scores on the individual tests in three steps. First, each score is standardized by age, such that the individual test scores have means of zero and standard deviations of 100 for each age group. Second, we calculate weighted means of the age-standardized scores on each test, placing equal weight on the literacy and numeracy components. This gives a synthetic or overall test score, the primary outcome of interest hereafter. Third, to facilitate interpretation and to address potential differences in the test difficulty between countries and rounds, we normalize this measure for each country and

---

<sup>13</sup> In some survey rounds, however, administrative difficulties meant that certain districts could not be surveyed. Throughout, (adjusted) survey weights are used that take into account these implementation issues.

<sup>14</sup> We recognize the contribution of different variance components may vary across age cohorts and we investigate this in our empirical analysis.

round such that the final standardized score has an overall mean of zero and standard deviation of one.<sup>15</sup> Table 2 reports regional means and standard deviations of the test scores, calculated at each step. The first (column I) are the weighted means of the raw competency tests (before any standardization); column II reports the age-standardized versions; and column III reports the final measures. As can be seen, movement from the second to the third metric constitutes a simple monotone transformation.

The test scores reported in Table 2 refer to the final sample used in the subsequent analysis, which pools all survey rounds. This is a slightly reduced sample of the original Uwezo data. Specifically, observations have been dropped that can be perfectly predicted using either household or school fixed effects – i.e., all singletons were removed. The objective of restricting the data in this way is to mitigate upward bias of the variance contribution estimates. As per the methodological discussion of Section 2.2, the analytical focus is on the variance components of the test score; and there is no evidence to suggest these dropped observations are distributed in a systematic pattern over regions or districts.<sup>16</sup> The (sample) standard deviations of the test scores, which can be directly interpreted as normalized measures of educational inequality (Van de Werfhorst and Mijs, 2010; Hanushek and Wößmann, 2006), are reported in parentheses in the table. It merits noting that the rank position of each region according to its test score variance is largely preserved, regardless of the transformation applied.

To implement the variance decomposition, the household and school indexes must be defined. The former is trivial – unique indexes are ascribed to all sibs in the same household in each year.<sup>17</sup> The school effects are less straightforward. In the present data, detailed information about the particular school each child attends is not provided. Nonetheless, we can identify the grade of attendance and the location of the household. Consequently, for each enumeration area (containing approximately 20 surveyed households), we categorize children into one of three school-grade categories based on their highest grade of completed or current schooling – namely: those attending grades 1-2; grades 3-4; and grades 5 or higher. For children who have never attended school, we hypothesize that the quality of schools (teachers) available in their locale may have an effect on their ability in numeracy and literacy. This may work directly, through the choice not to attend school, as well as indirectly through both sibling and peer effects – e.g., what other children learn can spill-over to non-attenders. In order to allow for these effects in the data, we allocate never-attenders to the

---

<sup>15</sup> Due to the absence of equating or anchor items in the Uwezo tests, we cannot distinguish between changes in test difficulty over time and changes in (average) learning outcomes. However, from the perspective of a variance decomposition, standardization by year ensures that the variance contributions retain a consistent meaning in relation to the overall distribution of outcomes in each age cohort and year.

<sup>16</sup> Details available on request from the authors.

<sup>17</sup> The Uwezo surveys are cross-sectional in nature and no explicit attempt is made to track the same children over time.

median grade category (index  $k$ ) of children with the same age and location.<sup>18</sup> Additionally, individual-level controls can be included in the variance decomposition (see Section 3.3) to account for children's specific educational status. Overall, this classification ensures that children within each household are unlikely to share the same school-grade index – i.e., the school-grade and household effects are crossed as opposed to nested. A downside is that if there are multiple schools of the same type in a given enumeration area, or children travel to another location to attend school, then children may be incorrectly treated as attending the same school. Consequently, school-grade fixed effects capture average school quality of a given type in a given location for each schooling level, but children may be subject to classification error.<sup>19</sup>

Descriptive statistics for the data set are reported in Table 3. This shows the number of unique children ( $i$ ), households ( $j$ ) and school-grade effects ( $k$ ) covered in the data set. Additionally, the table reports summary statistics of child characteristics (age, gender), schooling status indicators, and a normalized measure of socio-economic status (SES) based on observed household assets. Overall, these indicate the sample is comprehensive and balanced (by age and gender).<sup>20</sup> It also reveals that, although the vast majority of children are attending school, there are systematic differences in (mean) school status among regions within each country, and these seem to relate closely to differences in mean socio-economic status. For instance, in Kenya, there is an average difference of 1.5 grades between the (wealthier) Central and (poorer) North Eastern regions. These also appear to map into large differences in the magnitude of inequality in test scores.

### 3.2 Unconditional decomposition

The first empirical model we implement closely follows the framework outlined in Section 2, incorporating only the two fixed effects of interest. As also described earlier, we focus on three initializations of our implementation of the partitioned iterative algorithm. These are: (1)  $\pi = 0$ , which first allocates variation in the outcome to the household effect and the residual to the school effect; (2)  $\pi = 0.5$ , which is agnostic about how the variation should be initially allocated; and (3)  $\pi = 1$  which first allocates variation in the outcome to the school effect and the residual to the household effect. Since we are primarily interested in converging on (stable) estimates of the estimated fixed-effects, we stop each run of the algorithm when the estimated scalars on the plugged-in fixed effects (see Appendix A) are both not significantly different from one, and the absolute change in root mean square error between

<sup>18</sup> We examine the robustness of this set-up in our empirical analysis.

<sup>19</sup> Where the ‘true’ schooling effects are mutually independent, then misclassification error in the definition of the fixed-effects would attenuate their estimates, biasing the variance contribution downward.

<sup>20</sup> Average ages are higher in Tanzania as the starting school age is seven, compared to six in the other countries.

iterations falls below 0.01. Convergence is typically achieved in less than 20 iterations.

A summary of the main results is set out in Table 4, which reports the absolute and relative variance contributions respectively. Four main insights stand out. First, the different initializations have the hypothesized effects, with the variance components varying in the directions expected. Looking more specifically at how these contributions change, we note that the household and school variance contributions, but not their covariance, vary monotonically with changes in  $\pi$ , with the extreme values for  $\pi$  corresponding to upper/lower bounds (as hypothesized). Concretely, when  $\pi = 0$  the household share is largest (at around 35% in relative terms), the household-school sorting component share is close to 0%, and the school share is moderate (around  $\frac{1}{4}$  of the household share or 7%). When  $\pi = 1$ , the magnitudes of the household and school shares are roughly reversed, although the household component remains somewhat larger, at around 13%; but the sorting component continues close to zero. This switch directly reflects the opposite ways in which the between-factor covariance is allocated – i.e., under the extreme choices we assume either all covariance emanates from the causal effect of households on schools ( $\pi = 0$ ), or it goes *vice versa* ( $\pi = 1$ ).

In contrast to these two extreme choices, the agnostic initialization does not assume one effect is uniquely driven by the other (see Table 1); rather, it retains the covariance (sorting) term as a separate and substantial contribution. In turn, the results show this estimator does not yield such a substantial mechanical negative covariance bias as is also associated with more conventional two-way fixed-effects estimators (Andrews et al., 2012; Gaure, 2014). Indeed, in our case, this initialization indicates sorting between households and schools accounts for up to around 8% of the variation in test score outcomes (or an effect size of 0.28 standard deviation units). Comparable estimates for the sorting contribution in other contexts are rare; but the magnitudes found here are broadly in line with the un-shrunken estimates for Carneiro (2008) for Portugal. The immediate implication is that the allocation of children and/or teachers to schools tends to aggravate rather than compensate for existing (familial) inequalities. In turn, this suggests there is ample scope for policies to enhance access to schools of the same quality.

Second, the results suggest that households are not necessarily the primary source of IOE. Of course, such a conclusion naturally holds under the household upper bound model ( $\pi = 0$ ), but we have highlighted that this pertains to specific assumptions that rule out the contribution of sorting. When sorting half-way between the extreme cases is explicitly allowed ( $\pi = 0.5$ ), we find households and schools make approximately equal contributions to outcome inequality, accounting for around 15% of the total variance each. That is, by

excluding the contribution of sorting, previous studies using household upper bound models may well have over-stated the relative importance of households to outcome inequalities. In turn, and contrary to the proposition that schools contribute little to test score outcomes, we find differences among schools are a material source of inequalities in educational achievement. Moreover, the magnitude of the variance attributable to schools (grades) is not an order of magnitude lower than that found elsewhere (Hanushek and Rivkin, 2012; Azam and Kingdon, 2015), including studies that report upper bound school variance estimates (e.g., Pritchett and Viarengo, 2015).

Third, summing together the estimated variance contributions, we note that IOE is substantial across all countries. The residual or unexplained component, which roughly captures effort and preferences for education, accounts for a little over half of the total outcome variance. This indicates that equalizing educational opportunities would be expected to reduce the variance of test score outcomes by at least 40%. In relation to existing literature on IOE, these magnitudes are substantial. And it is clear that a part of this may relate to the more comprehensive approach we have adopted – i.e., we cover multiple sources of IOE and do not exclusively rely on (partial) observed proxies. Again, an implication is that previous studies may well be underestimating IOE.

Fourth, while the broad pattern of results described above holds across the three countries, there are also some differences. Figure 1 shows that the relative variance estimates for Tanzania are more distinctive, suggesting a generally larger contribution of household and sorting effects, and a smaller contribution from schools. A complete interpretation of these differences falls outside the scope of the current analysis. However, it hints at heterogeneity that may extend below the national level, to which we turn later in Section 3.5.

### 3.3 Conditional decomposition

The specification considered above abstracts from a range of (observed) child characteristics, such as gender and school enrollment status. Where these are correlated with either the school or household fixed effects, the previous estimates may be contaminated by omitted variables bias. To address this, we extend the simple unconditional model in two directions. First, we add an individual-specific component. Without longitudinal data, we cannot treat this as a latent variable. Instead, we partition this component into observed and unobserved parts:  $a_i = a^o + a^u = x_i' \alpha + a^u$ ; where the vector  $x_i$  contains five dummy variables that take a value of one if: children are female, they are the first born (oldest observed child), they are currently enrolled in school, they have never enrolled, and they are attending private school. The unobserved individual component remains in the error term.

Second, the preceding definition of school fixed effects is somewhat crude. Specifically,

children who have never attended school are allocated to the same unit (school-grade effect) as their local peers; and we do not distinguish between school types (private versus public). Increasing the number of school fixed effects is problematic due to the limited number of available observations. Nonetheless, we can extend the specification to allow each of the existing school-grade effects to vary by a multiple (i.e., to be scaled upwards or downwards) across children attending public schools, children attending private schools, and never attenders. Putting these extensions together yields the following empirical specification:

$$t_{ijk} = x'_i \alpha + h_j + s_k(1 + \lambda n_i + \kappa p_i) + (a^u + \varepsilon_{ijk}) \quad (3)$$

where  $n_i$  and  $p_i$  are both elements of  $x$ , being the dummy variables for never attenders and private school pupils respectively. Note this specification nests a test for whether there is any spillover of school quality from local peers attending school to non-attenders – i.e., positive spillovers obtain as long as  $\hat{\lambda} > -1$ ; and we can also test for the extent to which variation in school quality between public and private schools is correlated across locations – i.e., we cannot reject that they are correlated if  $\hat{\kappa} = 0$ .

Under our proposed partitioned iterative algorithm (see Section 2; Appendix A), inclusion of these interaction terms is straightforward. For the purposes of the variance decomposition, however, the additional terms demand consideration of multiple extra covariance terms. To simplify matters, for each individual we calculate the individual-specific aggregate or final value for the school fixed effect ( $s_k$ ), which absorbs the estimated contribution(s) of the interaction terms. For instance, in the case of never attenders, the final estimated school effect is calculated as:  $\hat{s}_k(1 + \hat{\lambda})$ . Using these final school effect estimates, the remaining covariance terms are subsequently estimated. Thus, the variance decomposition we report now contains seven elements:

$$\text{Var}(t_{ijk}) \equiv \sigma_t^2 = \sigma_x^2 + \sigma_h^2 + \sigma_s^2 + 2\Sigma_{hs} + 2\Sigma_{xh} + 2\Sigma_{xs} + \sigma_\varepsilon^2 \quad (4)$$

$$1 = \frac{\sigma_x^2}{\sigma_t^2} + \frac{\sigma_h^2}{\sigma_t^2} + \frac{\sigma_s^2}{\sigma_t^2} + \frac{2\Sigma_{hs}}{\sigma_t^2} + \frac{2\Sigma_{xh}}{\sigma_t^2} + \frac{2\Sigma_{xs}}{\sigma_t^2} + \frac{\sigma_\varepsilon^2}{\sigma_t^2} \quad (5)$$

Turning to the results, it is informative to begin with the regression output from the three estimators. These are summarized in Tables B1-B3 in Appendix B, treating each of the three countries separately. For purposes of comparison, column (1) is a naïve estimator, which is just an OLS regression of the baseline model ignoring the fixed effects; column (2) reports results from a conventional implementation of the partitioned iterative algorithm incorporating the two fixed effects of interest (based on the user-written *reghdfe* command in Stata, due to Correia, 2017), where the household effect is initially swept out of the regression (for speed). Columns (3) to (5) report the regression results associated with the three

alternative initializations, now including the interaction terms.

Three main points merit attention. First, the estimated regression coefficients and the overall coefficients of determination (R-squared) change materially when moving from columns (1) to (2), indicating the latent fixed effects are both relevant and correlated with the observed covariates – e.g., in Kenya, introduction of the fixed-effects leads to a more than 50 percentage point increase in the model R-squared. Second, comparing results across columns (2)-(5), the reported regression coefficients are not statistically indistinguishable from each other and the R-squared statistics do not change.<sup>21</sup> Together, this indicates that when the fixed effects are not of inherent interest (e.g., are to be treated as nuisance parameters), then the choice of initialization is unimportant. That is, the initial allocation of the covariance across the fixed-effects is material for the variance decomposition (as established in Section 3.2), but not for the levels regression estimates. Additionally, the similarity of the coefficient estimates supports the specific implementation of the iterative algorithm we have employed here – i.e., regardless of the initialization parameter, the procedure yields regression coefficients that are consistent with established two-way fixed effects approaches.

Third, the coefficients on the interaction terms are generally negative and statistically significant. In particular, and as might be expected, the school effects are substantially scaled downwards (shrunk toward zero) for children who have never attended school – e.g., under the midpoint initialization, the school effect is halved in Uganda for those without school experience. However, in no case does the final school effect approximate zero, suggesting the presence of some learning spillovers from attenders to non-attenders. It follows that improvements in school quality may well have a broader effect on achievement beyond children directly exposed to any improvements. At the same time, we find much less of a systematic difference in the (level of the) school effects between public and private schools in the same locations. This implies these effects are correlated – we tend to find relatively better (worse) private schools alongside relatively better (worse) public schools; and that there is a roughly similar amount of heterogeneity in private school and public school quality (see further below). Nonetheless, note there remain marked differences in the *average* level contributions of private schools. This is given by the coefficient on the private attendance variable, which is positive in all countries and ranges from 0.10 to 0.20 standard deviation units.

Sticking with the estimates from the partitioned iterative algorithm based on the extended specification, Table 5 reports the associated variance decompositions, in which the joint contribution of the individual terms is aggregated (see Appendix Tables B4-B6 for the

---

<sup>21</sup> Estimates for the ‘never enrolled’ term do vary across estimators. However, this is due to the inclusion of the interaction term. When this is dropped there are no remaining differences. (Results available on request.)

complete disaggregation). An immediate observation is that the previously-excluded contribution of the individual terms is material, accounting for between 4% and 9% of the total test score variance (or 0.19 to 0.30 standard deviation units). Furthermore, the extended specification suggests the contributions of both the household and school factors are moderately smaller in comparison to unconditional results. For example, in Uganda the upper bound estimate ( $\pi = 0$ ) for the household effect contribution drops from 32% (0.57 s.d. units) to 28% (0.53 s.d. units) when we condition on the individual characteristics. Similarly, the sorting contribution at the mid-point initialization also marginally declines under the conditional model to around 6% (0.13 s.d. units). At the same time, it is clear that *only* the household and school variance contributions, plus the corresponding sorting term, are sensitive to the choice of initialization. Thus, consistent with the regression outputs (Tables B4-B6), the contributions of the individual component and the residual remain stable across the values of  $\pi$ .

In sum, these results suggest that there is substantial continuity between the unconditional and conditional (extended) models, but the latter provides a more nuanced picture of the variance components in which some variation due to individual characteristics is permitted. Moreover, the main substantive insights from the unconditional model are retained here. Namely, we find that IOE is substantive and is unlikely to be only due to the household component.

## 4 Validation and Further Analysis

The previous subsections reported results from our proposed general estimator of a two-way fixed-effects model, which provides a practical and unified framework for the decomposition of variance under alternative covariance assumptions. As hypothesized, the results effectively bound the main estimates of interest:  $\pi = 0$  gives upper bound estimates for the household contribution;  $\pi = 0.5$  gives upper bound estimates on the sorting component; and  $\pi = 1$  gives upper bound estimates for the school contribution. To validate these results, we now pursue two strategies. First, we consider a broad range of values for  $\pi$  (increasing at intervals of 0.1) and plot the corresponding relative variance contributions. These results are shown in Figure 2 and confirm that the chosen values do yield upper/lower bounds on the main components. Furthermore, the figures indicate that the lower bound household share in all countries is moderately greater than the lower bound school share; and that the sorting component follows a shallow inverted-U shape. Indeed, we note that  $\pi = 0.5$  only corresponds to an approximate upper bound on the between-factor variance. Marginally larger point estimates are obtained when  $\pi = 0.4$  in two countries (Kenya and Uganda), however such differences are within sampling variation. In fact, broadly similar estimates for the contribution of sorting are found across the range  $\pi \in [0.33, 0.66]$ . This supports the notion

that a positive and non-trivial contribution of sorting obtains over a material domain of the choice space.<sup>22</sup>

Second, we compare our results to those obtained from three alternative approaches.<sup>22</sup> Firstly, we consider a standard household upper bound model (denoted UBS), estimated via a high-dimensional fixed-effects estimator (also a partitioned algorithm), but where only a single fixed effect for households is included. In doing so, the contribution of schools is initially ignored. However, consistent with the data-generating process set out in Row 3 of Table 1, the remaining orthogonal contribution associated with schools can be derived by taking school-wise averages of the estimated residuals from the previous step regression. This amounts to a two-step estimation process.

For the second approach, an upper bound on the school effects (model UBS) can be estimated in similar fashion. Given the available data, a distinctive feature here is we are able to include *observed* household covariates (for a similar approach see Solon, 2000; Raaum, 2006). Splitting out the latent household effect into observed and unobserved components:  $h_j = h_j^o + h_j^u = z_j'\beta + h_j^u$ , a first-step model corresponds to the following specification (see also Carneiro, 2008):

$$t_{ijk} = x_i'\alpha + z_j'\beta + s_k + \tilde{e}_{ijk} \quad (6)$$

where  $\tilde{e}_{ijk} = h_j^u + e_{ijk}$

Explicit inclusion of the observed household effects in equation (6) means the school effects can be considered as being conditional on these proxies; thus, their estimated variance should yield a tight(er) upper bound. Deriving the unobserved household component from the first-stage residuals as before – i.e.,  $\hat{h}_j^u = (1/N_j) \sum_{i|J=j} \tilde{e}_{ij}$ , this approach implies an estimate for the lower bound variance contribution of household factors as follows:  $\text{Var}(h_j) = \beta' \sigma_z^2 \beta + \sigma_{\hat{h}^u}^2$ . As with the first model, with estimates of each component in hand, the corresponding covariance terms also can be derived.

The third procedure adopts a different strategy. Recognizing that the unobserved household component omitted from equation (6) (later derived from the residual) can be considered a random effect, simultaneous estimation using a mixed-linear model is viable, treating both unobserved components ( $h_j^u, s_k$ ) as random effects. In this case, the empirical model to be estimated is:

$$t_{ijk} = x_i'\alpha + z_j'\beta + [h_j^u] + [s_k] + e_{ijk} \quad (7)$$

---

<sup>22</sup> As indicated below (see equation 6), our implementations of the alternative approaches omit the interactions with the school effect. However, this does not materially affect the broad direction of results.

and where the square brackets  $[ \cdot ]$  indicate effects treated as random latent variables. A potential drawback of this strategy is that some parametric structure on the random effects and their covariances must be imposed. In principle, the two random effects can each be treated as correlated with the elements included as fixed effects (e.g., vector  $z_j$ ). However, in practice, such unrestricted covariance structures are not only computationally intensive but also show poor convergence properties in large datasets (Gurka, 2011; Chirwa, 2014). Thus, typically at least some covariance restrictions do need to be applied in the estimation procedure; and the random effects must be treated as mutually orthogonal. Together, these assumptions mean that the population variance components estimated via a mixed-linear model may not adequately approximate those of the true data-generating process under the linear unrestricted model.

Offsetting this concern, we note that the properties of the predicted effects from a mixed-linear model do not necessarily reflect the restrictions imposed on the population model. As discussed in Bates (2010) (also Stanek et al., 1999), the best linear unbiased predictors (BLUPs) of random effects in these estimators are derived from the residuals of the estimated model and can be understood as conditional means (i.e., being conditional on the observed data and model parameter estimates). Consequently, since they represent an (optimal) approximation to the unit-specific values of the latent variables in the observed sample, they do not necessarily conform to the properties assumed to hold in the population. For example, as shrinkage methods are applied in derivation of the BLUPs, the variance of the predicted conditional means of the random effects tend to be lower than the corresponding population variance estimates. Similarly, and as we find here, in order to provide an optimal fit to the actual data, the same BLUPs may not be mutually orthogonal. This suggests that in the present case, whereas imposing a zero covariance restriction may be too strong, a variance decomposition based on the properties of the random effects BLUPs and estimated fixed-effects from a mixed-linear model can nonetheless remain informative (and less restrictive) for the specific sample in hand.

Appendix Tables B4-B6 compare results from our modified partitioned iterative algorithm (as earlier) and the three alternative estimators described in this subsection. In all cases, the household and school upper bound variance components, derived from the (two-stage) UBH and UBS models respectively, are highly comparable in magnitude to the same estimates from the corresponding modified partitioned iterative algorithm. For instance, in Tanzania, the absolute variance contribution due to households when  $\pi = 0$  is equal to 0.59 standard deviation units, and 0.58 under the UBH model. Similarly, in Kenya, the school variance contribution is 0.53 units when  $\pi = 1$ , and 0.51 units under the UBS estimates. In other words, these results confirm that extreme choices for  $\pi$  map to corner assumptions about how the between-factor covariance is allocated.

Despite the similarity of the upper-bound estimates, the corresponding lower bounds in each model are not so similar. Again, taking the example of Tanzania, the estimate for the school variance component is 0.12 standard deviation units under the (two-step) UBH model but is 0.25 for  $\pi = 0$ . Effectively, since the individual and residual terms are (on aggregate) also similar under these two models, the difference is found in the covariance term, which is systematically positive under the two-step estimates but is always much closer to zero under the extreme initializations of the modified partitioned iterative algorithm. This supports the contention that the latter approach produces systematically smaller estimates of the sorting component. Furthermore, estimates from these extreme choices do not appear very credible. Recall the UBS model includes proxies for the household terms, and this model yields estimates for the magnitude of sorting that are always positive and significantly greater than those based on the extreme initializations. For instance, the between-factor correlation ( $\rho$ ) in Table B4 (Kenya) is 1% when  $\pi = 0$  but is 9% under the two-step procedure (model UBH). That is, as the two-step upper-bound model points to the presence of positive sorting, the absence of this finding under the extreme initializations of the partitioned iterative algorithm support the contention such initializations suffer from a mechanical negative bias on the sorting (covariance) term.

Finally, estimates based on the BLUPs from the mixed-linear model closely resemble the results from the agnostic initialization across the various components. This is perhaps most stark for the sorting component, which is also larger in magnitude under this approach than found under either the UBH or UBS methods (as well as for the extreme initializations). Admittedly, the contribution of the school effects is somewhat larger under the mixed-linear model compared to the agnostic initialization, but this in part reflects the influence of the interaction terms. Thus, while we are not in a position to claim the agnostic initialization corresponds to the ‘true’ data generating process, it appears to provide the most reasonable and well-supported (conditional) variance decomposition in the present case.

## 4.1 Sub-group heterogeneity

Having validated our proposed methodology, we now investigate the presence of heterogeneity in the variance components, focusing on the preferred agnostic initialization (upper bound sorting model) with the extended specification. To do so, we re-run the variance decomposition based on the earlier regression estimates but now, instead of reporting whole-sample results, we stratify individuals according to various individual characteristics (gender, age group, schooling level, attendance of public/private school, SES status, and maternal schooling).<sup>23</sup> Appendix Tables B7–B9 report these findings.

---

<sup>23</sup> This is based on the same aggregate (regression) results reported previously, the difference being that the

Notwithstanding differences *among* the three countries, there generally appear to be systematic differences in the magnitude and (relative) contributions of the components to learning inequality. In particular, we observe greater inequality across boys (versus girls), among children in lower grades (or out of school) and among children from poorer households. There is also evidence that IOE is (relatively) more significant among younger children – e.g., in Kenya the residual accounts for 55% of the variance for children aged 6–8, but over 66% for those aged 12 and above. Also, with the exception of Uganda, IOE due to schooling tends to be larger in both absolute and relative terms among children attending public as opposed to private schools – e.g., in Kenya, the variance contribution of schools is 0.38 standard deviation units (17%) for children attending public schools, but 0.30 (13.5%) for children attending private schools, implying less heterogeneity in learning outcomes across schools in the private sector.<sup>24</sup>

Looking more closely at children who have not attended school (indicated in the tables by the level zero of the grade category), two features are of interest. Consistent with the regression coefficient interaction terms reported in Tables B1-B3, the (relative) variance contribution of schools among non-attenders is considerably lower than for attenders. Even so, we note the sorting term remains positive and significant for the non-attenders, being in fact larger in magnitude (absolutely and relatively) in Kenya than for school attenders. This indicates a direct effect of school quality on the decision to (ever) enroll in school, which seems to disadvantage children from less privileged backgrounds. Figure B2 and Appendix Tables B10-B13 confirm the smaller contribution of schools as well as a larger (relative) contribution due to households among non-attenders versus attenders for alternative choices of the initialization parameter, now using only sub-samples of children for which all children in the same family share the same schooling status. Moreover, since the magnitude of learning inequalities is considerably lower among children who attend school than among those who do not, we conclude that even in low-income contexts, such as found in East Africa, access to schooling does go some way to addressing learning inequalities.

## 4.2 Spatial heterogeneity

A further form of heterogeneity is among geographical locations. Arguably, this is particularly relevant from the perspective of policy as it speaks to the possibility for targeted interventions. This is also motivated by the educational differentials within each country, shown in Table 2, which indicate large differences in the mean levels and variances

---

variance components are simply calculated separately for each sub-group. Other stratifying variables were also examined, such as the survey year, but were not found to be of substantive interest (results available on request).

<sup>24</sup> What may account for this finding lies beyond the scope of the present paper. However, private schools are more common in urban areas, where there is greater competition (choice) in school provision.

of test scores across regions, and which in cases are larger than those due to low socio-economic status alone.

In this spirit, Tables 6 and 7 report the variance decomposition at the regional level, using the preferred choice  $\pi = 0.5$ , and where we also report the upper bound estimates for the household and school effects (first two columns). As above, we find substantial heterogeneity in the size and relative importance of the different factors. This is most stark in Kenya where the absolute variance contribution of different factors can vary by a factor of around four. For instance, under the preferred agnostic initialization, the absolute contribution of household factors in the Central region is 0.30 standard deviation units, which is smaller than the absolute contribution of sorting in the North Eastern region at 0.36 units. Similarly, we see very large differences in the variance contribution of schools, ranging from 0.26 (Central) to 0.47 units (North Eastern). When considered in relative terms (Table 7), these differences are less pronounced; but, even here there remain material differences in the contributions of schools and sorting between regions (e.g., the sorting component accounts for a minimum of 2% and maximum of 7% across Kenyan regions). We also note that the regions containing the capital city of each country (defined here as Central in Kenya and Uganda, Dar Es Salaam in Tanzania) tend to display comparatively low overall test score inequalities as well as larger contributions from the household effects and smaller contributions from both school and sorting effects (in relative terms). This is consistent with the notion the capital cities provide more equal access to schools of a similar quality.

Finally, we proceed to an analysis at the district level. Figure 3 illustrates the cumulative distributions of the relative variance shares for various components taken from the preferred estimates. These confirm substantial variations within each country, but they also suggest country-specific differences continue to be evident, especially in the contributions of the household and school effects – i.e., the distribution functions display (near) first-order dominance. To investigate whether these geographic differences are systematic, we regress the absolute and relative variance shares of the same components against a number of district-level characteristics (essentially, means taken from the same dataset). This analysis, which is intended only to indicate conditional correlations, is found in Table 8, treating all countries together.

Three points should be highlighted. First, a part of the heterogeneity in the variance decomposition seems systematic. Aside from the country fixed effects, whether the mother has any schooling and whether the child attends private school show significant correlations across the different variance components. In absolute terms (super-column I), lower maternal schooling is associated with larger contributions from all components – i.e., a higher prevalence of maternal schooling maps to lower educational inequalities, which is also

consistent with the same children being out of school. Districts with a higher prevalence of children attending private school show higher inequalities associated with household factors, but lower relative inequalities associated with schools. The latter provides further indicative evidence that (greater) competition across schools may help narrow the distribution of schooling effects. Furthermore, and as indicated by the interaction term, this effect seems most acute at lower grade levels – i.e., there is substantial homogeneity in the contribution of private primary schools to learning outcomes in early grades.

Second, and relatedly, sorting effects appear be larger, in both absolute *and* relative terms, in districts with a higher prevalence of disadvantaged households (i.e., those with lower average SES values and a higher proportion of mothers with no education). The data do not indicate what lies behind this – e.g., it may be due to greater clustering or residential segregation in these districts, or a less equal distribution of school quality. However, this result reinforces the idea that a deeper investigation of sorting processes may be helpful. Equally, and third, country differences remain persistent after controlling for other covariates. One interpretation is that such persistence can be explained by differences in the overall organization of schooling systems (macro-policies) and that these differences are material for educational inequalities. This would be in line with previous studies that find policy differences, such as ability-tracking and the allocation of teachers among schools, can contribute to national differences in educational inequalities (Hanushek and Wößmann, 2006; Van de Werfhorst and Mijs, 2010). And this agenda merits further attention in the context of East Africa.

## 5 Conclusion

The purpose of this paper is to shed light on the sources of educational inequalities in East Africa. Starting from the proposition that household circumstances, school factors and assortative matching (sorting) are *all* potentially important components of IOE, we seek to parse out their respective contributions to the observed variation in test scores. To do so, we review various variance decomposition procedures and suggest how a partitioned iterative algorithm can be used to estimate the relevant variance components, treating the two effects of interest as fixed latent variables. In order to address technical challenges of estimation, namely the problem of a mechanical negative bias in the correlation between the fixed-effects, we elaborate a unified procedure that controls how the fixed effects are estimated (initialized) and that maps to alternative assumptions about the between-factor covariance.

We apply the approach to rich test score data from East Africa from which three main findings stand out. First, we confirm that how the fixed effects are initialized under a partitioned iterative algorithm matters for subsequent estimates of the variance components.

At the same time, the choice of initialization does not affect the regression estimates (coefficients) for the included observed covariates. This implies the choice of initialization is *not* material when the fixed effects are considered nuisance parameters; but, if scholars wish to extract and interpret the latent fixed effects, the initialization choice is fundamental.

Second, our proposed unified approach to estimation, in which different initializations of the latent fixed effects are examined, provides bounds on their respective variance contributions. That is, extreme choices of the initialization parameter yield upper and lower bounds on the household and school variance contributions; and a midpoint (agnostic) initialization provides an (approximate) upper bound on the between factor covariance, which is interpreted as the contribution of sorting. Methodologically, these insights are validated using both conventional (single) fixed effects and mixed-linear models. Consistent with the existing literature, we confirm that the extreme (corner) initializations of the fixed effects bias the covariance term toward zero. However, we show the agnostic (midpoint) initialization substantively mitigates this problem and provides estimates of sorting that are material, positive, and comparable in magnitude to those from a mixed-linear model.

Third, taking the agnostic initialization as our preferred approach, we find that household factors are an important source of inequality in educational opportunity. However, when school effects and sorting are also accounted for, family effects are not decisive and contribute only around 15% to the variance in outcomes. Indeed, unlike the findings of the neighborhood-effects literature (e.g., Solon, 2000), we find the combination of school and sorting components is generally larger than the standalone contribution of households. This means that conventional upper-bound estimates of the contribution of household factors (e.g., as captured by simple sibling correlations) may well overstate the unique contribution of family circumstances. Despite low average learning outcomes, it also indicates that variation in school quality is substantial and that positive sorting (matching) between households and schools aggravates extant learning inequalities. This conclusion is supported by evidence of substantial spatial heterogeneity in the variance components, in which regional differences in poverty and parental education play a role. Pulling these findings together, we find that inequality in educational opportunity is substantial, accounting for almost half of all variation in test scores. However, given the importance of schools and sorting within this total, it follows that educational (school) reforms that alter the distribution of school quality, such as via the allocation of teachers across schools, can enhance opportunities for the most disadvantaged.

## References

- Abowd, J.M., Creecy, R.H. and Kramarz, F. (2002). Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. *Longitudinal Employer-Household Dynamics Technical Papers 2002-06*, Center for Economic Studies, U.S. Census Bureau.
- Ahn, S. and Fessler, J.A. (2003). Standard errors of mean, variance, and standard deviation estimators. Technical report, University of Michigan, EECS Department Paper.
- Andrews, M.J., Gill, L., Schank, T. and Upward, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3):673–697.
- (2012). High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, 117(3):824–827.
- Azam, M. and Kingdon, G.G. (2015). Assessing teacher quality in India. *Journal of Development Economics*, 117:74–83.
- Bates, D.M. (2010). *Lme4: Mixed-Effects Modeling with R*. New York: Springer.
- Behrman, J.R. and Birdsall, N. (1983). The quality of schooling: quantity alone is misleading. *The American Economic Review*, 73(5):928–946.
- Behrman, J.R., Gaviria, A. and Székely, M. (2001). Intergenerational mobility in Latin America. *Economía*, 2(1):1–31.
- Behrman, J.R., Rosenzweig, M.R., and Taubman, P. (1994). Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment. *Journal of Political Economy*, 102(6): 1131-1174.
- Behrman, J.R. and Rosenzweig, M.R. (2004). Returns to birthweight. *Review of Economics and Statistics*, 86(2):586–601.
- Björklund, A. and Salvanes, K.G. (2011). Education and family background: Mechanisms and policies. In E. Hanushek, S. Machin and L. Wößmann (Eds.), *Handbook of the Economics of Education*, volume 3, chapter 3, pp. 201–247. Elsevier.
- Carneiro, P. (2008). Equality of opportunity and educational achievement in Portugal. *Portuguese Economic Journal*, 7(1):17–41.
- Chirwa, E.D., Griffiths, P.L., Maleta, K., Norris, S.A. and Cameron, N. (2014). Multi-level modelling of longitudinal child growth data from the Birth-to-Twenty Cohort: a comparison of growth models. *Annals of Human Biology*, 41(2):168–179.
- Combes, P.P., Duranton, G. and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742.
- Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, 27(3): 79-102.
- Correia, S. (2017). *reghdfe*: Stata module for linear and instrumental-variable/GMM

regression absorbing multiple levels of fixed effects. Technical Report Statistical Software Components s457874, Boston College Department of Economics. URL <https://EconPapers.repec.org/RePEc:boc:bocode:s457874>.

Dabalen, Andrew, Ambar Narayan, Jaime Saavedra-Chanduvi, and Alejandro Hoyos Suarez, with Ana Abras and Sailesh Tiwari (2015). *Do African Children Have an Equal Chance? A Human Opportunity Report for Sub-Saharan Africa*. Directions in Development. Washington, DC: World Bank. doi:10.1596/978-1-4648-0332-1.

Dang, H.-A. and Glewwe, P. (2018). "Well Begun, But Aiming Higher: A Review of Vietnam's Education Trends in the Past 20 Years and Emerging Challenges". *Journal of Development Studies*, 54(7): 1171-1195.

Davidoff, T. (2005). Income sorting: Measurement and decomposition. *Journal of Urban Economics*, 58(2):289–303.

Ferreira, F.H. and Gignoux, J. (2014). The measurement of educational inequality: Achievement and opportunity. *The World Bank Economic Review*, 28(2):210–246.

Ferreira, F. H., Molinas Vega, J. R., Paes de Barros, R., & Saavedra Chanduvi, J. (2008). *Measuring inequality of opportunities in Latin America and the Caribbean*. The World Bank.

Freeman, R.B. and Viarengo, M. (2014). School and family effects on educational outcomes across countries. *Economic Policy*, 29(79):395–446.

Gaure, S. (2014). Correlation bias correction in two-way fixed effects linear regression. *Stat*, 3(1):379–390.

Gibbons, S., Overman, H.G. and Pelkonen, P. (2014). Area disparities in Britain: Understanding the contribution of people vs. place through variance decompositions. *Oxford Bulletin of Economics And Statistics*, 76(5):745–763.

Glewwe, P. and Muralidharan, K. (2016). 'Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications', in Hanushek, E. A., Machin, S. and Woessmann, L. (eds.), *Handbook of Economics of Education*, Vol. 5, Amsterdam: North-Holland.

Guimarães, P. and Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. *Stata Journal*, 10(4):628.

Gurka, M.J., Edwards, L.J. and Muller, K.E. (2011). Avoiding bias in mixed model inference for fixed effects. *Statistics in Medicine*, 30(22):2696–2707.

Hanushek, E. and Rivkin, S. (2006). Teacher quality. *Handbook of the Economics of Education*, 2:1051–1078.

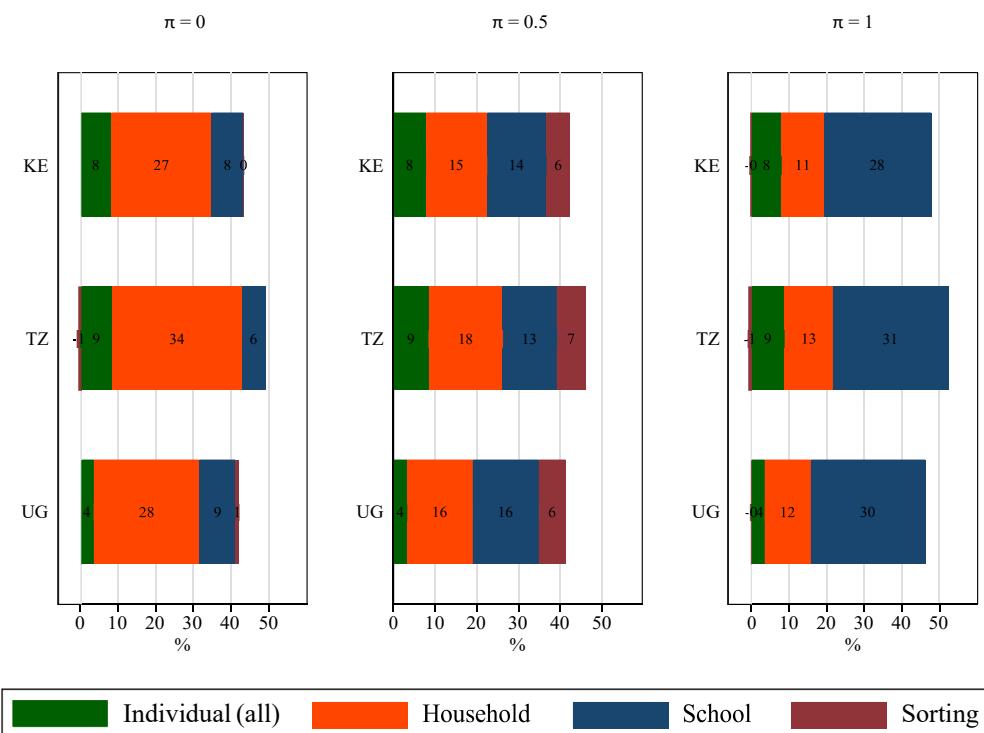
Hanushek, E.A. and Rivkin, S.G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics*, 4(1):131–157.

Hanushek, E. and Wößmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal*, 116(510):C63–C76.

- Hanushek, E. and Yilmaz, K. (2007). The complementarity of Tiebout and Alonso. *Journal of Housing Economics*, 16(2):243–261.
- Jackson, C.K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *Journal of Labor Economics*, 27(2):213–256.
- (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics*, 95(4):1096–1116.
- Jones, S., Schipper, Y., Ruto, S. and Rajani, R. (2014). Can your child read and count? Measuring learning outcomes in East Africa. *Journal of African Economies*, 23(5):643–672.
- Koedel, C., Mihaly, K. and Rockoff, J.E. (2015). Value-added modeling: A review. *Economics of Education Review*, 47:180–195.
- Lindahl, L. (2011). A comparison of family and neighborhood effects on grades, test scores, educational attainment and income – evidence from Sweden. *The Journal of Economic Inequality*, 9(2):207–226.
- Mazumder, B. (2008). Sibling similarities and economic inequality in the US. *Journal of Population Economics*, 21(3):685–701.
- (2011). Family and community influences on health and socioeconomic status: sibling correlations over the life course. *The B.E. Journal of Economic Analysis & Policy*, 11(3):Article 1.
- Mittag, N. (2012). New methods to estimate models with large sets of fixed effects with an application to matched employer-employee data from germany. FDZ Methodenreport 201201, Institute for Employment Research, Nuremberg, Germany.
- 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.
- Pop-Eleches, C. and Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *The American Economic Review*, 103(4):1289–1324.
- Pritchett, L. and Viarengo, M. (2015). Does public sector control reduce variance in school quality? *Education Economics*, 23(5):557–576.
- Raaum, O., Salvanes, K.G. and Sørensen, E.Ø. (2006). The neighbourhood is not what it used to be. *The Economic Journal*, 116(508):200–222.
- Rivkin, S.G., Hanushek, E.A. and Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458.
- Roemer, J.E. (1996). *Theories of distributive justice*. Harvard University Press: Cambridge MA.
- (2002). Equality of opportunity: A progress report. *Social Choice and Welfare*, 19(2):455–471.
- Sandefur, J. (2018). Internationally comparable mathematics scores for fourteen African

- countries. *Economics of Education Review*, 62: 267-286.
- Sass, T.R., Hannaway, J., Xu, Z., Figlio, D.N. and Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, 72(2):104–122.
- Schütz, G., Ursprung, H.W. and Wößmann, L. (2008). Education policy and equality of opportunity. *Kyklos*, 61(2):279–308.
- Sen, A. (2002). Why health equity? *Journal of Health Economics*, 11:659–666.
- Sen, A. (1985). A sociological approach to the measurement of poverty: a reply to Professor Peter Townsend. *Oxford Economic Papers*, 37(4): 669-676.
- Smyth, G.K. (1996). Partitioned algorithms for maximum likelihood and other non-linear estimation. *Statistics and Computing*, 6(3):201–216.
- Solon, G., Page, M.E. and Duncan, G.J. (2000). Correlations between neighboring children in their subsequent educational attainment. *Review of Economics and Statistics*, 82(3):383– 392.
- Stanek, E.J., Well, A. and Ockene, I. (1999). Why not routinely use best linear unbiased predictors (BLUPs) as estimates of cholesterol, per cent fat from Kcal and physical activity? *Statistics in Medicine*, 18(21):2943–2959.
- Uwezo (2012). Are our children learning? literacy and numeracy across East Africa. Technical report, Uwezo. URL [www.uwezo.net/wp-content/uploads/2012/08/RO\\_2012\\_UwezoEastAfricaREport.pdf](http://www.uwezo.net/wp-content/uploads/2012/08/RO_2012_UwezoEastAfricaREport.pdf).
- Walsh, P. (2009). Effects of school choice on the margin: The cream is already skimmed. *Economics of Education Review*, 28(2):227–236.
- Van de Werfhorst, H.G. and Mijs, J.J. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36:407–428.

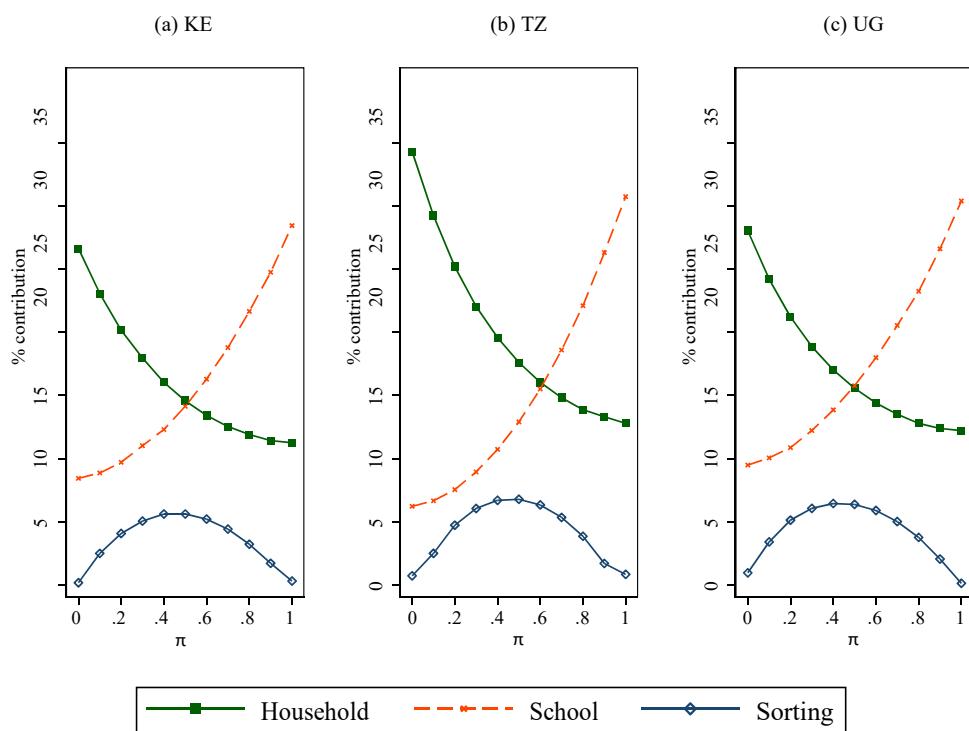
**Figure 1: Relative variance shares, by estimator**



Note: bars indicate relative shares reported in Table 4 for different models/estimators; component ‘individual (all)’ aggregates the three components including the observed individual characteristics; ‘sorting’ is the household-school covariance term; KE is Kenya; TZ is Tanzania (mainland); and UG is Uganda.

Source: own calculations.

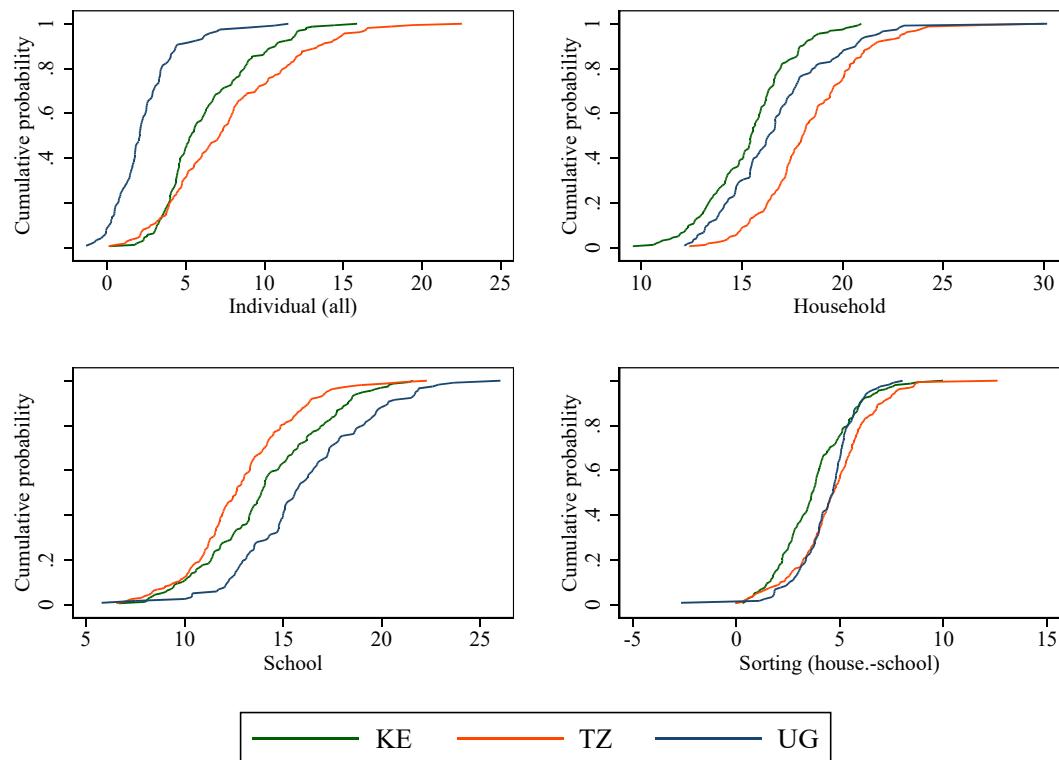
**Figure 2: Relative variance shares, by values of  $\pi$**



Note: ‘sorting’ is the household-school covariance term; KE is Kenya; TZ is Tanzania (mainland); and UG is Uganda.

Source: own calculations.

**Figure 3: Cumulative distribution functions of relative variance shares, by country**



Note: sub-figures show cumulative distribution functions (by country) for the relative variance shares calculated at the district-level; component ‘individual (all)’ aggregates the three components including the observed individual characteristics; KE is Kenya; TZ is Tanzania (mainland); and UG is Uganda.

Source: own calculations.

**Table 1: Alternative test score data-generating processes**

Model	Score level	Score variance	Description
1 Restricted linear	$h_j + s_k + e_{ijk}$	$\sigma_h^2 + \sigma_s^2 + \sigma_e^2$	Independent households & schools
2 Unrestricted linear	$h_j + s_k + e_{ijk}$	$\sigma_h^2 + \sigma_s^2 + 2\Sigma_{hs} + \sigma_e^2$	Correlated household & school factors
3 Household upper bound	$(1 + \gamma)h_j + v_k + e_{ijk}$	$(1 + \gamma)^2\sigma_h^2 + \sigma_v^2 + \sigma_e^2$	Households dominate school effects
4 School upper bound	$\omega_j + (1 + \theta)s_k + e_{ijk}$	$\sigma_\omega^2 + (1 + \theta)^2\sigma_s^2 + \sigma_e^2$	Schools dominate household effects

**Table 2: Description of synthetic test scores, by country & region**

		(1) Raw means		(2) Age std.ized		(3) Normalized	
Country & Region		Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
KE	Central	4.99	(1.52)	0.30	(0.67)	0.37	(0.73)
	Coast	4.25	(1.98)	-0.13	(0.93)	-0.10	(1.02)
	Eastern	4.51	(1.84)	-0.00	(0.87)	0.04	(0.95)
	North Eastern	3.67	(2.12)	-0.47	(1.23)	-0.48	(1.35)
	Nyanza	4.39	(1.90)	-0.05	(0.83)	-0.02	(0.91)
	Rift Valley	4.36	(1.95)	-0.06	(0.97)	-0.03	(1.06)
	Western	4.23	(1.98)	-0.16	(0.89)	-0.13	(0.98)
All		4.42	(1.91)	-0.04	(0.91)	-0.00	(1.00)
TZ	Arusha	3.69	(1.78)	0.15	(0.84)	0.20	(0.97)
	Dar Es Salaam	3.98	(1.64)	0.29	(0.76)	0.36	(0.88)
	Iringa	3.41	(1.83)	-0.00	(0.85)	0.03	(0.99)
	Kagera	3.24	(1.86)	-0.10	(0.87)	-0.09	(1.00)
	Kigoma	2.92	(1.90)	-0.25	(0.88)	-0.26	(1.02)
	Ruvuma	3.36	(1.79)	-0.03	(0.81)	-0.01	(0.94)
	Singida	3.47	(1.83)	0.03	(0.83)	0.07	(0.96)
	Tabora	2.94	(1.95)	-0.25	(0.91)	-0.26	(1.06)
	Tanga	3.43	(1.83)	0.01	(0.83)	0.04	(0.96)
	All	3.37	(1.86)	-0.03	(0.86)	-0.00	(1.00)
UG	Central	3.43	(1.87)	0.27	(0.89)	0.31	(1.00)
	Eastern	2.78	(1.89)	-0.15	(0.84)	-0.16	(0.94)
	Northern	2.57	(1.90)	-0.28	(0.88)	-0.30	(0.98)
	Western	3.13	(1.91)	0.06	(0.88)	0.08	(0.98)
	All	3.00	(1.92)	-0.01	(0.90)	-0.00	(1.00)

Note: synthetic test scores combine achievement in literacy and numeracy, as described in the text; ‘age std.’ scores are standardized within each age group (for each survey round and country) for the reference group defined as all children who are currently enrolled or have completed primary school; ‘normalized’ centers the age-standardized scores to have a mean of zero and standard deviation of one for the reference group in each country; removes mean level differences between districts; survey rounds are pooled; KE is Kenya; TZ is Tanzania (mainland); UG is Uganda; regions in Tanzania and Kenya are aggregated for clarity of presentation (for details see Appendix C).

Source: own calculations.

**Table 3: Descriptive sample statistics, by country & region**

Country & Region		Index count								
		<i>i</i>	<i>j</i>	<i>k</i>	Age	Female	SES	Attend	Grade	
KE	Central	37,870	15,194	6,353	11.2	50.6	62.5	95.5	5.1	
	Coast	49,921	17,013	5,084	11.1	49.9	-17.8	87.6	4.0	
	Eastern	84,133	30,596	10,160	11.2	50.0	-23.6	93.5	4.6	
	North Eastern	54,132	16,605	3,501	11.0	43.4	-61.7	80.7	3.6	
	Nyanza	76,471	27,014	8,410	11.2	49.5	-18.4	92.3	4.5	
	Rift Valley	169,672	58,737	17,095	11.1	49.0	-13.3	90.3	4.3	
	Western	83,720	28,392	8,296	11.2	50.0	-10.8	92.9	4.4	
All		555,919	193,551	58,899	11.1	49.4	-6.6	91.5	4.5	
TZ	Arusha	52,883	19,979	6,592	11.6	48.8	10.8	88.5	4.4	
	Dar Es Salaam	23,025	8,948	3,392	11.7	51.1	68.2	91.0	4.4	
	Iringa	48,617	19,414	6,981	11.6	49.9	-7.8	85.8	4.0	
	Kagera	50,476	17,815	5,921	11.6	49.6	-14.1	82.6	3.7	
	Kigoma	32,462	11,801	3,728	11.6	49.2	-28.6	80.8	3.6	
	Ruvuma	29,867	12,233	4,920	11.7	49.5	-16.0	86.6	4.1	
	Singida	32,269	12,379	4,168	11.6	49.7	-7.7	85.8	4.2	
	Tabora	49,936	17,289	5,311	11.5	49.1	-19.6	78.5	3.5	
	Tanga	39,928	15,071	5,010	11.6	48.6	-8.9	86.3	3.9	
	All	359,463	134,929	46,023	11.6	49.5	-3.1	84.8	4.0	
UG	Central	64,077	20,650	6,796	11.0	49.8	51.9	92.5	3.8	
	Eastern	120,142	36,183	9,762	11.1	49.2	-21.9	94.7	3.7	
	Northern	102,723	32,629	8,588	11.2	47.9	-38.5	86.3	3.2	
	Western	75,186	24,955	7,593	11.1	50.0	-12.8	91.7	3.5	
	All	362,128	114,417	32,739	11.1	49.3	-2.4	91.6	3.6	

Note: regions in Tanzania and Kenya are aggregated for clarity of presentation (see Appendix C); KE is Kenya; TZ is Tanzania (mainland); UG is Uganda; *i,j,k* refer to the number of unique observations for the individual, household and school-grade effects respectively; remaining columns are regional means (age, highest grade) or proportions; survey rounds are pooled.

Source: own calculations.

**Table 4: Unconditional absolute and relative variance contributions,  
alternative choices of  $\pi$**

	$\pi$	Absolute			Relative		
		Kenya	Tanzania	Uganda	Kenya	Tanzania	Uganda
Household	0	0.57	0.64	0.57	32.42	40.61	32.14
	0.5	0.41	0.45	0.42	17.11	20.7	17.63
	1	0.36	0.39	0.37	12.99	15.04	13.53
School	0	0.28	0.24	0.27	8.11	5.92	7.43
	0.5	0.38	0.36	0.36	14.37	13.24	13.25
	1	0.57	0.59	0.54	32.08	34.38	29.64
Sorting	0	-0.03	-0.11	0.04	-0.09	-1.14	0.18
	0.5	0.27	0.29	0.27	7.46	8.38	7.3
	1	-0.03	-0.08	0.05	-0.11	-0.71	0.25
Residual	.	0.77	0.74	0.77	58.55	54.53	59.56
Total	.	1.00	1.00	1.00	100.00	100.00	100.00

Note: following the logic of equations (4) and (5), but excluding the individual controls, the table sets out the absolute and relative variance contributions attributable to each component of the test score, reported in standard deviation units and percent respectively; different initializations of the partitioned iterative algorithm are indicated by column  $\pi$ ; standard errors not shown, but available on request.

Source: own calculations.

**Table 5: Conditional absolute and relative variance contributions, alternative choices of  $\pi$**

	$\pi$	Absolute			Relative		
		Kenya	Tanzania	Uganda	Kenya	Tanzania	Uganda
Individual	0	0.29	0.29	0.19	8.14	8.59	3.58
	0.5	0.28	0.29	0.19	7.91	8.65	3.51
	1	0.28	0.30	0.19	8.09	8.90	3.62
Household	0	0.52	0.59	0.53	26.63	34.31	28.03
	0.5	0.38	0.42	0.39	14.58	17.61	15.57
	1	0.34	0.36	0.35	11.26	12.80	12.22
School	0	0.29	0.25	0.31	8.44	6.23	9.49
	0.5	0.38	0.36	0.40	14.14	12.91	15.77
	1	0.53	0.55	0.55	28.45	30.74	30.39
Sorting	0	0.04	-0.08	0.10	0.18	-0.72	0.96
	0.5	0.24	0.26	0.25	5.62	6.78	6.38
	1	-0.06	-0.09	-0.04	-0.31	-0.84	-0.13
Residual	.	0.75	0.72	0.75	55.62	51.34	56.87
Total	.	1.00	1.00	1.00	100.00	100.00	100.00

Note: following equations (4) and (5), the table sets out the absolute and relative variance contributions attributable to each component of the test score, reported in standard deviation units and percent respectively; different initializations of the partitioned iterative algorithm are indicated by column  $\pi$ ; ‘individual’ component aggregates all observed individual effect components; full details found in Appendix Tables B4-B6.

Source: own calculations.

**Table 6: Absolute variance contributions (in s.d. units), by country & region**

		UBH	UBS	$\pi = 0.5$						
		House.	School	Indiv.	House.	School	Sorting	Resid.	Score	Correl.
		$\sigma_h^2$	$\sigma_s^2$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$	$\sigma_t^2$	$\rho_{hs}$
KE	Central	0.38	0.34	0.15	0.30	0.26	0.11	0.59	0.73	0.08
	Coast	0.54	0.50	0.31	0.39	0.38	0.24	0.77	1.02	0.19
	Eastern	0.50	0.49	0.23	0.37	0.37	0.21	0.73	0.95	0.17
	North Eastern	0.72	0.71	0.44	0.49	0.46	0.35	1.03	1.35	0.27
	Nyanza	0.46	0.46	0.22	0.35	0.37	0.16	0.70	0.91	0.10
	Rift Valley	0.57	0.54	0.33	0.40	0.39	0.27	0.80	1.06	0.23
	Western	0.51	0.49	0.20	0.40	0.39	0.18	0.76	0.98	0.11
	All	0.53	0.51	0.28	0.38	0.38	0.24	0.76	1.00	0.20
TZ	Arusha	0.59	0.54	0.22	0.42	0.35	0.27	0.72	0.97	0.25
	Dar Es Salaam	0.51	0.44	0.24	0.37	0.31	0.20	0.67	0.88	0.19
	Iringa	0.58	0.54	0.29	0.41	0.36	0.26	0.73	0.99	0.23
	Kagera	0.58	0.50	0.31	0.44	0.35	0.24	0.74	1.00	0.19
	Kigoma	0.60	0.53	0.32	0.45	0.37	0.25	0.74	1.02	0.19
	Ruvuma	0.54	0.50	0.23	0.40	0.35	0.22	0.71	0.94	0.17
	Singida	0.54	0.50	0.27	0.40	0.35	0.23	0.73	0.96	0.19
	Tabora	0.61	0.54	0.34	0.44	0.37	0.26	0.77	1.06	0.21
	Tanga	0.55	0.52	0.26	0.40	0.36	0.24	0.71	0.96	0.19
	All	0.58	0.53	0.29	0.42	0.36	0.26	0.74	1.00	0.22
UG	Central	0.55	0.50	0.19	0.39	0.38	0.24	0.77	1.00	0.18
	Eastern	0.51	0.49	0.12	0.38	0.37	0.22	0.73	0.94	0.17
	Northern	0.51	0.53	0.17	0.37	0.40	0.22	0.76	0.98	0.16
	Western	0.56	0.53	0.15	0.41	0.40	0.23	0.75	0.98	0.17
	All	0.56	0.53	0.19	0.39	0.40	0.25	0.77	1.00	0.20

Note: top-level column indicates the model, where UBH is the upper bound household model, UBS is the upper bound school model and  $\pi = 0.5$  is the (preferred) PIA estimator;  $\sigma_{a+}^2$  is the aggregate of all observed individual effect components; all other components are as before; values are reported in standard deviation units; regions in Tanzania and Kenya are aggregated for clarity of presentation (see Appendix C); KE is Kenya; TZ is Tanzania (mainland); UG is Uganda.

Source: own calculations.

**Table 7: Relative variance shares, by country & region**

		UBH	UBS	$\pi = 0.5$					
		House.	School	Indiv.	House.	School	Sorting	Resid.	Score
		$\sigma_h^2$	$\sigma_s^2$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$	$\sigma_t^2$
KE	Central	26.6	21.9	4.5	16.4	12.5	2.3	64.4	100.0
	Coast	27.9	24.4	9.1	14.8	13.9	5.4	56.7	100.0
	Eastern	28.2	26.6	6.1	15.0	14.9	5.0	58.9	100.0
	North Eastern	28.8	27.5	10.5	13.4	11.4	6.7	58.0	100.0
	Nyanza	25.7	25.8	5.7	15.0	16.7	3.3	59.4	100.0
	Rift Valley	28.7	26.2	9.7	14.1	13.4	6.4	56.4	100.0
	Western	27.4	25.1	4.2	16.4	15.6	3.6	60.2	100.0
All		28.2	25.8	7.9	14.6	14.1	5.6	57.7	100.0
TZ	Arusha	37.4	30.8	5.1	18.8	13.2	8.0	55.0	100.0
	Dar Es Salaam	32.9	25.3	7.1	17.2	12.1	5.4	58.1	100.0
	Iringa	34.4	29.5	8.8	17.2	13.4	6.9	53.8	100.0
	Kagera	33.5	25.0	9.5	18.9	11.9	5.8	54.0	100.0
	Kigoma	34.1	27.1	9.9	19.2	12.9	5.9	52.0	100.0
	Ruvuma	33.5	28.7	5.9	17.9	14.0	5.4	56.9	100.0
	Singida	31.8	26.8	7.9	16.9	12.9	5.6	56.7	100.0
	Tabora	32.8	26.2	10.4	17.5	12.1	6.2	53.8	100.0
	Tanga	33.4	29.5	7.6	17.7	14.5	6.2	54.1	100.0
	All	34.1	27.9	8.7	17.6	12.9	6.8	54.0	100.0
UG	Central	30.5	25.4	3.8	15.7	14.9	5.6	60.0	100.0
	Eastern	30.1	27.1	1.6	16.4	15.5	5.4	61.2	100.0
	Northern	27.6	28.9	3.1	14.6	16.4	5.0	60.9	100.0
	Western	32.4	29.7	2.5	17.3	16.4	5.7	58.0	100.0
	All	31.0	28.1	3.5	15.6	15.8	6.4	58.8	100.0

Note: top-level column indicates the model, where UBH is the upper bound household model, UBS is the upper-bound school model and  $\pi = 0.5$  is the (preferred) PIA estimator;  $\sigma_{a+}^2$  is the aggregate of all observed individual effect components; all other components are as before; regions in Tanzania and Kenya are aggregated for clarity of presentation (see Appendix C); KE is Kenya; TZ is Tanzania (mainland); UG is Uganda.

Source: own calculations.

**Table 8: Analysis of systematic patterns in variance components, by district**

	(I) Absolute shares					(II) Relative shares				
	Indiv.	House.	School	Sorting	Resid.	Indiv.	House.	School	Sorting	Resid.
	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$
Female	-0.00 (0.21)	-0.13 (0.16)	-0.33** (0.14)	-0.32** (0.14)	-0.49** (0.22)	1.31 (5.63)	10.29 (8.95)	-9.45 (8.97)	-6.11 (4.70)	3.96 (13.16)
Never enrolled	0.12 (0.28)	0.19 (0.19)	-0.30 (0.19)	-0.20 (0.18)	0.36 (0.25)	-4.84 (10.75)	14.15 (9.16)	-26.92*** (9.97)	-10.00* (6.01)	27.61 (17.09)
Current attending	-0.78*** (0.20)	-0.06 (0.14)	0.07 (0.15)	-0.12 (0.16)	-0.08 (0.17)	-33.32*** (8.29)	8.35 (8.01)	12.96 (8.64)	-2.18 (5.85)	14.19 (14.73)
Highest grade	-0.01 (0.02)	-0.00 (0.01)	-0.06*** (0.01)	-0.02 (0.01)	0.01 (0.02)	-0.77 (0.48)	0.91 (0.68)	-4.21*** (0.70)	-0.49 (0.37)	4.57*** (0.99)
Attends private sch.	0.30*** (0.05)	0.16*** (0.04)	0.02 (0.04)	0.07** (0.03)	0.18*** (0.06)	6.91*** (1.44)	4.58** (2.24)	-7.15*** (2.30)	0.33 (1.04)	-4.68* (2.66)
Private $\times$ grade	-0.20** (0.08)	-0.04 (0.07)	0.10 (0.06)	-0.04 (0.05)	0.06 (0.11)	-9.69*** (2.36)	-3.94 (4.09)	8.50*** (2.71)	-1.09 (1.63)	6.22 (4.43)
SES index	-0.00 (0.01)	-0.01 (0.01)	-0.02 (0.01)	-0.01 (0.01)	-0.01 (0.02)	0.86* (0.45)	-0.21 (0.75)	-0.96* (0.58)	-0.52 (0.34)	0.83 (0.95)
Mother no schooling	0.15*** (0.02)	0.07*** (0.02)	0.07*** (0.02)	0.15*** (0.02)	0.25*** (0.03)	2.57*** (0.91)	-3.37*** (1.08)	-3.41*** (1.31)	2.99*** (0.72)	1.22 (1.78)
Test score (percentile)	0.07 (0.11)	-0.24*** (0.07)	-0.04 (0.07)	-0.02 (0.07)	-0.31*** (0.11)	3.29 (3.73)	-7.88* (4.60)	8.42** (3.98)	2.49 (2.43)	-6.32 (6.91)
Tanzania	0.02*** (0.01)	0.04*** (0.00)	-0.01*** (0.00)	0.03*** (0.01)	-0.02** (0.01)	1.48*** (0.24)	2.91*** (0.28)	-1.52*** (0.31)	1.28*** (0.20)	-4.15*** (0.45)
Uganda	-0.09*** (0.01)	0.02** (0.01)	0.02*** (0.01)	0.02*** (0.00)	0.01 (0.01)	-3.38*** (0.18)	0.91*** (0.32)	1.13*** (0.34)	0.84*** (0.17)	0.51 (0.42)
Constant	0.23*** (0.00)	0.37*** (0.00)	0.35*** (0.00)	0.18*** (0.00)	0.72*** (0.01)	6.03*** (0.12)	15.67*** (0.19)	14.25*** (0.23)	3.65*** (0.12)	60.40*** (0.27)
Obs.	434	434	434	434	434	434	434	434	434	434

	(I) Absolute shares					(II) Relative shares				
	Indiv.	House.	School	Sorting	Resid.	Indiv.	House.	School	Sorting	Resid.
	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$
R <sup>2</sup> (adj.)	0.79	0.52	0.50	0.53	0.69	0.79	0.31	0.46	0.33	0.41
RMSE	0.05	0.04	0.04	0.04	0.05	1.56	2.10	2.22	1.35	3.05

Note: the table sets out OLS regression results for the conditional correlates of the district-level variance component estimates, based on the preferred estimator  $\pi=0.5$ ; dependent variable is indicated in the columns, where the absolute share is in standard deviation units; robust standard errors are reported in parentheses. Source: own calculations.

## SUPPLEMENTARY MATERIAL

### A Technical details on empirical methods

As noted in the main text, two primary challenges arise when estimating high-dimensional two-way fixed effects. First, to deal with bias from measurement error, empirical Bayes shrinkage approaches are often employed. This involves adjusting each estimated effect toward a common prior, where the adjustment factor is proportional to the estimated noise-to-signal ratio in the original estimates. Following Stanek et al. (1999), and to ensure consistency across the various methods deployed, we apply the approach typically used to adjust predictions of the random effects. Concretely, for a given estimated fixed effect (e.g.,  $\hat{h}_j$ ) we shrink it toward a global mean as follows:

$$\tilde{h}_j = \bar{h}_j + \left( \frac{\sigma_{\hat{h}}^2}{\sigma_{\hat{h}}^2 + \sigma_{\hat{\varepsilon}}^2 / (N_j - 1)} \right) (\hat{h}_j - \bar{h}_j) \quad (\text{A1})$$

where  $N_j - 1$  is the effective degrees of freedom available to estimate each of the  $j$  effects;  $\sigma_{\hat{h}}^2$  is the variance of the estimated effect;  $\sigma_{\hat{\varepsilon}}^2$  is the estimated residual variance; and  $\bar{h}_j$  is the population mean, typically zero under conventional normalization restrictions.

The second challenge is the (mechanical) negative covariance bias of the two estimated fixed effects. While this may be partially mitigated by the aforementioned empirical Bayes shrinkage, since this procedure simply modifies both sets of effects by a (varying) scalar bound between zero and one, it should have little effect on their correlation. A closer look at the nature of this bias indicates it may be driven (at least in part) by how the fixed effects are initialized under the iterative algorithm. While the latent fixed effects are adjusted iteratively based on model residuals, the assumed starting values for the two effects fundamentally determine their final estimated levels and variance shares. Referring to the unrestricted linear model (without additional controls), the regression specification (estimated via simple OLS) used in the first step of the iterative algorithm is just:

$$t_{ijk} = \delta_{h(1)} \hat{h}_{j(0)} + \delta_{s(1)} \hat{s}_{k(0)} + e_{ijk(1)} \quad (\text{A2})$$

where  $\delta_h, \delta_s$  are parameters to be estimated;  $\hat{h}, \hat{s}$  are initial estimates for the fixed effects (see below); and the numeric indexes in the subscripts represent the iteration number. In the second step, the model to be estimated is updated using the residual from equation (A2), as:

$$t_{ijk} = \delta_{h(2)} \left( \hat{\delta}_{h(1)} \hat{h}_{J(0)} + \frac{1}{N_j} \sum_{i|J=j} \hat{e}_{ij(1)} \right) + \delta_{s(2)} \left( \hat{\delta}_{s(1)} \hat{s}_{K(0)} + \frac{1}{N_k} \sum_{i|K=k} \hat{e}_{ik(1)} \right) + e_{ijk(2)} \quad (\text{A3})$$

from where the algorithm iterates until some convergence criterion is reached, such as when:  $|\sum \hat{e}_{ijk(n+1)}^2 - \sum \hat{e}_{ijk(n)}^2| \leq N\varepsilon$ . From this, the proposed starting values for the two fixed effects that enter equation (A2) appear fundamental. Typically, these are approximated using the group-specific means of the residuals taken from a (zero step) naïve model. Continuing with our simple case, without additional covariates, the general expression for these is:

$$\begin{aligned} \hat{h}_{j(0)} &= \frac{1}{N_j} \sum_{i|J=j} \left( t_{ij} - \frac{\pi}{N_k} \sum_{i|K=k} t_{ik} \right) \\ \hat{s}_{k(0)} &= \frac{1}{N_k} \sum_{i|K=k} \left( t_{ik} - \frac{1-\pi}{N_j} \sum_{i|J=j} t_{ij} \right) \end{aligned}$$

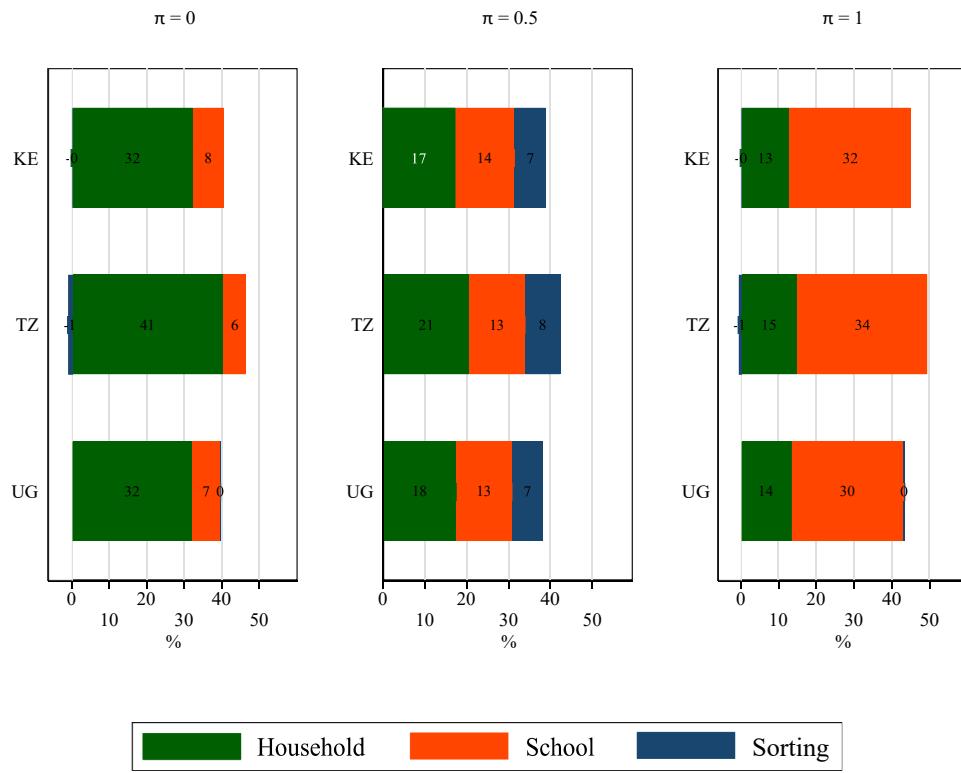
where  $\pi \in [0,1]$  serves as an initialization scalar that apportions the variation in  $t$  across the school and household effects. For instance, if  $\pi = 0$  then the observed variation in  $t$  is allocated primarily

to the initial estimate for the household fixed effect:  $\hat{h}_{j(0)} = (1/N_j) \sum_{ilj} t_{ij}$ ; in turn, the initial estimate for the school fixed effect captures only any residual variation in  $t$ , averaged by index  $k$ . The same initialization is also implicit when one of the fixed effects is initially 'swept out' of the regression, leaving the iterative adjustment to focus on the remaining effect (Guimaraes and Portugal 2010). This clarifies the aforementioned concern that any over-estimation (upward bias) of the initial values for one factor will be mechanically reflected by an under-estimation in the other (and *vice versa*). Indeed, since the assumed starting values are derived directly from the dependent variable (or residuals thereof), they always contain relevant information and effectively become locked-in as the algorithm proceeds – i.e., regression estimates for  $\hat{\delta}_{h(1)}$ ,  $\hat{\delta}_{s(1)}$  derived from equation (A3) should always be close to one.

These mechanics demonstrate that the initialization of the fixed effects embeds specific presumptions about how variation in the outcome is to be allocated across the fixed effects. Our working hypothesis is that this translates into specific assumptions about the form of the between-factor covariance. Specifically, as extreme choices for the initial values (e.g.,  $\pi = 0, \pi = 1$ ) treat the second effect as a residual term, the implicit assumption is that the two factors are orthogonal. Thus, these corner choices are expected to correspond to upper bound models in which one factor is dominant (Rows 3 and 4 of Table 1). A midpoint or agnostic choice ( $\pi = 0.5$ ), however, is likely to behave conversely to the extreme choices. By giving equal weight to both effects in the initialization they are no longer assumed *a priori* to be orthogonal, which would correspond to a case where sorting (between-factor covariance) is not ruled out from the outset.

## B Additional figures and tables

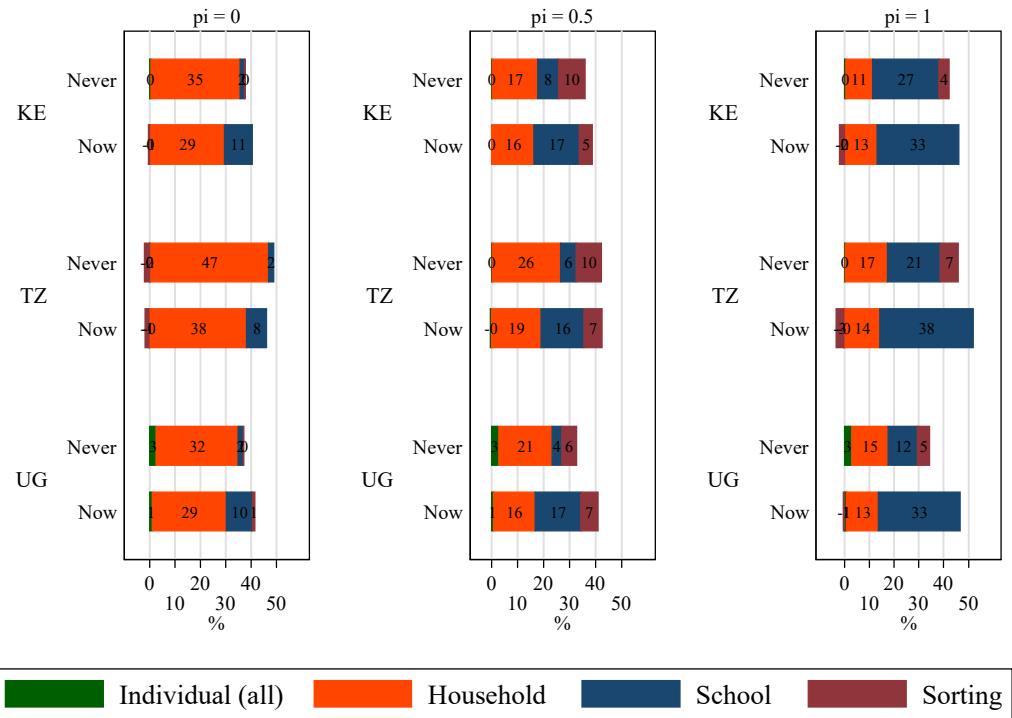
Figure B1: Relative unconditional variance shares, by estimator



Note: bars indicate relative variance contributions based on the same variance decomposition reported in Tables B4-B6 but without individual-specific controls; ‘sorting’ is the household-school covariance term; KE is Kenya; TZ is Tanzania (mainland); and UG is Uganda.

Source: own calculations.

Figure B2: Relative variance shares, by estimator and schooling status



Note: bars indicate relative variance contributions based on the same variance decomposition reported in Tables B4-B6 for all school-age children in the household either out of school ('never') or attending school ('now'); 'sorting' is the household-school covariance term; KE is Kenya; TZ is Tanzania (mainland); and UG is Uganda.

Source: own calculations.

Table B1: Regression results for alternative models/estimators, Kenya

	Naïve (1)	reghdfe (2)	$\pi=0$ (3)	$\pi=0.5$ (4)	$\pi=1$ (5)	UBH (6)	UBS (7)
Child is female	0.10 (0.00)	0.06 (0.00)	0.06 (0.00)	0.06 (0.00)	0.06 (0.00)	0.08 (0.00)	0.07 (0.00)
Oldest sib	0.06 (0.00)	-0.15 (0.00)	-0.15 (0.00)	-0.15 (0.00)	-0.15 (0.00)	0.04 (0.00)	-0.17 (0.00)
Never enrolled	-0.52 (0.02)	-0.15 (0.02)	-0.24 (0.02)	-0.22 (0.02)	-0.20 (0.02)	-0.24 (0.02)	-0.19 (0.02)
Currently enrolled	0.69 (0.02)	0.48 (0.02)	0.47 (0.02)	0.47 (0.02)	0.48 (0.02)	0.49 (0.02)	0.59 (0.02)
Attends private	0.35 (0.01)	0.13 (0.01)	0.14 (0.00)	0.15 (0.00)	0.16 (0.00)	0.09 (0.01)	0.22 (0.01)
Never $\times s_k$			-0.37 (0.02)	-0.21 (0.02)	-0.10 (0.01)		
Private $\times s_k$			-0.19 (0.01)	-0.16 (0.01)	-0.11 (0.01)		
Obs.	555,919	555,919	555,919	555,919	555,919	555,919	555,919
R <sup>2</sup>	0.13	0.72	0.72	0.72	0.72	0.64	0.48

Note: columns refer to different estimators/models; UBH is the household upper bound (excludes school effects); ‘Naïve’ is simple OLS excluding fixed effects; reghdfe reports results based on the (improved, accelerated) partitioned iterative algorithm due to Correia (2017) in which household effects enter first (using the Stata command of the same name); columns  $\pi = (0, 0.5, 1)$  are taken from the partitioned iterative algorithm set out in the text; UBH is the household upper bound model and UBS is the school upper bound models, in which household effects are proxied by observed characteristics (not shown); all reported coefficients are significantly different from zero; cluster robust standard errors reported in parentheses.

Source: own calculations.

Table B2: Regression results for alternative models/estimators, Tanzania

	Naïve (1)	reghdfe (2)	$\pi=0$ (3)	$\pi=0.5$ (4)	$\pi=1$ (5)	UBH (6)	UBS (7)
Child is female	0.08 (0.00)	0.04 (0.01)	0.05 (0.00)	0.04 (0.00)	0.05 (0.00)	0.06 (0.00)	0.05 (0.00)
Oldest sib	0.05 (0.01)	-0.13 (0.01)	-0.12 (0.00)	-0.13 (0.00)	-0.12 (0.00)	0.01 (0.00)	-0.14 (0.00)
Never enrolled	-0.29 (0.02)	0.02 (0.02)	-0.02 (0.01)	-0.06 (0.01)	-0.07 (0.01)	-0.06 (0.02)	-0.07 (0.02)
Currently enrolled	0.74 (0.02)	0.58 (0.02)	0.58 (0.01)	0.58 (0.01)	0.58 (0.01)	0.56 (0.02)	0.67 (0.01)
Attends private	0.28 (0.01)	0.10 (0.01)	0.10 (0.01)	0.10 (0.01)	0.11 (0.01)	0.11 (0.01)	0.09 (0.01)
Never $\times s_k$			-0.31 (0.02)	-0.35 (0.02)	-0.25 (0.01)		
Private $\times s_k$			0.02 (0.02)	-0.02 (0.02)	-0.02 (0.01)		
Obs.	359,463	359,463	359,463	359,463	359,463	359,463	359,463
R <sup>2</sup>	0.12	0.77	0.77	0.77	0.77	0.69	0.51

Note: columns refer to different estimators/models; UBH is the household upper bound (excludes school effects); ‘Naïve’ is simple OLS excluding fixed effects; reghdfe reports results based on the (improved, accelerated) partitioned iterative algorithm due to Correia (2017) in which household effects enter first (using the Stata command of the same name); columns  $\pi = (0, 0.5, 1)$  are taken from the partitioned iterative algorithm set out in the text; UBH is the household upper bound model and UBS is the school upper bound models, in which household effects are proxied by observed characteristics (not shown); all reported coefficients are significantly different from zero; cluster robust standard errors reported in parentheses.

Source: own calculations.

Table B3: Regression results for alternative models/estimators, Uganda

	Naïve (1)	reghdfe (2)	$\pi=0$ (3)	$\pi=0.5$ (4)	$\pi=1$ (5)	UBH (6)	UBS (7)
Child is female	0.06 (0.01)	0.02 (0.00)	0.03 (0.00)	0.02 (0.00)	0.02 (0.00)	0.04 (0.00)	0.04 (0.00)
Oldest sib	0.09 (0.01)	-0.22 (0.01)	-0.21 (0.00)	-0.21 (0.00)	-0.21 (0.00)	0.05 (0.01)	-0.23 (0.01)
Never enrolled	-0.16 (0.02)	0.07 (0.02)	-0.02 (0.02)	-0.06 (0.02)	-0.04 (0.02)	-0.01 (0.02)	0.06 (0.02)
Currently enrolled	0.48 (0.02)	0.35 (0.02)	0.34 (0.01)	0.34 (0.01)	0.35 (0.01)	0.39 (0.02)	0.43 (0.01)
Attends private	0.43 (0.01)	0.18 (0.01)	0.19 (0.00)	0.19 (0.00)	0.20 (0.00)	0.12 (0.01)	0.25 (0.01)
Never $\times s_k$			-0.48 (0.04)	-0.55 (0.03)	-0.43 (0.02)		
Private $\times s_k$			-0.03 (0.01)	-0.02 (0.01)	-0.01 (0.01)		
Obs.	362,128	362,128	362,128	362,128	362,128	362,128	362,128
R <sup>2</sup>	0.08	0.71	0.71	0.71	0.71	0.62	0.45

Note: columns refer to different estimators/models; UBH is the household upper bound (excludes school effects); ‘Naïve’ is simple OLS excluding fixed effects; reghdfe reports results based on the (improved, accelerated) partitioned iterative algorithm due to Correia (2017) in which household effects enter first (using the Stata command of the same name); columns  $\pi = (0, 0.5, 1)$  are taken from the partitioned iterative algorithm set out in the text; UBH is the household upper bound model and UBS is the school upper bound models, in which household effects are proxied by observed characteristics (not shown); all reported coefficients are significantly different from zero; cluster robust standard errors reported in parentheses.

Source: own calculations.

Table B4: Full variance decomposition for KE

	Absolute contributions (in s.d. units)						Relative shares (in %)																							
	(a) $\pi = 0$			(b) $\pi = 0.5$			(c) $\pi = 1$			UBH			MLM			(a) $\pi = 0$			(b) $\pi = 0.5$			(c) $\pi = 1$			UBH			MLM		
	$\pi = 0$	$\pi = 0.5$	$\pi = 1$	UBH	UBS	MLM	$\pi = 0$	$\pi = 0.5$	$\pi = 1$	UBH	UBS	MLM	$\pi = 0$	$\pi = 0.5$	$\pi = 1$	UBH	UBS	MLM	$\pi = 0$	$\pi = 0.5$	$\pi = 1$	UBH	UBS	MLM						
$\sigma_{ao}^2$	0.21 (0.00)	0.21 (0.00)	0.20 (0.00)	0.21 (0.00)	0.24 (0.00)	0.19 (0.00)	4.39 (0.06)	4.27 (0.06)	4.16 (0.05)	4.38 (0.06)	5.73 (0.06)	3.63 (0.05)																		
$\sigma_h^2$	0.52 (0.00)	0.38 (0.00)	0.34 (0.00)	0.53 (0.00)	0.29 (0.00)	0.37 (0.00)	26.63 (0.19)	14.58 (0.14)	11.26 (0.13)	28.22 (0.20)	8.54 (0.11)	13.79 (0.14)																		
$\sigma_s^2$	0.29 (0.00)	0.38 (0.00)	0.53 (0.00)	0.14 (0.00)	0.51 (0.00)	0.42 (0.00)	8.44 (0.18)	14.14 (0.23)	28.45 (0.33)	2.00 (0.09)	25.75 (0.31)	17.81 (0.26)																		
$2\Sigma_{hs}$	0.04 (0.00)	0.24 (0.00)	-0.06 (0.00)	0.12 (0.00)	0.16 (0.00)	0.25 (0.00)	0.18 (0.02)	5.62 (0.11)	-0.31 (0.03)	1.39 (0.06)	2.48 (0.07)	6.41 (0.12)																		
$2\Sigma_{ao_h}$	0.19 (0.00)	0.15 (0.00)	0.12 (0.00)	0.21 (0.00)	0.11 (0.00)	0.12 (0.00)	3.44 (0.06)	2.28 (0.05)	1.46 (0.04)	4.52 (0.07)	1.19 (0.03)	1.48 (0.04)																		
$2\Sigma_{ao_s}$	0.06 (0.00)	0.12 (0.00)	0.16 (0.00)	0.07 (0.00)	0.17 (0.00)	0.14 (0.00)	0.31 (0.02)	1.36 (0.04)	2.47 (0.06)	0.50 (0.03)	2.90 (0.07)	2.02 (0.05)																		
$\sigma_e^2$	0.75 (0.00)	0.76 (0.00)	0.72 (0.00)	0.77 (0.00)	0.73 (0.00)	0.74 (0.00)	56.61 (0.20)	57.75 (0.20)	52.51 (0.19)	58.98 (0.21)	53.40 (0.20)	54.86 (0.20)																		
$\sigma_t^2$	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	100.00 . . .	100.00 . . .	100.00 . . .	100.00 . . .	100.00 . . .	100.00 . . .																		
$\rho_{hs}$	0.01 (0.00)	0.20 (0.00)	-0.01 (0.00)	0.09 (0.00)	0.08 (0.00)	0.20 (0.00)	. . .	. . .	. . .	. . .	. . .	. . .																		

Note: following equations (3b) and (3c), the table sets out the absolute and relative variance shares attributable to each component of the test score; absolute shares are in standard deviation units; different models/estimators are indicated in the columns – (a) to (c) refer to results from a partitioned iterative algorithm for different choices of initialization scalar  $\pi$ , (d) is the household upper bound, (e) is the school upper bound, and (f) is a mixed linear model;  $\rho_{hs}$  is the estimated correlation coefficient between household and school effects; standard errors are reported in parentheses, calculated using the asymptotic approximation due to Ahn and Fessler (2003).

Source: own calculations.

Table B5: Full variance decomposition for TZ

	Absolute contributions (in s.d. units)						Relative shares (in %)											
	(a) $\pi = 0$			(b) $\pi = 0.5$			(c) $\pi = 1$			(d) UBH			(e) UBS			(f) MLM		
	(a)	(b)	(c)	(d)	(e)	(f)	(a)	(b)	(c)	(d)	(e)	(f)	(a)	(b)	(c)	(d)	(e)	(f)
$\sigma_{a^o}^2$	0.23	0.23	0.23	0.22	0.27	0.19	5.06	5.38	5.51	4.84	7.11	3.63						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)	(0.08)	(0.08)	(0.07)	(0.09)	(0.06)						
$\sigma_h^2$	0.59	0.42	0.36	0.58	0.30	0.43	34.31	17.61	12.80	34.11	9.25	18.55						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.26)	(0.19)	(0.16)	(0.26)	(0.14)	(0.19)						
$\sigma_s^2$	0.25	0.36	0.55	0.12	0.53	0.42	6.23	12.91	30.74	1.53	27.90	17.24						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.17)	(0.25)	(0.39)	(0.09)	(0.37)	(0.29)						
$2\Sigma_{hs}$	-0.08	0.26	-0.09	0.10	0.16	0.28	-0.72	6.78	-0.84	0.92	2.63	7.74						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.14)	(0.05)	(0.05)	(0.09)	(0.15)						
$2\Sigma_{a^o h}$	0.18	0.16	0.13	0.21	0.10	0.11	3.16	2.41	1.71	4.29	1.03	1.15						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.07)	(0.06)	(0.05)	(0.08)	(0.04)	(0.04)						
$2\Sigma_{a^o s}$	0.06	0.09	0.13	0.04	0.15	0.16	0.37	0.87	1.67	0.16	2.22	2.53						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.04)	(0.06)	(0.02)	(0.07)	(0.07)						
$\sigma_e^2$	0.72	0.74	0.70	0.74	0.71	0.70	51.59	54.04	48.39	54.15	49.85	49.17						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.24)	(0.25)	(0.23)	(0.25)	(0.24)	(0.23)						
$\sigma_t^2$	1.00	1.00	1.00	1.00	1.00	1.00	100.00	100.00	100.00	100.00	100.00	100.00						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	.	.	.	.	.	.						
$\rho_{hs}$	-0.02	0.22	-0.02	0.06	0.08	0.22	.	.	.	.	.	.						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	.	.	.	.	.	.						

Note: following equations (3b) and (3c), the table sets out the absolute and relative variance shares attributable to each component of the test score; absolute shares are in standard deviation units; different models/estimators are indicated in the columns – (a) to (c) refer to results from a partitioned iterative algorithm for different choices of initialization scalar  $\pi$ , (d) is the household upper bound, (e) is the school upper bound, and (f) is a mixed linear model;  $\rho_{hs}$  is the estimated correlation coefficient between household and school effects; standard errors are reported in parentheses, calculated using the asymptotic approximation due to Ahn and Fessler (2003).

Source: own calculations.

Table B6: Full variance decomposition for UG

	Absolute contributions (in s.d. units)						Relative shares (in %)											
	(a) $\pi = 0$			(b) $\pi = 0.5$			(c) $\pi = 1$			(d) UBH			(e) UBS			(f) MLM		
	(a)	(b)	(c)	(d)	(e)	(f)	(a)	(b)	(c)	(d)	(e)	(f)	(a)	(b)	(c)	(d)	(e)	(f)
$\sigma_{ao}^2$	0.17	0.18	0.18	0.13	0.20	0.19	2.94	3.12	3.13	1.75	4.05	3.59						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.06)	(0.06)	(0.06)	(0.04)	(0.07)	(0.06)						
$\sigma_h^2$	0.53	0.39	0.35	0.56	0.31	0.38	28.03	15.57	12.22	31.03	9.35	14.72						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.25)	(0.19)	(0.17)	(0.27)	(0.15)	(0.18)						
$\sigma_s^2$	0.31	0.40	0.55	0.14	0.53	0.44	9.49	15.77	30.39	2.03	28.11	19.44						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.25)	(0.32)	(0.45)	(0.12)	(0.43)	(0.36)						
$2\Sigma_{hs}$	0.10	0.25	-0.04	0.13	0.18	0.22	0.96	6.38	-0.13	1.73	3.15	4.78						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.06)	(0.16)	(0.02)	(0.08)	(0.11)	(0.14)						
$2\Sigma_{ao}h$	0.14	0.11	0.09	0.16	0.09	0.11	1.89	1.31	0.86	2.47	0.74	1.31						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.04)	(0.04)	(0.06)	(0.03)	(0.04)						
$2\Sigma_{aos}$	-0.11	-0.10	-0.06	0.04	-0.08	0.18	-1.25	-0.92	-0.37	0.19	-0.59	3.34						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.05)	(0.03)	(0.02)	(0.04)	(0.09)						
$\sigma_e^2$	0.76	0.77	0.73	0.78	0.74	0.73	57.94	58.77	53.90	60.79	55.20	52.82						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.25)	(0.25)	(0.24)	(0.26)	(0.25)	(0.24)						
$\sigma_t^2$	1.00	1.00	1.00	1.00	1.00	1.00	100.00	100.00	100.00	100.00	100.00	100.00						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	.	.	.	.	.	.						
$\rho_{hs}$	0.03	0.20	-0.00	0.11	0.10	0.14	.	.	.	.	.	.						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	.	.	.	.	.	.						

Note: following equations (3b) and (3c), the table sets out the absolute and relative variance shares attributable to each component of the test score; absolute shares are in standard deviation units; different models/estimators are indicated in the columns – (a) to (c) refer to results from a partitioned iterative algorithm for different choices of initialization scalar  $\pi$ , (d) is the household upper bound, (e) is the school upper bound, and (f) is a mixed linear model;  $\rho_{hs}$  is the estimated correlation coefficient between household and school effects; standard errors are reported in parentheses, calculated using the asymptotic approximation due to Ahn and Fessler (2003).

Source: own calculations.

Table B7: Variance contributions for Kenya, by sub-group

Strata	Level	Absolute (s.d. units)					Relative (in %)				
		$\sigma_t^2$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$
Female	0	1.02	0.21	0.39	0.38	0.25	4.21	14.49	13.92	5.79	61.59
	1	0.98	0.20	0.38	0.37	0.23	4.28	14.76	14.46	5.43	61.06
Age	6	0.98	0.29	0.38	0.39	0.21	9.08	15.38	15.66	4.83	55.05
	9	1.01	0.16	0.38	0.39	0.28	2.58	14.24	15.12	7.47	60.59
	12	1.01	0.15	0.38	0.33	0.26	2.36	14.15	10.80	6.68	66.01
	15	1.01	0.18	0.38	0.30	0.24	3.08	14.38	8.96	5.63	67.95
Grade level	0	1.34	0.15	0.54	0.38	0.37	1.25	16.38	8.14	7.62	66.61
	1	1.07	0.10	0.41	0.39	0.23	0.87	14.76	13.29	4.47	66.61
	3	1.07	0.11	0.39	0.40	0.25	1.01	13.58	13.83	5.26	66.32
	5	0.65	0.11	0.32	0.24	0.08	3.01	23.93	13.55	1.42	58.09
SES tercile	1	0.91	0.10	0.36	0.38	0.21	1.25	15.97	17.34	5.35	60.08
	2	0.82	0.08	0.33	0.30	0.14	0.92	16.67	13.50	2.90	66.01
	3	1.17	0.25	0.44	0.43	0.29	4.47	14.17	13.11	6.16	62.09
Mother primary	0	0.97	0.20	0.37	0.38	0.21	4.22	14.98	15.46	4.75	60.58
	1	0.85	0.18	0.33	0.33	0.16	4.74	15.15	14.90	3.62	61.60
All	.	1.19	0.25	0.44	0.42	0.31	4.54	13.61	12.39	6.54	62.92

Note: for each stratifying variable, indicated in the first column, sub-groups (second column) are mutually exclusive and span the entire dataset; female is a dummy variable (i.e., 0 = male, 1 = female); age and grade levels are grouped (e.g., age level 6 indicates children aged 6-8; age level 15 is children 15 and above; grade level 0 contains never enrolled children; grade level 5 is all those with highest grade 5 and above); for SES tercile, level 1 is the poorest group; mother edu. takes a value of 1 if the mother has attended primary school; variance components are as per equations (3b)-(3c) and the individual effects are aggregated for simplicity (denoted,  $\sigma_{a+}^2$ ); absolute and relative contributions are as per earlier tables, reported in standard deviation units and percentages respectively.

Source: own calculations.

Table B8: Variance contributions for Tanzania, by sub-group

Strata	Level	Absolute (s.d. units)					Relative (in %)				
		$\sigma_t^2$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$
Female	0	1.01	0.24	0.42	0.36	0.26	5.53	17.34	12.63	6.76	57.72
	1	0.98	0.22	0.42	0.36	0.26	5.17	17.99	13.27	6.82	56.76
Age	6	0.98	0.28	0.42	0.36	0.23	8.03	18.57	13.54	5.74	54.12
	9	1.00	0.20	0.42	0.37	0.27	4.06	17.72	13.93	7.32	56.97
	12	1.01	0.21	0.42	0.34	0.27	4.47	17.48	11.56	7.37	59.12
	15	1.02	0.25	0.42	0.32	0.27	5.95	16.76	10.04	7.15	60.10
Grade level	0	1.03	0.15	0.51	0.26	0.28	2.13	24.89	6.26	7.59	59.13
	1	1.02	0.15	0.43	0.37	0.23	2.13	17.68	13.51	5.19	61.50
	3	1.07	0.17	0.43	0.38	0.28	2.41	16.00	12.74	6.88	61.95
	5	0.77	0.17	0.37	0.29	0.18	4.81	22.93	13.78	5.49	52.98
SES tercile	1	0.95	0.17	0.41	0.37	0.25	3.04	18.16	15.04	6.77	56.98
	2	0.89	0.07	0.38	0.32	0.25	0.56	18.52	13.04	7.72	60.16
	3	1.03	0.25	0.44	0.36	0.25	6.07	18.34	12.46	5.68	57.44
Mother primary	0	0.99	0.24	0.42	0.36	0.24	5.80	17.91	13.09	5.99	57.22
	1	0.91	0.20	0.38	0.34	0.23	5.03	17.25	13.87	6.34	57.51
All	.	1.06	0.25	0.45	0.37	0.28	5.64	17.81	12.05	6.70	57.80

Note: for each stratifying variable, indicated in the first column, sub-groups (second column) are mutually exclusive and span the entire dataset; female is a dummy variable (i.e., 0 = male, 1 = female); age and grade levels are grouped (e.g., age level 6 indicates children aged 6-8; age level 15 is children 15 and above; grade level 0 contains never enrolled children; grade level 5 is all those with highest grade 5 and above); for SES tercile, level 1 is the poorest group; mother edu. takes a value of 1 if the mother has attended primary school; variance components are as per equations (3b)-(3c) and the individual effects are aggregated for simplicity (denoted,  $\sigma_{a+}^2$ ); absolute and relative contributions are as per earlier tables, reported in standard deviation units and percentages respectively.

Source: own calculations.

Table B9: Variance contributions for Uganda, by sub-group

Strata	Level	Absolute (s.d. units)					Relative (in %)				
		$\sigma_t^2$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_{a+}^2$	$\sigma_h^2$	$\sigma_s^2$	$2\Sigma_{hs}$	$\sigma_e^2$
Female	0	0.99	0.18	0.39	0.40	0.25	3.15	15.66	15.84	6.28	59.08
	1	1.00	0.18	0.40	0.40	0.26	3.07	15.50	15.73	6.47	59.23
Age	6	0.99	0.18	0.40	0.35	0.25	3.44	16.43	12.59	6.29	61.26
	9	1.00	0.14	0.39	0.39	0.29	2.06	15.33	15.40	8.63	58.58
	12	1.00	0.16	0.39	0.38	0.28	2.53	15.16	14.63	7.89	59.80
	15	1.01	0.18	0.39	0.35	0.23	3.37	14.91	11.99	5.26	64.47
Grade level	0	0.96	0.21	0.42	0.18	0.19	4.78	19.44	3.50	3.75	68.53
	1	0.93	0.13	0.40	0.33	0.21	1.97	17.98	12.53	4.82	62.70
	3	1.08	0.15	0.40	0.38	0.27	2.04	14.14	12.28	6.41	65.14
	5	0.82	0.16	0.37	0.29	0.14	3.80	19.92	12.71	2.82	60.75
SES tercile	1	0.97	0.13	0.38	0.41	0.25	1.75	15.87	18.01	6.50	57.87
	2	0.97	0.10	0.40	0.40	0.25	0.98	16.77	16.52	6.60	59.15
	3	0.97	0.18	0.39	0.39	0.22	3.35	15.79	15.95	4.99	59.91
Mother primary	0	0.96	0.17	0.39	0.39	0.23	3.17	16.22	16.73	5.76	58.12
	1	0.99	0.17	0.39	0.39	0.24	3.09	15.55	15.86	6.10	59.40
All	.	1.01	0.19	0.40	0.39	0.24	3.58	15.44	15.12	5.77	60.09

Note: for each stratifying variable, indicated in the first column, sub-groups (second column) are mutually exclusive and span the entire dataset; female is a dummy variable (i.e., 0 = male, 1 = female); age and grade levels are grouped (e.g., age level 6 indicates children aged 6-8; age level 15 is children 15 and above; grade level 0 contains never enrolled children; grade level 5 is all those with highest grade 5 and above); for SES tercile, level 1 is the poorest group; mother edu. takes a value of 1 if the mother has attended primary school; variance components are as per equations (3b)-(3c) and the individual effects are aggregated for simplicity (denoted,  $\sigma_{a+}^2$ ); absolute and relative contributions are as per earlier tables, reported in standard deviation units and percentages respectively.

Source: own calculations.

Table B10: Summary of absolute variance contributions, alternative choices of  $\pi$ , never enrolled children only

	$\pi$	Kenya		Tanzania		Uganda	
		$\sigma$	(s.e.)	$\sigma$	(s.e.)	$\sigma$	(s.e.)
Individual	0	0.10	(0.001)	0.06	(0.002)	0.17	(0.002)
	0.5	0.10	(0.001)	0.06	(0.002)	0.17	(0.002)
	1	0.10	(0.001)	0.05	(0.002)	0.17	(0.002)
Household	0	0.92	(0.012)	0.76	(0.012)	0.59	(0.011)
	0.5	0.64	(0.008)	0.57	(0.009)	0.47	(0.009)
	1	0.51	(0.007)	0.46	(0.007)	0.40	(0.008)
School	0	0.20	(0.003)	0.16	(0.003)	0.15	(0.003)
	0.5	0.44	(0.007)	0.27	(0.005)	0.20	(0.004)
	1	0.80	(0.012)	0.51	(0.009)	0.36	(0.008)
Sorting	0	0.11	(0.001)	-0.15	(0.003)	0.03	(0.001)
	0.5	0.49	(0.007)	0.35	(0.006)	0.25	(0.005)
	1	0.32	(0.004)	0.30	(0.005)	0.23	(0.005)
Residual	.	1.21	(0.017)	0.82	(0.014)	0.83	(0.017)
Total	.	1.55	(0.013)	1.11	(0.011)	1.03	(0.012)

Note: the table sets out the absolute variance contribution (in standard deviation units) attributable to each component of the test score, where ‘sorting’ is the contribution of the between-factor covariance; different initializations of the partitioned iterative algorithm are indicated by column  $\pi$ ; standard errors are reported in parentheses, calculated using the asymptotic approximation due to Ahn and Fessler (2003).

Source: own calculations.

Table B11: Summary of relative variance contributions, alternative choices of  $\pi$ , never enrolled children only

	$\pi$	Kenya		Tanzania		Uganda	
		$\sigma^2/\sigma_t^2$	(s.e.)	$\sigma^2/\sigma_t^2$	(s.e.)	$\sigma^2/\sigma_t^2$	(s.e.)
Individual	0	0.44	(0.10)	0.26	(0.22)	2.66	(0.31)
	0.5	0.43	(0.11)	0.30	(0.21)	2.75	(0.30)
	1	0.41	(0.11)	0.24	(0.22)	2.77	(0.30)
Household	0	35.23	(0.90)	46.59	(1.29)	32.30	(1.28)
	0.5	17.31	(0.63)	26.17	(0.97)	20.52	(1.02)
	1	10.80	(0.50)	17.18	(0.78)	14.72	(0.87)
School	0	1.68	(0.23)	2.19	(0.30)	2.06	(0.37)
	0.5	8.10	(0.50)	5.96	(0.49)	3.62	(0.49)
	1	26.80	(0.90)	21.22	(0.92)	12.01	(0.89)
Sorting	0	0.47	(0.11)	-1.95	(0.27)	0.10	(0.08)
	0.5	10.01	(0.51)	9.73	(0.61)	5.84	(0.58)
	1	4.18	(0.33)	7.33	(0.53)	4.83	(0.53)
Residual	.	61.38	(1.56)	54.93	(1.85)	65.28	(2.39)
Total	.	100.00	(1.15)	100.00	(1.44)	100.00	(1.71)

Note: the table sets out the relative variance contribution (in percent) attributable to each component of the test score, where ‘sorting’ is the contribution of the between-factor covariance; different initializations of the partitioned iterative algorithm are indicated by column  $\pi$ ; standard errors are reported in parentheses, calculated using the asymptotic approximation due to Ahn and Fessler (2003).

Source: own calculations.

Table B12: Summary of absolute variance contributions, alternative choices of  $\pi$ , children attending school only

	$\pi$	Kenya		Tanzania		Uganda	
		$\sigma$	(s.e.)	$\sigma$	(s.e.)	$\sigma$	(s.e.)
Individual	0	0.05	(0.000)	-0.04	(0.000)	0.10	(0.001)
	0.5	0.04	(0.000)	-0.05	(0.000)	0.09	(0.001)
	1	0.04	(0.000)	-0.05	(0.000)	0.09	(0.001)
Household	0	0.48	(0.002)	0.56	(0.002)	0.52	(0.002)
	0.5	0.36	(0.001)	0.39	(0.002)	0.39	(0.002)
	1	0.32	(0.001)	0.34	(0.001)	0.35	(0.002)
School	0	0.29	(0.002)	0.25	(0.002)	0.31	(0.002)
	0.5	0.37	(0.002)	0.36	(0.002)	0.40	(0.003)
	1	0.51	(0.003)	0.55	(0.004)	0.56	(0.004)
Sorting	0	-0.07	(0.000)	-0.11	(0.001)	0.10	(0.001)
	0.5	0.20	(0.001)	0.24	(0.001)	0.25	(0.001)
	1	-0.12	(0.001)	-0.16	(0.001)	-0.07	(0.000)
Residual	.	0.68	(0.002)	0.67	(0.003)	0.73	(0.003)
Total	.	0.88	(0.002)	0.90	(0.002)	0.97	(0.002)

Note: the table sets out the absolute variance contribution (in standard deviation units) attributable to each component of the test score, where ‘sorting’ is the contribution of the between-factor covariance; different initializations of the partitioned iterative algorithm are indicated by column  $\pi$ ; standard errors are reported in parentheses, calculated using the asymptotic approximation due to Ahn and Fessler (2003).

Source: own calculations.

Table B13: Summary of relative variance contributions, alternative choices of  $\pi$ , children attending school only

	$\pi$	Kenya		Tanzania		Uganda	
		$\sigma^2/\sigma_t^2$	(s.e.)	$\sigma^2/\sigma_t^2$	(s.e.)	$\sigma^2/\sigma_t^2$	(s.e.)
Individual	0	0.29	(0.06)	-0.24	(0.05)	1.01	(0.09)
	0.5	0.18	(0.06)	-0.28	(0.05)	0.84	(0.08)
	1	0.17	(0.06)	-0.28	(0.05)	0.91	(0.08)
Household	0	29.08	(0.21)	38.26	(0.31)	29.16	(0.27)
	0.5	16.21	(0.16)	19.14	(0.22)	16.04	(0.20)
	1	13.04	(0.14)	14.20	(0.19)	12.68	(0.18)
School	0	11.02	(0.21)	7.93	(0.20)	10.44	(0.27)
	0.5	17.22	(0.26)	16.22	(0.29)	17.30	(0.34)
	1	32.87	(0.36)	37.85	(0.44)	33.03	(0.47)
Sorting	0	-0.62	(0.04)	-1.47	(0.07)	0.96	(0.06)
	0.5	5.11	(0.11)	7.02	(0.16)	6.72	(0.16)
	1	-2.01	(0.07)	-2.99	(0.10)	-0.50	(0.05)
Residual	.	59.14	(0.38)	54.88	(0.47)	57.14	(0.46)
Total	.	100.00	(0.29)	100.00	(0.37)	100.00	(0.35)

Note: the table sets out the relative variance contribution (in percent) attributable to each component of the test score, where ‘sorting’ is the contribution of the between-factor covariance; different initializations of the partitioned iterative algorithm are indicated by column  $\pi$ ; standard errors are reported in parentheses, calculated using the asymptotic approximation due to Ahn and Fessler (2003).

Source: own calculations.

## C List of aggregated regions

Country	Aggregated region	Actual region	Obs.
KE	Central	Central	33,306
KE	Central	Nairobi	4,564
KE	Coast	Coast	49,921
KE	Eastern	Eastern	84,133
KE	North Eastern	North Eastern	54,132
KE	Nyanza	Nyanza	76,471
KE	Rift Valley	Rift Valley	169,672
KE	Western	Western	83,720
TZ	Arusha	Arusha	18,609
TZ	Arusha	Kilimanjaro	16,102
TZ	Arusha	Mara	18,172
TZ	Dar Es Salaam	Dar Es Salaam	5,889
TZ	Dar Es Salaam	Pwani	17,136
TZ	Iringa	Dodoma	16,991
TZ	Iringa	Iringa	15,594
TZ	Iringa	Morogoro	13,500
TZ	Iringa	Njombe	2,532
TZ	Kagera	Geita	4,600
TZ	Kagera	Kagera	20,595
TZ	Kagera	Mwanza	25,281
TZ	Kigoma	Katavi	1,994
TZ	Kigoma	Kigoma	14,228
TZ	Kigoma	Rukwa	16,240
TZ	Ruvuma	Lindi	10,611
TZ	Ruvuma	Mtwara	6,594
TZ	Ruvuma	Ruvuma	12,662
TZ	Singida	Mbeya	18,232
TZ	Singida	Singida	14,037
TZ	Tabora	Shinyanga	25,162
TZ	Tabora	Simiyu	4,576
TZ	Tabora	Tabora	20,198
TZ	Tanga	Manyara	16,836
TZ	Tanga	Tanga	23,092
UG	Central	Central	64,077
UG	Eastern	Eastern	120,142
UG	Northern	Northern	102,723
UG	Western	Western	75,186