UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2021/101

How far does the apple really fall from the tree?

Practical guidance on measuring intergenerational mobility from a simulation framework

Marcelo Cardona and Sam Jones*

June 2021

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: Despite the importance attributed to intergenerational educational mobility in the process of development, there remains little consensus on how mobility should be measured. We present analytical and empirical evidence regarding the sensitivity of alternative estimators to different forms of measurement error and data transformations. To do so, we develop a flexible simulation model, allowing us to quantify the bias associated with different empirical choices. Our evidence suggests that use of an upwards mobility estimator, complemented by an out-performance metric, based on a reference distribution transformation is comparatively most robust. Applying these recommendations to the case of Mozambique, using complete census records from 1997, 2007, and 2017, we find a high degree of provincial heterogeneity, with a marked north-south gradient in mobility. Moreover, we identify a clear slowdown in educational mobility between 2007 and 2017, especially in the northern region, suggesting the education system is not delivering consistent gains to children across the country.

Key words: mobility, measurement errors, data truncation, sample truncation, co-residency, simulations

JEL classification: C15, I24, I25, O15

* UNU-WIDER; corresponding author: jones@wider.unu.edu

This study has been prepared within the UNU-WIDER project Inclusive growth in Mozambique – scaling-up research and capacity implemented in collaboration between UNU-WIDER, University of Copenhagen, University Eduardo Mondlane, and the Mozambican Ministry of Economics and Finance. The project is financed through specific programme contributions by the governments of Denmark, Finland, and Norway.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-041-2

https://doi.org/10.35188/UNU-WIDER/2021/041-2

Typescript prepared by Mary Lukkonen.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Intergenerational educational mobility (IGM) refers to the relationship between child and parental educational outcomes. In large part, the concern for mobility reflects the critical role accorded to education as a means to enhance social opportunities and advance society more generally. As Horace Mann put it long ago:

> Education, then, beyond all other devices of human origin, is the great equalizer of the conditions of men – the balance-wheel of the social machinery. ... It does better than to disarm the poor of their hostility towards the rich: it prevents being poor (Mann 1868: 669).

Evidence suggests that countries with higher IGM are usually countries with lower levels of income inequality (Andrews and Leigh 2009; Glewwe 2012). Because education also is an important determinant of health and wealth, among other things, greater IGM can drive improvements in overall living standards through time, while lower mobility can be understood as a deprivation of opportunities for those with less favourable family backgrounds.

The first generation of empirical research on IGM focused primarily on the slope of the linear relationship between child and parental outcomes, capturing the extent to which educational outcomes are 'inherited' (Becker and Tomes 1979, 1986; Solon 1992; Hertz et al. 2007; S. E. Black and Devereux 2011). While this linear relation is important, upward shifts in average levels of education (for all children), *without* changes in heritability, may also be of interest. This may be especially pertinent in low-income contexts that have embarked on large-scale reforms to enhance access to schooling (e.g., via abolition of user fees; see Fox et al. 2012; Glewwe and Muralidharan 2016). Such gains in (absolute) attainment may be missed where mobility is defined exclusively in terms of heritability. Reflecting on this insight, various recent contributions widen the scope of analysis, using alternative metrics of mobility, including the degree of (absolute or relative) upward mobility (e.g., Chetty et al. 2014; Asher et al. 2020).

In addition to debates regarding the relevant dimensions of intergenerational educational mobility and its corresponding metrics, there are long-standing concerns that such estimators may be badly biased (Emran et al. 2018; Emran and Shilpi 2019). Parent and child educational outcomes are typically only observed in co-resident samples, and various forms of measurement error may be common when administrative data are not used—e.g., individuals may exaggerate their attainment or enumerators can miscode educational categories (Bertrand and Mullainathan 2001; D. Black et al. 2003). Data transformations, such as use of rank percentiles, constitute one means to address measurement errors and facilitate comparison of metrics over space and time. However, the sensitivity of different estimators both to alternative data transformations and different sources of measurement error is not well understood, meaning empirical researchers face what may seem to be a bewildering range of choices (e.g., Emran and Shilpi 2019).

This bewilderment represents our point of departure. Our primary contribution is to develop a highly flexible simulation framework to investigate the magnitude of bias associated with different measurement choices. Specifically, we explore the combination of three mobility estimators (the heritability coefficient, out-performance, and upwards mobility) and three data transformations (raw data in levels, a reference distribution transformation, and rank-rank percentiles). Allowing for classical and multiplicative measurement errors, as well as bias from missing observations and incomplete schooling, we run simulations comparing 'true' and 'observed' mobility estimates. Based on this evidence we draw a set of five practical implications, recommending use of the upwards mobility estimator with a reference distribution transformation as a reliable basis for analysis of IGM. Finally, we apply these lessons to three waves of census data from Mozambique, revealing low levels of mobility on average as well as a very strong (and worsening) north-south gradient to educational progress.

2 Measures of mobility

The canonical approach to analysing intergenerational mobility, whether defined in terms of income or education, has focused on the slope of the relationship between parent and child outcomes. For example, consider the linear conditional expectation:

$$\mathrm{E}(c_i^* \mid p_i^*) = \alpha + \beta p_i^* \tag{1}$$

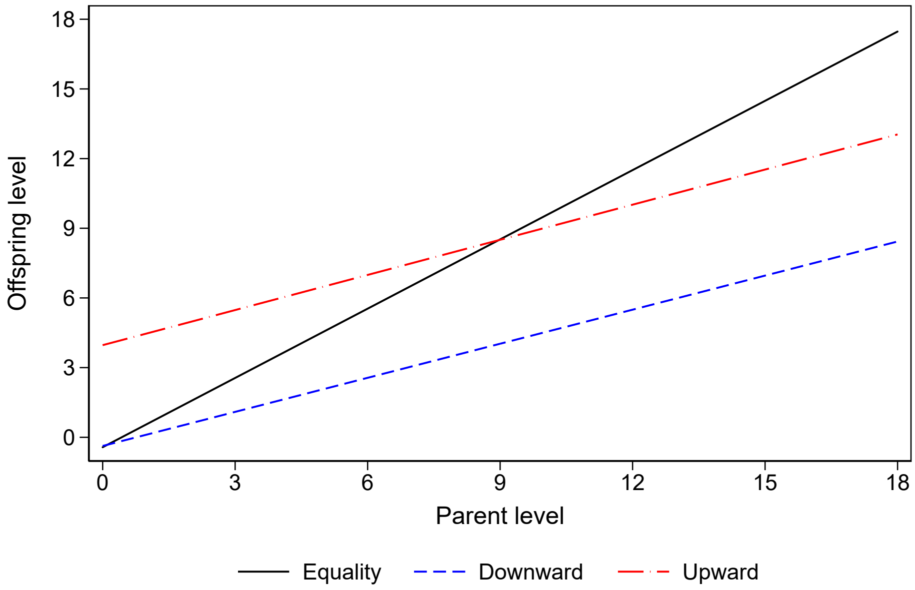
where c^* is the final educational achievement of child i, and p^* is that of their parent.[1] Thus, parameter β—or some standardized version thereof, such as the pairwise correlation coefficient—indicates the degree to which education is inherited across generations. Lower values of β imply less heritability and, thus, greater scope for intergenerational mobility on average (Solon 1992; Hertz et al. 2007; S. E. Black and Devereux 2011; Emran et al. 2018).

Understanding mobility in terms of the degree of heritability in outcomes is intuitive and simple. Nonetheless, as noted by Asher et al. (2020) among others, a unique focus on the slope of the relationship between parents and their offspring is likely to provide at best an insufficient and at worst a misleading characterization of the extent to which the attainment of children (of low-outcome parents) exceeds that of their parents. For instance, if lower heritability of outcomes is achieved at the cost of poorer overall (absolute) outcomes, either on average or for some previously disadvantaged group, it is not self-evident this would represent a genuine *improvement* in intergenerational mobility.

To see this, consider the stylized sets of relations depicted in Figure 1, where the solid line indicates the benchmark case of zero mobility (perfect heritability), given in equation (1) by $\alpha = 0, \beta = 1$. The other two cases both halve the degree of heritability ($\beta = 0.5$). However, in the 'downward' case there is no shift in the intercept, meaning offspring outcomes are never better than their parent's. In contrast, under the 'upward' case, the intercept shifts upward ($\alpha = 4.5$), meaning children born to parents with less than around nine years of education—i.e. observations to the left of the intersection with the line of equality—are expected to outperform their parents, which is plausibly consistent with substantial upward mobility, defined as the difference in outcomes between offspring and parents. However, because of upper limits on the number of years of education that can be attained (here 18), it must be the case that for most values of α, some degree of downward mobility must be observed among the offspring of the highest-achieving parents (for $\beta \geq 1$).

[1] For simplicity, we abstract here from possible differences in education between mothers and fathers.

Figure 1: Stylized expectation functions of offspring education conditional on parental education



Source: authors' elaboration.

In this light, a unique focus on the slope parameter would seem to provide inadequate descriptive insights into the relation between offspring and parental outcomes. The immediate alternative is to define mobility with respect to the full linear conditional expectation function. For instance, from equation (1), we see that for any fixed level of parental education p^*, the expression: $\Delta_{p^*} = \alpha + (\beta - 1)p^*$ gives the expected (absolute) difference between offspring and parental outcomes at that point in the parental distribution.[2] Or, $\alpha/(1-\beta)$ indicates the point of intersection with the line of equality and, thus, the highest level of parental education below which offspring outcomes are expected to exceed those of their parents, on average. So, consideration of both β and α (as in Δ_{p^*}) may provide a more complete picture of overall mobility—capturing the degree of both heritability and out-performance, which refer to *relative* and *absolute* aspects of intergenerational mobility, respectively.

There are two main disadvantages of the above measures. First, they impose a linear relationship that may hide important heterogeneity at the bottom (or top) of the distribution of parental education. Indeed, conventional ordinary least squares (OLS) estimates of equation (1) generally do not correspond to any meaningful weighted average of what may be heterogeneous subgroup parameters (Gibbons et al. 2019). To avoid this, scholars have also used non-parametric estimators, such as the share of children who achieve higher education than their parents, which is just the empirical probability of absolute upward mobility (e.g., Card et al. 2018).

The second disadvantage is that mobility estimates calculated from raw outcome levels will be sensitive to monotonic transformations, such as level shifts in either parental or child education or (order-preserving) changes in the shape of their distributions (e.g., more dispersion), in turn making them unsuited for making consistent subgroup comparisons. Thus, rank-transformed metrics of mobility are often preferred (e.g., Landersø and Heckman 2016; Eriksen and Munk 2020; Karlson and Landersø 2021)—i.e. we re-express equation (1) in rank-rank terms:

$$\mathrm{E}(c_i^{*r} \mid p_i^{*r}) = \alpha^r + \beta^r p_i^{*r} \quad (2)$$

[2] In keeping with Chetty et al. (2014), setting p^* at a specified percentile, such as the median or below, constitutes a natural reference point. See further below.

where r superscripts refer to rank-transformed variables or the corresponding parameters of the conditional linear expectation function. So, β^r gives the rank correlation, capturing relative rank mobility; rank out-performance is given by the expected percentile rank of the child (in her distribution of outcomes) calculated at the 25th percentile of the parent outcome distribution, $\Delta^r_{p25} = \alpha^r + .25(\beta^r - 1)$; and a non-linear alternative is the share of children with a higher percentile rank than their parents, conditional on their parents being located in the first two quartiles of the parental distribution.

Despite these attractive properties, it is worth noting that the properties of a rank-rank specification mean that *any* difference in ranks between offspring and parents will be reflected in a heritability coefficient below unity ($\beta^r < 1$), regardless of differences in absolute outcomes. This has substantive implications. Consider the following thought experiment: if the offspring of all parents holding less than a full primary school education accumulate no schooling (e.g., because of some exogenous shock), then the rank position of these individuals will tend to be higher than that of their parents simply because of the concentration of mass at zero years of schooling; even if all other offspring accumulate the same education as their parents, their rank positions will be mechanically lower, implying the overall rank correlation is below unity.[3] That is, there is nothing to prohibit a child from attaining a higher (lower) rank than their parent but a lower (higher) outcome level. Indeed, the insensitivity of the rank transformation to differences in (average) levels is reinforced from the OLS definition: $\alpha^r = 0.5(1-\beta^r) \implies \Delta^r_{p25} = 0.25(1-\beta^r) \geq 0$.[4] So, counterintuitively, *any* deviation from perfect heritability guarantees the rank out-performance will be in the positive domain.

The above demonstrates that rank-rank type mobility measures, as recommended by Chetty et al. (2014), depend exclusively on the rank-rank correlation coefficient (i.e. only β^r matters). To avoid this, a potential middle ground is to rank all outcomes according to some fixed (meaningful) reference distribution, such as that of the older generation.[5] This represents a form of reference distribution normalization and has the advantage of straightforward interpretation and easy use, especially because the moments of the parental ranks—when used as the reference—will follow those of a uniform distribution. For instance, under this latter transformation, denoted with d superscripts, we have: $\alpha^d = \mathrm{E}(c^{*d}) - 0.5\beta^d \implies \Delta^d_{p25} = \mathrm{E}(c^{*d}) - 0.25(1+\beta^d)$, which shows the average percentile rank of children within the distribution of parental outcomes contributes to the out-performance metric. And, more generally, the empirical distribution of the offspring relative ranks will clearly indicate the extent and location of deviations from the reference ranking (we return to this below).

3 Bias characterization

Thus far, we have established that different empirical measures employed in the literature capture distinct attributes of the relationship between offspring and parental outcome distributions and, in so doing, effectively respond to quite different dimensions of mobility (see also Emran et al. 2018; Emran and Shilpi 2019). Table B1 summarizes the various metrics of IGM, distinguishing between those based on different transformations and those that concern relative versus absolute dimensions. Aside from the theoretical properties of these different measures, however, a related concern refers to the extent to which these may be biased in typical empirical applications. With the exception of a few high-income countries (Björklund, Lindahl, and Plug 2006; Björklund and Jäntti 2009; Landersø and Heckman 2016; Karlson

[3] The same result will hold for any arbitrary years of schooling below primary that all such children accumulate.

[4] This comes from the fact that, under the rank transformation, the expected value of both child and parental outcomes will be 0.5.

[5] For related applications, see the method of 'relative distribution' analysis (Morris, Bernhardt, and Handcock 1994; Handcock and Morris 1998).

and Landersø 2021), detailed records linking the outcomes of adult children to their parents are scant, meaning co-resident samples—such as population and housing census data—are most often used for the analysis of intergenerational educational mobility.

Within the context of this kind of co-resident data, at least three generic challenges arise. First, children living with their parents may not yet have completed all their education, in which case their current reported level of completed schooling will be biased downward.[6] Second, linked parent-child outcomes are only measured for children residing with their parents, implying a missing data problem, and there is no *a priori* reason to assume that the decision to leave the household is random. Third, both parent and child outcomes may be measured with error, particularly where interval-type scales are used to record completed education (e.g., primary, secondary).

To capture the implications of these challenges, an analytical characterization of the expected bias in our mobility measures is helpful. To do so, we extend the generalized errors-in-variables approach suggested by Nybom and Stuhler (2017), which applies to a wide range of data transformations, taking account the additional complexities associated with education data. Concretely, consider the following setup:

$$c^*_{ij} = \alpha_j + \beta_j p^*_i + \varepsilon^*_{ij} \tag{3a}$$

$$\beta = \sum_j \pi_j \beta_j \tag{3b}$$

$$c_{ij} = c_0 + (1+\lambda_c)c^*_{ij} - \delta m^*_i + \nu_i \tag{3c}$$

$$p_i = p_0 + (1+\lambda_p)p^*_i + \mu_i \tag{3d}$$

where $j \in \{0,1\}$ indexes co-residency status (or other relevant subgroups), such that the aggregate slope coefficient (β) in equation (3b) is a weighted average of the sub-group-specific slopes, but where the weights (π) do not necessarily coincide with their respective sample proportions.[7] Equations (3c) and (3d) represent linear projections of the observed outcomes on the unobserved (true) variables, which we assume are only available for the co-resident subgroup. Here, $m^* \geq 0$ represents the magnitude of 'missing' years of education where the child has not yet completed her schooling (recall, c^* is the final level of education accumulated), which is plausibly a function of the child's current age and other factors that influence the probability of enrolment—hereafter, we refer to this as the age effect. In turn, ν_i, μ_i are mean-zero white noise terms, $\mathrm{E}(\nu_i) = 0 = \mathrm{E}(\nu_i)$. And while, by definition, they are uncorrelated with all true variables, we permit their mutual correlation to be non-zero: $\mathrm{E}(\nu_i \mu_i) \neq 0$.[8] Finally, c_0, p_0 are intercept terms, which generally are expected to be non-zero under any of the rank transformations, and ε^*_{ij} is pure white noise, reflecting the notion that equation (3a) is a linear projection.

The above structure is shown for data expressed in levels. It also applies naturally to the two-rank transforms; however, when moving from one transform to another (with the same data), the error parameters will not be identical. For example, under the level's specification we expect $\delta \equiv 1$, while for the rank transform, $\delta < 1$ captures the mapping of the missing years onto the ranks of the observed outcome. Similarly, the properties of the rank-rank transform imply we expect λ_p always to be less than or equal to zero, regardless of the corresponding parameter magnitude under a level's specification (see Nybom and Stuhler 2017), but this does not automatically hold for λ_c, unless $\delta = 0$.

---

[6] This problem is similar to the life cycle difficulties facing measures of intergenerational income mobility.

[7] In general, these weights will reflect both sample frequencies *and* the relative magnitude of the variance of the explanatory variable in the subsample to that in the full sample. So, here, subgroups with relatively greater variance in parental education will attract a higher weight relative to their sample frequency.

[8] One rationale for allowing some mutual correlation is that, in many surveys, a single person reports on behalf of the entire household. Thus, if the individual misunderstands or has a specific propensity to bias, then this may reveal itself as correlated additive measurement error.

Using the above system, and assuming $j \in \{0,1\}$, note that the overall bias of the slope coefficient will be as follows:

$$\hat{\beta}_1 - \beta = (1-\pi_1)(\hat{\beta}_1 - \beta_0) + \pi_1(\hat{\beta}_1 - \beta_1) \quad (4)$$

which is a weighted combination of bias driven by missing data (the first term on the RHS) and the bias of the estimator for the observed subgroup ($j = 1$, in the second term). Thus, the larger the effective weight associated with the observed subgroup, the smaller the overall contribution of bias from the missing observations (sample truncation). Further elaboration indicates the technical sources of bias from least squares estimates of the observed linear relationship between child and parent outcomes. To see this, insert the true relationship from equation (3a) into equation (3c), then substitute for the true parental outcome using equation (3d), yielding:

$$c_i = c_0 + (1+\lambda_c)\left(\alpha + \beta\left(\frac{p_i - \mu_i - p_0}{1+\lambda_p}\right) + \varepsilon_i^*\right) - \delta m_i^* + \upsilon_i \quad (5)$$

Using this, we obtain an expression for the least squares estimate of the relationship of interest for the observed (co-resident) subgroup:

$$\hat{\beta}_1 = \beta_1 \frac{(1+\lambda_c)}{(1+\lambda_p)}\left(1 - \frac{\text{var}(\mu_i)}{\text{var}(p_i)}\right) - \delta\frac{\text{cov}(m_i^*, p_i)}{\text{var}(p_i)} + \frac{\text{cov}(\nu_i, p_i)}{\text{var}(p_i)} \quad (6)$$

Applying earlier assumptions, this can be further clarified as:

$$\hat{\beta}_1 = \beta_1 \frac{(1+\lambda_c)}{(1+\lambda_p)}\left(1 - \frac{\text{var}(\mu_i)}{\text{var}(p_i)}\right) - \delta(1+\lambda_p)\frac{\text{cov}(m_i^*, p_i^*)}{\text{var}(p_i)} + \frac{\text{cov}(\nu_i, \mu_i)}{\text{var}(p_i)} \quad (7)$$

Here, the first term on the right-hand side (RHS) combines the contributions of the multiplicative measurement errors (λ_c, λ_p) and attenuation bias associated with the additive error in the parental outcome (μ_i). The second term captures the contribution of the age effect, driven exclusively by the covariance between the incomplete years of education and the true level of parental education and which will be exacerbated by any multiplicative error in the latter. The third term captures the covariance between the additive error terms.

Analytically, equation (7) indicates the likely direction of the bias of the estimated slope coefficient ($\hat{\beta}_1 - \beta_1$) for the general structure proposed in equations (3a)-(3d). Importantly, this not only shows there are multiple different potential sources of bias, but also these can run in different directions, so the overall direction of bias is unlikely to be clear without imposing very strong assumptions *ex ante*. For instance, while classical measurement is expected to lead to downward bias in the estimated coefficient, this could be offset by other forms of measurement error, but it also may be exacerbated, such as from a large positive covariance between the missing years of the child's education and the parent's education or where the multiplicative error in parental education is positive. Derivations in Appendix 6 confirm the ambiguous net direction of bias holds for both the reference distribution and rank-rank estimators of the slope coefficient. And, in both these cases, even though the multiplicative error associated with the parental outcome will be fixed below unity [e.g., $(1+\lambda_p^r) < 1$], this provides no guarantee of fully offsetting either attenuation bias or other sources of error, not least because of the dependence on the ratio $(1+\lambda_c^r)/(1+\lambda_p^r)$. Moreover, any bias in the estimated slope coefficient will also directly affect the estimate for the intercept.

The theoretically indeterminate bias of the heritability coefficient provides a primary justification for moving to the empirical realm, using simulation evidence to explore the properties of the different estimators under specified conditions. Simulations are also particularly well-suited to allow for truncation bias from non-co-resident children, as well to inspect the bias of $\Delta_{p_{25}}$ and upward mobility estimators. As indicated in Appendix 6, for both the levels and reference distribution estimators, these generally show a more complex structure because of the inclusion of additional terms. However, in all cases, we note the contribution of any bias associated with the slope coefficient enters with a negative sign in the $\Delta_{p_{25}}$ estimators, being directly proportional in the case of the rank-rank transformation.

4 Simulation approach

The previous section indicated that analytical investigation of bias in estimates of IGM can provide concrete insights but only under very specific and limited assumptions regarding the data-generating process. Moreover, such exercises merely indicate the expected direction of bias, and this quickly becomes ambiguous once multiple sources of error are permitted, which is almost unavoidable when rank transformations are applied. Consequently, we contend that a calibrated simulation-based approach can provide additional guidance regarding how the different estimators and transformations will perform in real-world settings.

4.1 Setup

Our simulation framework places additional structure on the general errors-in-variables system presented earlier. For simplicity, and consistent with empirical practice, we first estimate the system in levels and subsequently apply alternative transformations to the simulated variables of interest (p_i, p_i^*, c_i, c_i^*), from which the true and observed mobility measures are derived. The full system, in levels, is as follows:

$$p_i^* \sim 18 \cdot \mathscr{B}(a,b) \tag{8a}$$
$$\alpha_0 = \alpha_1 + \Delta_\alpha; \beta_0 = \beta_1 + \Delta_\beta \tag{8b}$$
$$c_{ij}^* = \alpha_j + \beta_j p_{ij}^* + \varepsilon_{1ij} \tag{8c}$$
$$c_{ij} = c_0 + (1+\lambda_c)c_{ij}^* - \delta m_i^* + \nu_i,\, \delta \in [0,1] \tag{8d}$$
$$m_i^* = \max\left[0, c_i^* - .9(y_i^* - 6)\right] \tag{8e}$$
$$y_i = 12 + (24-12)\mathscr{U}^{1.1} \tag{8f}$$
$$p_i = p_0 + (1+\lambda_p)p_i^* + \mu_i \tag{8g}$$
$$\nu_i = \varepsilon_{2i},\ \sigma_\nu^2 = \theta_1 \sigma_{p^*}^2 \tag{8h}$$
$$\mu_i = \gamma \nu_i + \varepsilon_{3i},\ \sigma_\mu^2 = \theta_2 \sigma_{p^*}^2 \tag{8i}$$
$$\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}, \sim \mathscr{N}(0, \sqrt{.5} \cdot \sigma_{p^*}) \tag{8j}$$

Each realization (iteration) of the simulation represents a specific combination of: (a) choices for the so-called calibrated parameters: $a, b, f_1, \alpha_0 \geq 0, 0 < \beta_0 < 1$, where $f_1 = N_1/N$ represents the share of the sample in the observed co-resident group; (b) assumptions on the measurement error structure: $\Delta_\alpha, \Delta_\beta, \theta_1, \theta_2, p_0, c_0, \lambda_c, \lambda_p, \gamma, \delta$, where $\delta = 1$ fully switches on the age effect term; and (c) corresponding random draws for $p_i^*, a_i, \varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}$. Throughout, we fix $N = 5{,}000$, and all age (denoted y_i) and education variables are rounded to the nearest integer, consistent with the data typically obtained from surveys.[9]

To evaluate the performance of the different estimators, we follow Harwell (2019) and report two standard metrics: the mean bias (MB) and the mean fractional bias (MFB), where the latter is particularly valuable to compare across alternative transformations with different measurement scales. Taking the example of

[9] The system assumes children start school at age six years, thus the arithmetic difference $y_i - 6$ indicates the child's (maximum) current years of accumulated schooling. Also, consistent with empirical practice (Emran et al. 2018; Card et al. 2018; Alesina, Hohmann, Michalopoulos, and Papaioannou 2019), we only consider offspring aged 12–24 in the analysis.

the slope estimator estimated in levels, they are calculated as :

$$\text{MB}(\beta) = \frac{1}{N}\sum_{n=1}^{N}(\hat{\beta}_n - \beta_n) \tag{9a}$$

$$\text{MFB}(\beta) = \frac{1}{N}\sum_{n=1}^{N}\frac{2(\hat{\beta}_n - \beta_n)}{(\mid \hat{\beta}_n \mid + \mid \beta_n \mid)} \tag{9b}$$

where $n = \{1, \ldots, N\}$ indexes simulation iterations (for a given scenario).[10] In addition, to further simplify comparisons, we report the mean rank of the absolute fractional bias (RFB). Specifically, under each iteration, for *either* a given estimator (heritability, out-performance, or upwards mobility) *or* for a given transformation (levels, reference distribution, or rank-rank distribution), the lowest absolute fractional bias is ranked first.

4.2 Stylized scenarios

To begin the exercise, we consider a set of simple stylized scenarios, focusing on the implications of specific sources of measurement error. Concretely, and as summarized in Table 1, we run eight separate scenarios, which we initially calibrate to a typical low-income country with low average levels of parental education (see Appendix Table B2). For each chosen scenario we run 50 iterations of the simulation model and, for each iteration, five steps are followed. First, we simulate the true distribution of parental education, drawing from a *Beta* distribution, and then use equation (3a) to derive the 'true' final value of the child's education. Second, taking the measurement error assumptions for a given simulation, we derive the 'observed' child and parental outcome values. Third, we apply the rank transformations to both the true and observed series, and fourth, for each transformation, we calculate the associated set of 'true' and 'observed' mobility estimates, making sure to only include co-resident children for the observed estimates. Fifth, the magnitude of bias is estimated for the full set of nine estimates (3 estimators × 3 transformations). Finally, after running each set of iterations, we derive the mean and standard error of the bias estimates, which we report henceforth.

Table 1: Summary of stylized scenarios

Scenario		Assumptions
1	Classical measurement error	$\theta_1 > 0, \theta_2 > 0$
2	Correlated additive measurement error	$\theta_1 > 0, \theta_2 > 0, \gamma > 0$
3	Multiplicative measurement error, parent	$\lambda_p > 0$
4	Multiplicative measurement error, child	$\lambda_c > 0$
5	Multiplicative measurement error, both	$\lambda_p > 0, \lambda_c > 0$
6	Incomplete education	$\delta = 1$
7	Truncation bias	$\Delta_\alpha < 0, \Delta_\beta > 0$
8	Combined	As per rows 2, 5, 6, & 7

Note: in each scenario, given by a row of the table, all measurement error parameters not explicitly declared are set to zero.
Source: authors' elaboration.

Tables 2 and 3 present results for the mean bias and mean fractional bias associated with each estimator and transformation. To further assist interpretation, Tables 4 and 5 report the rank absolute fractional bias, comparing within estimators (across transforms) and within transforms (across estimators), respectively.

[10] The estimated parameter is included in the denominator of the MFB to address scaling issues if the true parameter is close to zero. In practise, this measure is effectively equivalent to the relative bias.

Four main insights stand out. First, consistent with theoretical expectations, different bias components—as captured by the individual scenarios (1–7)—work in different directions. As expected, classical measurement error attenuates the slope coefficient but generates upward bias in the out-performance metric (in which the slope enters with a negative sign). However, the presence of positive multiplicative error in the observed child's education tends to generate upwards bias in the slope coefficient. Nonetheless, bias in the upwards estimator tends to be negative on average across most scenarios and transforms.

Table 2: Mean bias of alternative IGM estimators under different data transforms and simulated error structures, low-income example

		Heritability ($\hat{\beta}$)			Out-performance ($\hat{\Delta}$)			Upwards mobility		
Scenario	|| Transform →	Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
1	Classical ME	-5.2	-6.9	-8.5	1.0	1.4	2.2	1.2	-1.3	0.5
		(0.2)	(0.2)	(0.2)	(0.0)	(0.1)	(0.1)	(0.2)	(0.3)	(0.5)
2	Correlated additive MEs	-2.6	-3.6	-4.9	0.6	0.5	1.2	0.0	-2.0	0.5
		(0.2)	(0.2)	(0.2)	(0.0)	(0.1)	(0.1)	(0.2)	(0.3)	(0.5)
3	Multiplicative ME, parent	-4.5	0.5	0.2	-0.3	-3.8	-0.0	-0.1	-0.2	-0.2
		(0.1)	(0.1)	(0.2)	(0.0)	(0.1)	(0.1)	(0.2)	(0.2)	(0.1)
4	Multiplicative ME, child	6.2	3.8	0.3	1.5	1.7	-0.1	-0.1	-0.2	-0.1
		(0.2)	(0.2)	(0.2)	(0.0)	(0.1)	(0.1)	(0.2)	(0.2)	(0.2)
5	Multiplicative ME, both	0.5	-0.0	-0.0	1.3	-0.0	-0.0	-0.2	-0.2	-0.2
		(0.1)	(0.2)	(0.2)	(0.0)	(0.1)	(0.1)	(0.1)	(0.2)	(0.2)
6	Incomplete education	-4.8	-2.5	-0.9	-0.2	-0.3	0.2	-0.1	-0.2	-0.2
		(0.1)	(0.2)	(0.2)	(0.0)	(0.1)	(0.1)	(0.2)	(0.2)	(0.1)
7	Truncation bias	-3.2	-1.8	-2.0	0.1	0.0	0.1	-0.1	-0.1	-0.1
		(0.1)	(0.2)	(0.2)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
8	Combined	-9.9	-7.7	-8.2	1.9	0.5	2.2	1.5	-1.2	1.8
		(0.2)	(0.2)	(0.2)	(0.0)	(0.1)	(0.1)	(0.3)	(0.2)	(0.4)

Note: cells show the mean bias of each estimator, calculated as per equation (9a) for different estimators and transformations (as indicated in the columns); each row indicates a specific stylized set of assumptions (see Appendix B2), for which 50 separate simulations are run and the mean bias and its standard error (in parentheses) are reported.

Source: authors' estimates.

Table 3: Mean fractional bias of alternative IGM estimators under different data transforms and simulated error structures, low-income example

		Heritability ($\hat{\beta}$)			Out-performance ($\hat{\Delta}$)			Upwards mobility		
Scenario	|| Transform →	Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
1	Classical ME	-13.2	-16.4	-18.5	15.8	11.0	16.3	2.1	-2.3	0.7
		(0.5)	(0.6)	(0.6)	(0.5)	(0.5)	(0.5)	(0.3)	(0.5)	(0.7)
2	Correlated additive MEs	-6.3	-8.1	-10.1	9.3	3.9	9.6	0.0	-3.5	0.7
		(0.4)	(0.4)	(0.4)	(0.6)	(0.6)	(0.5)	(0.4)	(0.6)	(0.7)
3	Multiplicative ME, parent	-11.3	1.1	0.3	-4.6	-38.9	-0.4	-0.2	-0.3	-0.3
		(0.3)	(0.3)	(0.3)	(0.6)	(0.7)	(0.5)	(0.3)	(0.3)	(0.2)
4	Multiplicative ME, child	13.4	7.9	0.6	22.3	13.0	-1.0	-0.2	-0.4	-0.2
		(0.3)	(0.3)	(0.3)	(0.5)	(0.5)	(0.5)	(0.3)	(0.3)	(0.2)
5	Multiplicative ME, both	1.2	-0.1	-0.1	19.0	-0.4	-0.4	-0.3	-0.3	-0.3
		(0.3)	(0.4)	(0.3)	(0.5)	(0.6)	(0.5)	(0.2)	(0.3)	(0.2)
6	Incomplete education	-12.1	-5.7	-1.9	-2.9	-2.2	1.5	-0.3	-0.4	-0.4
		(0.4)	(0.4)	(0.4)	(0.5)	(0.6)	(0.6)	(0.3)	(0.3)	(0.2)
7	Truncation bias	-7.1	-3.9	-4.0	1.7	0.3	0.4	-0.3	-0.3	-0.2
		(0.3)	(0.4)	(0.3)	(0.6)	(0.6)	(0.6)	(0.2)	(0.2)	(0.2)
8	Combined	-24.1	-17.6	-17.1	28.5	4.0	16.7	2.6	-2.0	2.6
		(0.4)	(0.4)	(0.4)	(0.5)	(0.5)	(0.4)	(0.5)	(0.3)	(0.6)

Note: cells show the mean fractional bias of each estimator, calculated as per equation (9b) for different estimators and transformations (as indicated in the columns); each row indicates a specific stylized set of assumptions (see Appendix B2), for which 50 separate simulations are run and the mean fractional bias and its standard error (in parentheses) are reported.

Source: author's estimates.

Table 4: Mean within-estimator rank absolute fractional bias of alternative IGM estimators under different data transforms and simulated error structures, low-income example

Scenario ‖ Transform →		Heritability ($\hat{\beta}$)			Out-performance ($\hat{\Delta}$)			Upwards mobility		
		Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
1	Classical ME	1.0	2.0	3.0	2.5	1.0	2.5	1.7	2.1	2.2
		(0.0)	(0.0)	(0.0)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)
2	Correlated additive MEs	1.1	1.9	3.0	2.4	1.0	2.5	1.5	2.3	2.1
		(0.0)	(0.0)	(0.0)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)
3	Multiplicative ME, parent	3.0	1.5	1.5	1.8	3.0	1.2	2.0	2.2	1.7
		(0.0)	(0.1)	(0.1)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)
4	Multiplicative ME, child	3.0	2.0	1.0	3.0	2.0	1.0	1.9	2.2	1.7
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
5	Multiplicative ME, both	2.2	2.1	1.8	3.0	1.6	1.4	1.9	2.2	1.8
		(0.1)	(0.1)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
6	Incomplete education	3.0	2.0	1.0	2.3	2.0	1.7	2.1	2.2	1.6
		(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
7	Truncation bias	3.0	1.5	1.5	2.3	1.8	1.9	2.0	2.1	1.8
		(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
8	Combined	3.0	1.8	1.2	3.0	1.0	2.0	2.2	1.7	2.1
		(0.0)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)

Note: cells show the rank absolute fractional bias across transformation, and its standard error (in parentheses) is reported. The closer the value to the unity, the smaller the absolute bias; each row indicates a specific stylized set of assumptions (see Appendix B2), for which 50 separate simulations are run and the absolute bias is calculated as the mean fractional bias absolute value.

Source: authors' estimates.

Table 5: Mean within-transform rank absolute fractional bias of alternative IGM estimators under different data transforms and simulated error structures, low-income example

Scenario ‖ Transform →		Levels			Reference distribution			Rank-rank		
		$\hat{\beta}$	$\hat{\Delta}$	up	$\hat{\beta}^d$	$\hat{\Delta}^d$	up^d	$\hat{\beta}^r$	$\hat{\Delta}^r$	up^r
1	Classical ME	2.2	2.8	1.0	2.9	2.1	1.0	2.8	2.2	1.0
		(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.0)
2	Correlated additive MEs	2.1	2.7	1.2	2.7	1.6	1.8	2.5	2.3	1.2
		(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
3	Multiplicative ME, parent	2.8	2.0	1.2	1.6	3.0	1.4	2.0	2.4	1.6
		(0.1)	(0.1)	(0.1)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)
4	Multiplicative ME, child	2.0	3.0	1.0	2.1	2.8	1.1	1.9	2.4	1.7
		(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)
5	Multiplicative ME, both	1.7	3.0	1.3	1.9	2.5	1.6	1.9	2.6	1.5
		(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
6	Incomplete education	2.9	1.9	1.1	2.6	2.3	1.2	2.1	2.4	1.5
		(0.0)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
7	Truncation bias	2.8	2.0	1.2	2.5	2.1	1.4	2.5	2.2	1.3
		(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
8	Combined	2.1	2.9	1.0	3.0	1.8	1.2	2.6	2.4	1.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.0)

Note: cells show the rank absolute fractional bias across estimators, and its standard error (in parentheses) is reported. The closer the value to the unity, the smaller the absolute bias; each row indicates a specific stylized set of assumptions (see Appendix B2), for which 50 separate simulations are run and the absolute bias is calculated as the mean fractional bias absolute value.

Source: authors' estimates.

Second, the magnitude of the mean bias varies considerably across the scenarios and is often substantial. Nearly half of the estimates for the heritability estimator in Table 3 suggest a (mean) fractional bias falling outside a range of +/- 10 per cent of the true value, and one-third of the estimates for the out-performance metric also fall outside this range. Furthermore, there is no indication that the different bias components will necessarily offset one another—the mean fractional bias under the combined scenarios (row 8) is almost always at least as large as under any of the individual scenarios. At the same time, all estimates for the upwards estimator show a (mean) fractional bias within the +/- 10 per cent range. So, this estimator generally appears less biased in general, and, even under the combined scenario (row 8), where the heritability and out-performance estimators generally show significant bias, the fractional bias in the upwards estimator remains within +/- 3 per cent.

Third, confirming the above, Table 5 indicates that within each transform, the upwards mobility estimator is generally ranked best. However, the mean rank is not unity, implying there are cases in which other estimators yield a lower absolute fractional bias. In terms of the different transformations, there also is no unambiguous 'winner'. As Table 4 shows, the absolute fractional bias rankings vary according to the specific scenario and estimator chosen. Broadly, the rank-rank transform performs best in the presence of multiplicative errors (rows 3–5). This reflects the point that such errors constitute monotonic transformations of the true values to which a rank transform will be indifferent.[11] However, the rank of this transform tends to be worst where classical errors are present. Overall, the reference distribution transform appears to perform well, achieving the lowest average ranking under the combined scenario for both the upwards mobility and out-performance estimators.

We recognize these results may be driven by the particular (fixed) assumptions used for the calibrated parameters. Thus, we re-run the same set of measurement error scenarios for a hypothetical middle-income country (see Table B3 for details). The results are summarized in Tables B5–B8 and support our main general conclusions. In particular, the upwards mobility estimator consistently shows lowest (mean fractional) bias, and the reference distribution transform tends to be ranked best. Indeed, for the latter choice we note that the mean fractional bias for both the out-performance and upwards mobility estimators are within +/- 5 per cent of the true value under the combined scenario.

4.3 General scenarios

The previous sub-section considered specific (albeit plausible) regions of the parameter space. We now loosen these restrictions further, permitting all calibrated parameters and measurement error components to vary simultaneously within a broad range (see Appendix Table B4 for details). In addition, to mimic the real world even more closely, we permit a richer structure to the measurement errors. First, we allow the multiplicative errors to vary quadratically with the true values:

$$c_{ij} = c_0 + (1 + \lambda_c - \lambda_q \cdot c_i^* \lambda_c^2 / 3) c_{ij}^* - \delta m_i^* + \nu_i \tag{10}$$

$$p_i = p_0 + (1 + \lambda_p - \lambda_q \cdot p_i^* \lambda_p^2 / 3) p_i^* + \mu_i \tag{11}$$

where parameter $\lambda_q \in (0,1)$ determines the magnitude of this feature, which we vary randomly across simulations using a uniform distribution.

Second, following the point raised by Asher et al. (2020), we consider a restriction on the measurement errors whereby their net contribution can never be so large as to shift an individual across aggregate schooling levels, at least before applying any age effect adjustment. In other words, measurement errors excluding an age effect cannot move an individual from, say, a true grade in upper primary to an observed grade in lower secondary. Equally, if a parent's true highest level of education was upper primary school,

[11] Any bias in this specific case is therefore uniquely driven by bottom- or top-coding observations at 0 and 18 years of education.

then their observed level of schooling can only be within the range of school grades associated with upper primary. This amounts to top- and bottom-coding the observed outcomes (before an age effect) to remain within the true schooling level.

Third, to address bias driven by age effects, in several empirical applications analysts restrict the analytical sample to co-resident children aged over 18 years. While this may focus attention on children who are likely to have completed their schooling, this is not necessarily a panacea. As stated above, this may induce new bias because of systematic selection into leaving the household—e.g., co-resident children aged 18 years may be more likely to be studying. Consequently, we compare results with and without an age restriction. Concretely, in the latter case we seek to maximize the sample size by applying a dynamic age restriction, excluding observations where the age of the child is not sufficient to have accumulated more schooling than their parents.[12]

We run 500 iterations of the simulation and for each iteration we separately apply: none of the above restrictions, each one individually (within-interval errors and dynamic age effect), and both together. For each iteration, parameters/errors are simulated within the chosen ranges using a scrambled Halton sequence, which assures broad coverage of the parameter space. Our results are summarized in Tables 6 and 7, structured as before. Taken as a whole, they corroborate the earlier insights—the combination of the upwards mobility estimator and reference distribution transform is generally preferred. There are also mixed insights in regards to the contribution of the restrictions. For the slope coefficient, each restriction tends to shrink the bias towards zero, such that in combination the bias is much smaller on average (in raw and fractional terms; see Table 6). However, for the out-performance and upwards mobility estimators, the separate effects are ambiguous and in combination do not lead to any dramatic reduction in mean bias relative to the case of no restrictions.

Table 6: Mean bias of alternative IGM estimators under different simulated error structures

	Slope ($\hat{\beta}$)			Percentile diff. ($\hat{\Delta}$)			Upwards mobility		
Assumptions || Transform →	Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
(a) *Mean bias*:									
None	-14.8	-10.5	-14.9	0.2	-0.5	3.9	-0.9	-0.9	3.4
	(0.8)	(0.7)	(0.7)	(0.4)	(0.5)	(0.2)	(0.8)	(0.7)	(0.4)
Within-interval errors	-9.6	-6.3	-8.9	0.8	0.1	4.8	1.6	1.2	5.2
	(0.7)	(0.7)	(0.9)	(0.4)	(0.5)	(0.3)	(0.8)	(0.7)	(0.5)
Dynamic age effect	-10.1	-6.2	-8.0	-0.7	-1.5	2.2	-2.4	-1.8	2.4
	(0.6)	(0.6)	(0.5)	(0.2)	(0.3)	(0.2)	(0.5)	(0.5)	(0.4)
Both	-4.4	-1.7	-1.4	-0.2	-0.8	3.2	0.3	0.6	4.3
	(0.4)	(0.4)	(0.8)	(0.2)	(0.3)	(0.3)	(0.4)	(0.4)	(0.5)
(b) *Mean fractional bias*:									
None	-36.1	-30.2	-34.5	6.0	3.8	28.2	2.1	-1.2	5.1
	(1.9)	(2.2)	(1.7)	(3.9)	(3.7)	(1.7)	(2.0)	(1.8)	(0.6)
Within-interval errors	-23.0	-18.2	-22.1	11.6	8.0	32.7	8.2	4.6	7.5
	(1.8)	(2.1)	(1.9)	(3.9)	(3.7)	(1.8)	(2.0)	(1.6)	(0.7)
Dynamic age effect	-22.2	-15.7	-16.3	-5.1	-5.8	15.8	-3.8	-3.6	3.4
	(1.3)	(1.5)	(1.1)	(3.2)	(3.1)	(1.8)	(1.3)	(1.2)	(0.7)
Both	-8.9	-4.0	-4.6	1.6	0.2	21.4	4.6	3.1	6.1
	(1.0)	(1.2)	(1.4)	(3.0)	(3.0)	(1.9)	(1.3)	(0.9)	(0.7)

Note: cells in panel (a) show the mean bias of each estimator, calculated as per equation (9a); cells in panel (b) show the mean fractional bias of each estimator, calculated as per equation (9b); each row indicates a draw from the general scenarios defined by Appendix Table B4 and based on specific restrictions; 'none' implies no additional restrictions applied; 500 iterations are run, and the mean bias, mean fractional bias, and its respective standard error (in parentheses) are reported.

Source: authors' estimates.

[12] Specifically, in addition to only considering children aged 12 to 24, under this restriction we also exclude children from the observed sample when $y_i < 7 + p_i$. Of course, in the simulated 'true' sample this restriction is not applied because all true values are known.

Table 7: Mean rank absolute bias of alternative IGM estimators under different generalized error structures

(a) *Within estimators*:

	Heritability ($\hat{\beta}$)			Out-performance ($\hat{\Delta}$)			Upwards mobility		
Assumptions ‖ Transform →	Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
None	2.2	1.9	1.9	2.1	1.8	2.0	2.2	1.8	1.8
	(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(0.1)	(0.1)	(0.0)	(0.1)
Within-interval errors	1.9	2.0	2.0	2.0	1.8	2.0	2.1	1.8	1.9
	(0.1)	(0.1)	(0.1)	(0.1)	(0.0)	(0.1)	(0.1)	(0.0)	(0.1)
Dynamic age effect	2.4	1.9	1.6	2.0	1.8	2.1	2.0	1.8	1.9
	(0.1)	(0.1)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Both	2.0	2.0	2.0	1.9	1.8	2.2	2.0	1.7	2.1
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)

(b) *Within transformations*:

	Levels			Reference distribution			Rank-rank		
Assumptions ‖ Estimator →	$\hat{\beta}$	$\hat{\Delta}$	up	$\hat{\beta}^d$	$\hat{\Delta}^d$	up^d	$\hat{\beta}^r$	$\hat{\Delta}^r$	up^r
None	2.3	2.3	1.3	2.4	2.3	1.3	2.5	2.4	1.1
	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Within-interval errors	2.2	2.4	1.4	2.2	2.4	1.4	2.3	2.5	1.2
	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)
Dynamic age effect	2.3	2.3	1.3	2.3	2.4	1.3	2.2	2.5	1.3
	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)
Both	2.1	2.4	1.5	2.1	2.4	1.4	2.0	2.6	1.4
	(0.1)	(0.1)	(0.0)	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)

Note: cells in panel (a) show the rank absolute fractional bias across transformation; cells in panel (b) show the rank absolute fractional bias across estimator; each row indicates a draw from the general scenarios defined by Appendix Table B4 and based on specific restrictions; 'none' implies no additional restrictions applied; 500 iterations are run, and the mean bias, mean fractional bias, and its respective standard error (in parentheses) are reported.

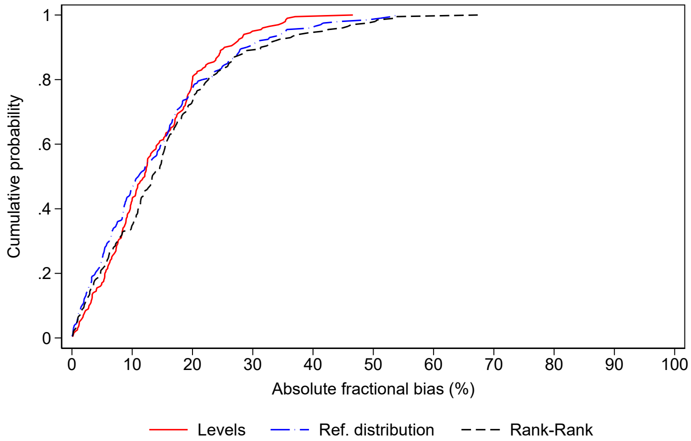
Source: authors' estimates.

Figure 2 plots the empirical cumulative distribution functions of the absolute fractional errors from the three different estimators across the 500 iterations, distinguishing between the three different transformations, based on applying both restrictions. In all plots the rank-rank transform lies furthest to the right for at least two-thirds of its cumulative distribution, implying it is associated with a somewhat larger absolute bias. In contrast, the reference distribution tends to lie furthest to the left. We also observe a sharper incline to the plots for the upwards mobility estimator (panel c), indicating relatively more mass lies closer to zero (e.g., less than 20 per cent absolute bias).

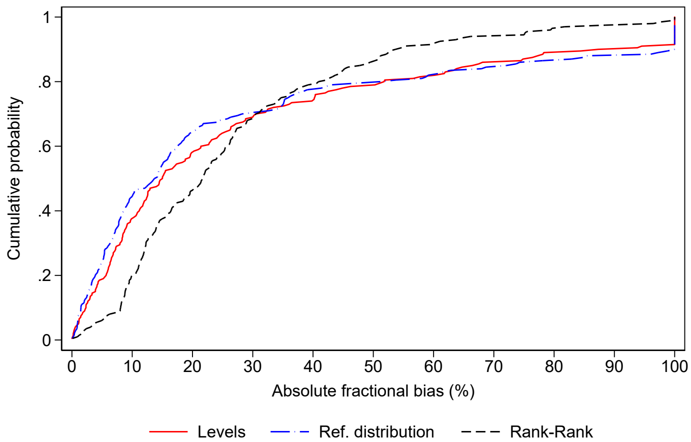
Box plots in Figure 3, which capture the mean fractional bias, underpin this conclusion. They show the reference distribution is most consistently centered on zero, implying it does not show any systematic bias across our simulations. And the bias is clearly smallest for the upwards mobility estimator. However, differences across measurement choices are not completely clear-cut, and no combination of estimator or transform is guaranteed to yield the lowest absolute fractional bias relative to the alternatives. Indeed, calculating ranks across all nine combinations of transforms and estimators, we find our preferred choice is ranked first on 25 per cent of iterations (and in the top three two-thirds of the time) but is placed in the bottom three on 4 per cent of iterations.

Figure 2: Comparison of empirical cumulative distribution functions of absolute fractional bias from alternative transformations, by estimator

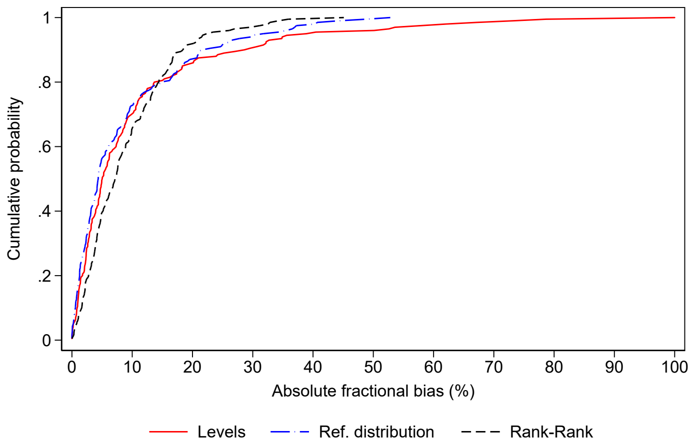
(a) *Heritability*:



(b) *Out-performance*:



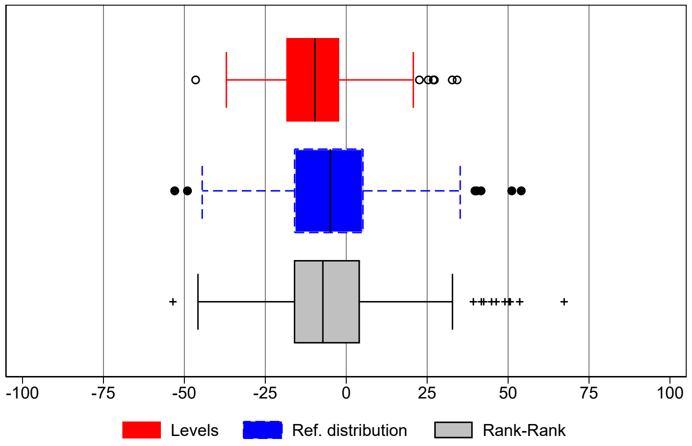
(c) *Upwards mobility*:



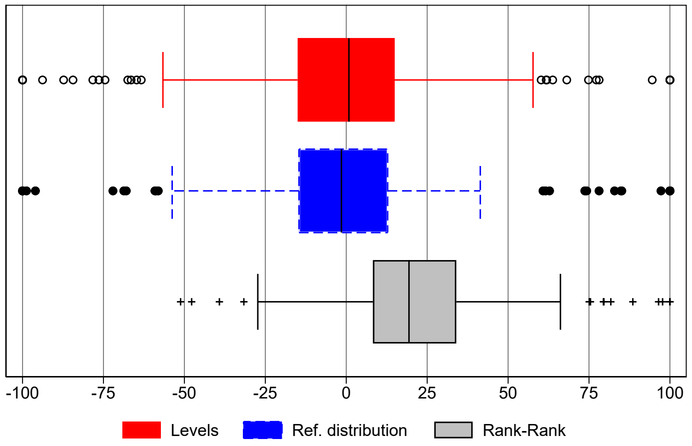
Source: authors' estimates.

Figure 3: Comparison of empirical distributions of mean fractional bias from alternative transformations, by estimator

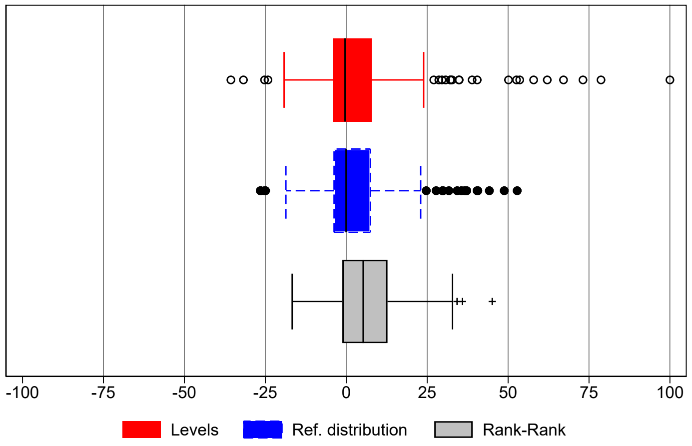
(a) *Heritability*:



(b) *Out-performance*:



(c) *Upwards mobility*:



Source: authors' estimates.

4.4 Practical implications

In this section, we developed a flexible simulation model to investigate the implications of measurement error, including sample truncation, for studying intergenerational mobility. Reflecting on the variety of metrics applied in the literature, we considered three different estimators of mobility as well as three alternative data transformations. This enabled us to quantify how the magnitude of bias varies under different stylized scenarios and whether there are differences in bias associated with specific measurement choices.

To pull this together, we summarize some of the *practical* implications of our findings. The main lessons are as follows:

1. Despite its continued widespread use in the literature, the heritability estimator (slope coefficient) tends to perform poorly. Across our unrestricted general simulations we found a mean absolute fractional bias of over 30 per cent (or less than -30 per cent on average; also see Figure 3a) in these estimates. Moreover, contrary to recommendations in the literature, applying some form of rank transform to scale the education variables did not yield significant reductions in the bias of this estimator. As such, unique reliance on the heritability coefficient to inform IGM analysis may be misguided.

2. Use of a rank-rank transformation, as advocated in some recent studies (Chetty et al. 2014; Emran and Shilpi 2019; Eriksen and Munk 2020; Karlson and Landersø 2021), appears to perform adequately according to the measures of bias used herein. However, in removing all level effects from the data, application of this transform exclusively provides insights into relative aspects of mobility. Also, in the specific case of the out-performance estimator, this transform is associated with a particularly substantial mean fractional bias in the positive domain—i.e. it overestimates the degree of out-performance by around 20 per cent (on average over our simulations). So, use of this transform may not always be appropriate, particularly where absolute dimensions of mobility are of interest.

3. In contrast, the reference distribution transform appears to combine advantages of the two others—mapping child outcomes to the percentile ranks of the parental education distribution provides a consistent means to scale the data while also permitting meaningful absolute comparisons (over time and space). Furthermore, our simulation evidence indicated this transform is robust, in the sense of being most often associated with the lowest bias of all transforms for any given estimator.

4. Across the three estimators, we found that the upwards mobility metric is least biased—in over two-thirds (half) of simulations it delivered an absolute fractional bias of less than 10 (5) per cent. Nonetheless, it should be recognized that this property is likely to be a function of the binary score underlying this metric. Upwards mobility simply indicates the share of children born to parents with below-median education who are expected to accumulate more education than their parents. The out-performance metric effectively captures how much more and, as such, is a natural complement.

5. We showed that assuming errors remain restricted within broad schooling intervals can reduce the bias of some estimators, and similar gains are also obtained from applying a dynamic age restriction. However, these gains are mostly concentrated in the slope coefficient; both the out-performance and upwards mobility metrics are largely unaffected, further reinforcing their general robustness.

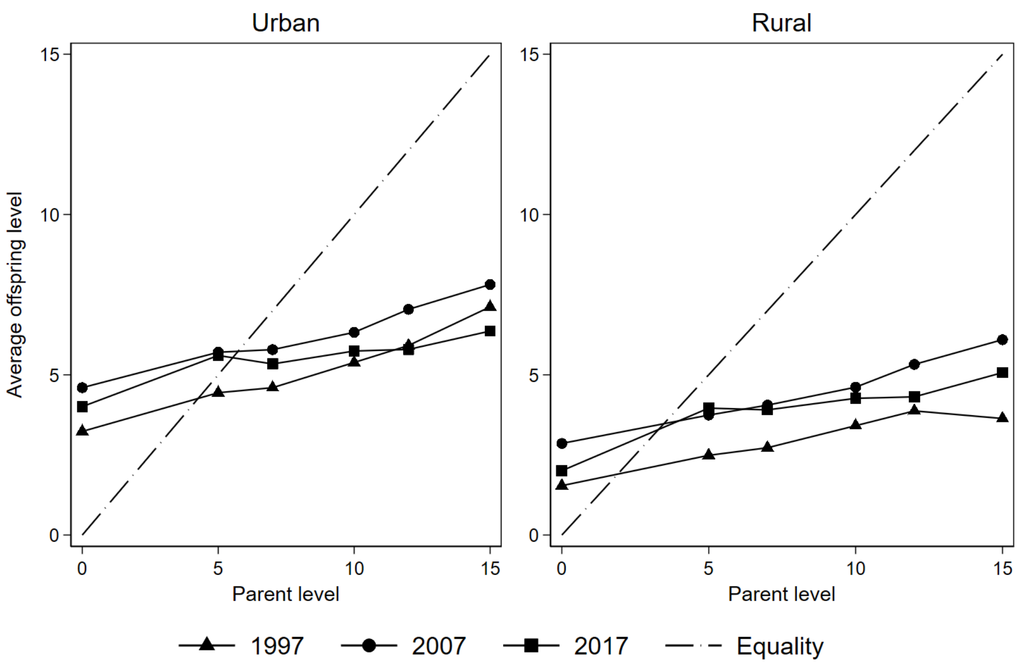
In sum, combining both theoretical insights and simulation evidence, we recommend IGM is studied by applying a reference distribution transformation (at the level of analysis of interest) with an upwards mobility estimator, complemented by the out-performance estimator. Nonetheless, as this combination of metrics is neither guaranteed to be unbiased nor even to minimize bias, careful sensitivity checks are likely to be helpful. For instance, if results vary dramatically across transformations or when applying a (dynamic) age restriction, this should be considered a red flag. Moreover, it must be recalled that different estimators measure different dimensions of mobility. So, with due attention to possible bias, a comparative analysis deploying different estimators is likely to be informative.

5 Application to Mozambique

Thus far we have set out practical guidance for undertaking empirical studies of IGM. In this section we apply this guidance, focusing on the case of Mozambique where we use the complete records of censuses undertaken in 1997, 2007, and 2017 to explore the magnitude of IGM and the extent to which this has varied over time and place. To begin, we echo Figure 1 and collapse the data in each census to obtain the average level of education among children (aged 12–24) for each level of parental education. We plot these relationships in Figure 4 by census year, treating rural and urban areas separately.

The plots yield immediate insights. First, the child-parent education relationship is approximately linear, suggesting the general framework based on equation (1) is likely to be sufficient. Second, while the slope seems fairly stable over time, there are material differences across years in the intercepts, which indicate the average years of schooling of children born to parents who did not attend school. In both areas, the intercept shifts upward by around one year from 1997–2007, but from 2007–17 there is no clear improvement. The change in intercept is reflected in changes in the intersection point with the 45 degree line. In urban areas in 1997, children born to parents with less than four years of formal education were expected to out-perform their parent's educational attainment. In 2007 and 2017, this had increased to just over five years. In rural areas, however, educational attainment remains much lower. Even in 2017, only children born to parents with less than three years of formal education were expected to accumulate more schooling (out-perform) than their parents. So, not only does this highlight the ongoing scale of the education challenge in the country but also that a unique focus on heritability is likely to be incomplete.

Figure 4: Mean education of children vs. parents in Mozambique, by area of residence, 1997–2017



Source: authors' estimates.

Table 8 digs deeper, looking at IGM across provinces over time. Following the guidance set out in Section 4.4, we apply the reference distribution transform, calculated separately for each province and census, from which we then calculate all three mobility estimators.[13] Taking the (most robust) upwards mobility metric first, there is almost no change on aggregate between 1997 and 2017. Namely, just as in 1997, about 50 per cent of children living with parents who had no more than a median education in 2017 were expected to accumulate *more* education than their parents. Critically, this represents a reduction since 2007, when the upwards mobility metric was around 72 per cent, thus indicating that gains in access to schooling achieved in the first period (1997–2007) have not been sustained, despite low average levels of attainment in the population.

Table 8: Relative and absolute mobility measures using a reference distribution transformation (no age restriction), at province level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Upwards mobility			Out-performance			Heritability		
Province	1997	2007	2017	1997	2007	2017	1997	2007	2017
Niassa	0.372	0.601	0.395	0.223	0.264	0.138	0.380	0.334	0.375
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Cabo Delgado	0.363	0.632	0.336	0.230	0.281	0.118	0.288	0.271	0.377
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Nampula	0.392	0.637	0.352	0.201	0.248	0.126	0.355	0.320	0.415
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Zambezia	0.420	0.682	0.351	0.251	0.292	0.156	0.352	0.291	0.440
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Tete	0.387	0.625	0.425	0.250	0.300	0.170	0.392	0.323	0.399
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Manica	0.462	0.810	0.527	0.256	0.366	0.214	0.398	0.246	0.304
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Sofala	0.429	0.766	0.523	0.211	0.317	0.208	0.460	0.312	0.349
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Inhambane	0.664	0.860	0.736	0.364	0.427	0.329	0.260	0.220	0.237
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Gaza	0.656	0.854	0.685	0.352	0.409	0.319	0.277	0.255	0.287
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Maputo Province	0.729	0.750	0.694	0.298	0.351	0.254	0.311	0.245	0.237
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Maputo City	0.548	0.806	0.780	0.274	0.338	0.271	0.276	0.219	0.186
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Mozambique	0.483	0.719	0.493	0.243	0.300	0.200	0.408	0.335	0.361
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)

Note: cells report mean of mobility estimators, calculated separately for each census year and province using a reference distribution; 'Mozambique' is calculated for the whole country; no dynamic age effect restriction applied; see Table B10 for corresponding estimates with dynamic age effect; standard errors in parentheses.

Source: authors' estimates.

The same general pattern of results is reflected in the out-performance metric, although here we observe a moderate improvement between 1997 and 2007 and a larger from 2007–17. Taking the country as a whole, we see that a child born to a parent located at the 25th percentile of parental educational attainment in 2017 was only expected to be placed at the 45th percentile on the latter distribution. This constitutes an expected 20 point gain, compared to an expected 30 point improvement in 2007. However, national-level results mask very substantial provincial heterogeneity, with a strong and marked north-south gradient in mobility. A number of provinces show improvements in upwards mobility from 1997–2017; however, only the southern provinces (Inhambane, Gaza, Maputo City, and Maputo Province) have consistently maintained rates of upwards mobility in excess of 50 per cent throughout the observation period. This is despite the fact that these same provinces have always shown the highest levels of parental education on average (Table B9). Provinces in the north, such as Cabo Delgado, have shown much lower rates of upward mobility. These rates were marginally lower in 2017 than in 1997. Today, around one-third of children born to parents located at or below the 50th percentile in the north are expected to out-perform

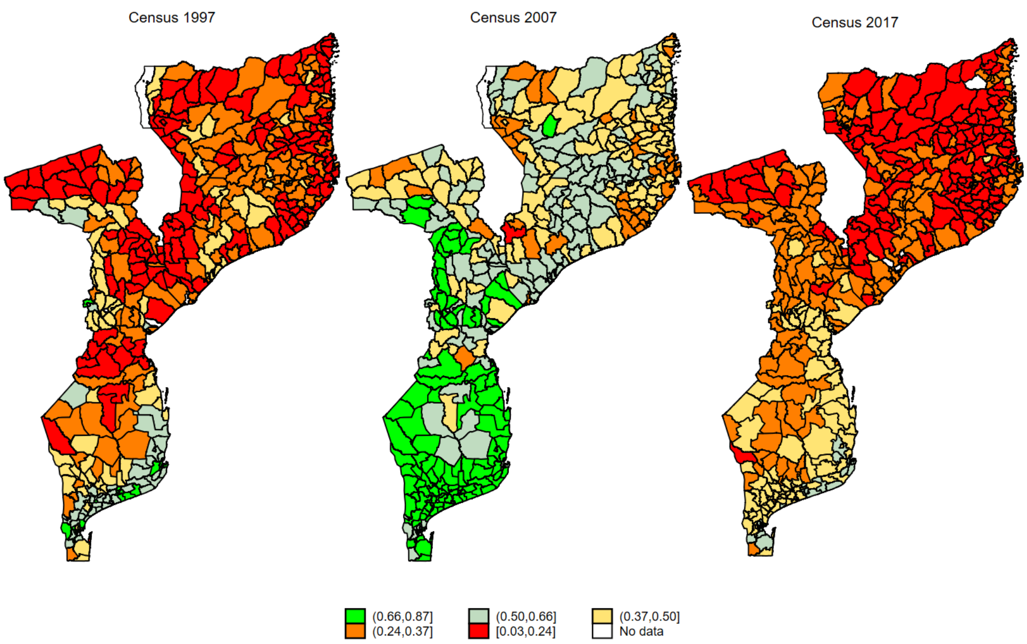
[13] As such, all estimates are benchmarked to the distribution of parental outcomes in a given province and census year. It would also be feasible to fix the reference distribution over time and/or place.

their parents versus two-thirds in the south. Put simply, the education system is not delivering consistent gains to children, and those outside the south of the country are especially disadvantaged.

The out-performance metrics further nuance these findings. In 2017, not only was the degree (magnitude) of out-performance nearly three times higher in Maputo City compared to Cabo Delgado (0.271 vs. 0.118), but this compares to a similar multiple of less than 1.5 in 1997 and 2007. From this perspective, the weakening of IGM in the north (and, to a somewhat lesser extent, the centre) over the latest period is particularly dramatic. Although more likely to be biased downwards, estimates for the heritability in outcomes are consistent with this narrative. On aggregate, heritability seems to have improved over time (fallen in magnitude), but this appears to be driven exclusively by developments in the south, especially in Maputo City where heritability has halved over time. But in the north (and centre), heritability has actually worsened—e.g., increasing from 0.288 to 0.377 in Cabo Delgado over the period 1997–2007.

To complement our analysis and further exploit the disaggregated nature of our data, we present a graphical analysis. Figure 5 presents results for the (preferred, robust) upwards mobility estimator calculated at the third administrative level (*posto administrativo*) for each census. This starkly displays that children in the coastal south of the country enjoy the highest likelihood of out-performing their parents (conditional on their parents having no more than a median-level of education for that location), and this likelihood has remained comparatively high across all periods. In contrast, despite a general improvement from 1997–2007, upwards mobility has been and remains lowest in the northern provinces (Cabo Delgado, Niassa, Nampula, and Tete). Similar differences stand out from both the out-performance and heritability mapping exercises, shown in Figures 6 and 7. These underline the findings of a general slowdown in mobility—especially in the north—over the most recent period.

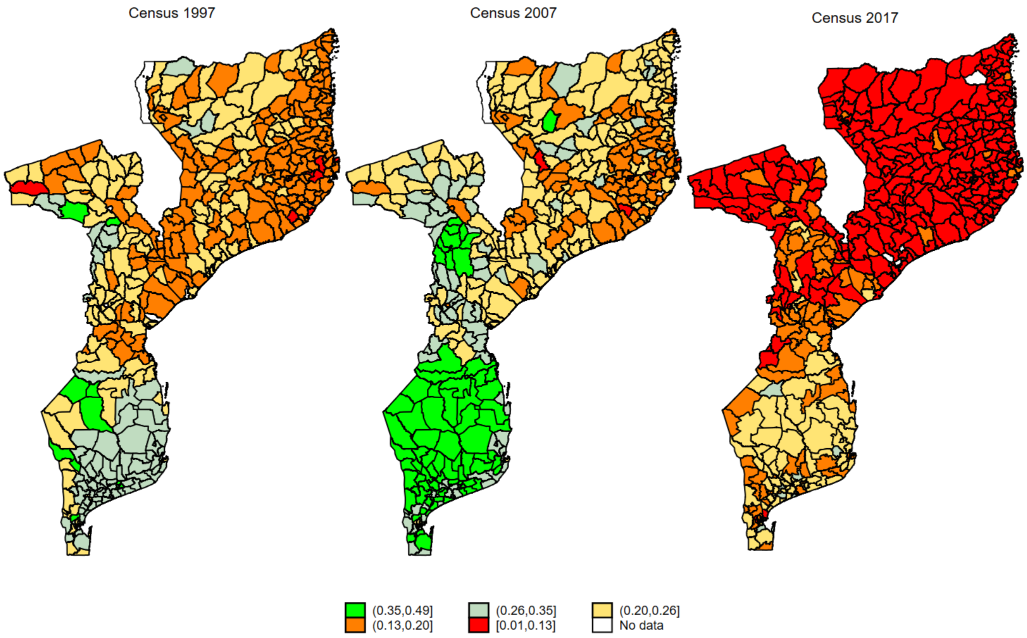
Figure 5: Upwards mobility in Mozambique, using a reference distribution transformation (no dynamic age restriction), at *posto administrativo* level



Note: graphs report mean of upwards mobility estimator, calculated separately for each census year and posto administrativo using a reference distribution; no dynamic age effect restriction applied; see Figure B1 for corresponding estimates with dynamic age effect.

Source: authors' estimates.

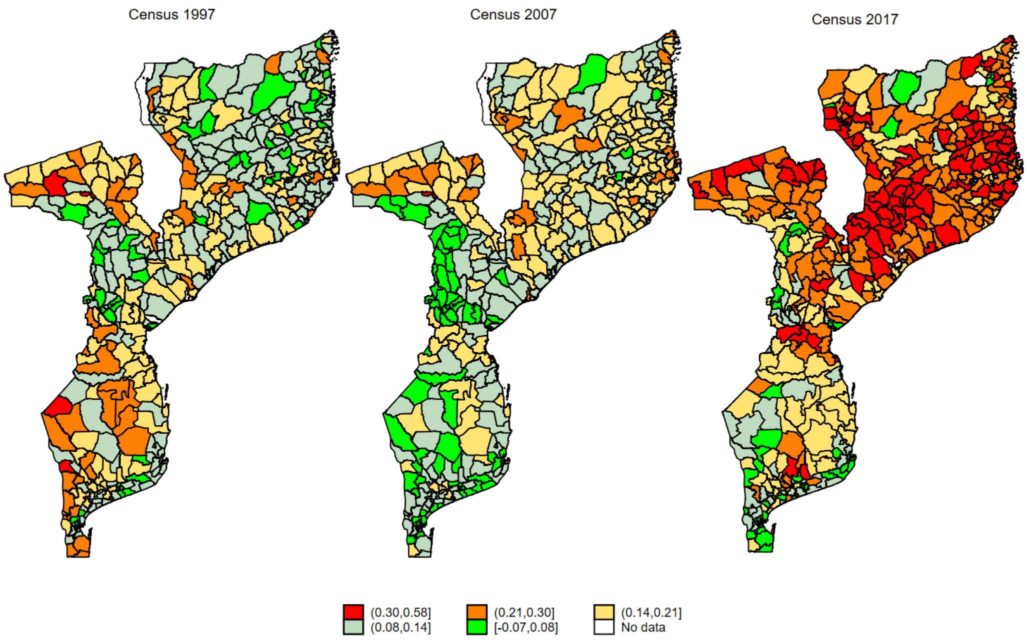
Figure 6: Out-performance in Mozambique using a reference distribution transformation (no dynamic age restriction), at *posto administrativo* level



Note: graphs report mean of out-performance mobility estimator, calculated separately for each census year and posto administrativo using a reference distribution; no dynamic age effect restriction applied; see Figure B2 for corresponding estimates with dynamic age effect.

Source: authors' estimates.

Figure 7: Heritability in Mozambique using a reference distribution transformation (no dynamic age restriction), at *posto administrativo* level



Note: graphs report mean of the heritability estimator, calculated separately for each census year and posto administrativo using a reference distribution; no dynamic age effect restriction applied; see Figure B3 for corresponding estimates with dynamic age effect.

Source: authors' estimates

As a final exercise, we calculate the intersection point for each province and census, indicating the level of parental education below which the average child is expected to out-perform their parents. As reported in Table 9, the intersection point increased from 1997–2007 but fell dramatically in 2007–17, with larger falls in the north than the south. In raw-level terms (also shown in the table), we should not ignore general improvements over time in the intersection point. But even here we see regression in the most recent period—today, on average, we only expect children born to parents with less than 6.1 years of schooling to accumulate more education than their parents versus a threshold of 6.6 years in 2007.

Table 9: Point of intersection below which children are expected to accumulate more education than their parents on average (no dynamic age restriction), by province

	Reference distribution			Levels		
Province	1997	2007	2017	1997	2007	2017
Niassa	0.610	0.646	0.470	3.202	5.387	5.180
	(0.001)	(0.001)	(0.001)	(0.017)	(0.014)	(0.017)
Cabo Delgado	0.573	0.636	0.440	2.709	5.114	4.481
	(0.001)	(0.001)	(0.001)	(0.012)	(0.011)	(0.016)
Nampula	0.561	0.614	0.466	3.160	5.271	4.900
	(0.001)	(0.001)	(0.001)	(0.009)	(0.007)	(0.010)
Zambezia	0.637	0.662	0.528	3.461	5.470	4.860
	(0.001)	(0.001)	(0.001)	(0.008)	(0.007)	(0.009)
Tete	0.661	0.692	0.532	3.568	5.615	5.564
	(0.001)	(0.001)	(0.001)	(0.014)	(0.011)	(0.012)
Manica	0.675	0.735	0.557	4.248	6.685	6.084
	(0.001)	(0.001)	(0.001)	(0.014)	(0.010)	(0.011)
Sofala	0.640	0.711	0.570	4.533	6.933	6.379
	(0.001)	(0.001)	(0.001)	(0.014)	(0.010)	(0.011)
Inhambane	0.741	0.797	0.682	5.000	7.279	6.654
	(0.001)	(0.001)	(0.001)	(0.012)	(0.013)	(0.011)
Gaza	0.737	0.799	0.698	4.908	7.314	6.534
	(0.001)	(0.001)	(0.001)	(0.011)	(0.013)	(0.013)
Maputo Province	0.682	0.715	0.582	5.766	8.046	7.455
	(0.001)	(0.001)	(0.001)	(0.011)	(0.009)	(0.008)
Maputo City	0.628	0.683	0.583	6.747	8.925	8.325
	(0.001)	(0.001)	(0.001)	(0.008)	(0.009)	(0.010)
Mozambique	0.661	0.701	0.563	4.474	6.585	6.091
	(0.000)	(0.000)	(0.000)	(0.004)	(0.003)	(0.003)

Note: cells report mean of mobility estimators, calculated separately for each census year and province using a reference distribution; 'Mozambique' is calculated for the whole country; no dynamic age effect restriction applied; the intersection point indicates the intersection with the line of equality and, thus, the highest level of parental education below which offspring outcomes are expected to exceed those of their parents, on average; see Table B10 for corresponding estimates with dynamic age effect; standard errors in parentheses.

Source: authors' estimates.

6 Conclusion

We began this paper noting that, despite the importance attributed to intergenerational mobility in the process of socio-economic development, empirical researchers face a bewildering set of choices. Moreover, despite recognition that conventional metrics may be (badly) biased, there has been little consensus on how this should be addressed. As such, our objective was to contribute analytical and empirical evidence regarding the sensitivity of alternative IGM estimators to different forms of measurement errors and data transformations. Put simply, our primary contribution was to investigate the magnitude of bias associated with different measurement choices.

Analytically, we showed that the bias associated with different estimators of IGM can run in different directions, depending on the specific source and valence of measurement error. When multiple sources of error are combined, including sample truncation, the net bias cannot be easily determined *ex ante*. To make progress, we developed a flexible simulation model. For a wide range of plausible scenarios, this allowed us to explore the direction and magnitude of bias associated with different combinations of mobility estimators (the heritability coefficient, out-performance, and upwards mobility) and three data transformations (raw data in levels, a reference distribution transformation, and rank-rank percentiles). Using this framework, we found that the heritability estimator (the slope coefficient), widely used in the literature, tends to perform poorly. In our generalized simulation, covering a broad parameter space, we found this estimator is biased downwards by around 30 per cent on average. Furthermore, rank-rank-based measures, such as suggested by Chetty et al. (2014), did not seem to yield significant reductions in the bias of this estimator. We also found that the out-performance estimator presents an upwards bias around 20 per cent when using a rank-rank transformation. We concluded that reliance on estimates from the heritability coefficient as well as use of the rank-rank transformation may be particularly misleading.

In contrast, our results suggest that a reference distribution transform is associated with the lowest bias, on average, across all estimators. This transform also has the advantage of permitting meaningful (absolute) comparisons over time and space, such as where mean levels of education differ. Evidence from simulations suggests that an upwards mobility metric is generally least biased across the three estimators—indeed, in over two-thirds of the simulations the absolute fractional bias associated with this estimator is less than 10 per cent. Nevertheless, it merits note that the upwards mobility estimator simply indicates the share of children born to parents with below-median education who are expected to accumulate at least some more education than their parents. Thus, to better understand IGM, we suggest that results from an upwards mobility estimator should be complemented with an out-performance metric, which captures how much more education than their parents children born to the median parent is expected to achieve.

Taking guidance from the simulations, we moved into an empirical analysis using complete records of censuses undertaken in 1997, 2007, and 2017 in Mozambique. Our results suggest a high degree of provincial heterogeneity, with a strong and marked north-south gradient in mobility. Moreover, there has been a clear slowdown in IGM between 2007 and 2017—especially in the northern region—suggesting the education system is not delivering consistent gains to children across the country. When only focusing on the heritability metric, there have been improvements over time, driven by developments in the south. This is our point, derived from the simulation exercises, that the heritability metric is not a reliable or comprehensive guide to IGM.

In sum, this article sought to shed light on the challenges of estimating IGM, especially in developing countries where data are limited. Our evidence suggests that a using the upwards mobility estimator, complemented by an out-performance metric, based on a reference distribution transformation, is comparatively most robust. That is, while these two metrics are unlikely to be unbiased, the bias is generally centred on zero and tends to be small relative to other measurement choices. Using these methods revealed the significant challenges facing Mozambique today.

References

Alesina, A., Hohmann, S., Michalopoulos, S., and Papaioannou, E. (2019). 'Intergenerational Mobility in Africa'. NBER Working Paper 25534. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w25534

Andrews, D., and Leigh, A. (2009). 'More Inequality, Less Social Mobility'. *Applied Economics Letters*, 16(15): 1489–92. https://doi.org/10.1080/13504850701720197

Asher, S., Novosad, P., and Rafkin, C. (2020). 'Intergenerational Mobility in India: New Methods and Estimates Across Time, Space, and Communities'. Unpublished manuscript. Available at: https://paulnovosad.com/

pdf/anr-india-mobility.pdf
Becker, G. S., and Tomes, N. (1979). 'An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility'. *Journal of Political Economy*, 87(6): 1153. https://doi.org/10.1086/260831
Becker, G. S., and Tomes, N. (1986). 'Human Capital and the Rise and Fall of Families'. *Journal of Labor Economics*, 4(3): 1. https://doi.org/10.1086/298118
Bertrand, M., and Mullainathan, S. (2001). 'Do People Mean What They Say? Implications for Subjective Survey Data'. *American Economic Review*, 91(2): 67–72. https://doi.org/10.1257/aer.91.2.67
Björklund, A., and Jäntti, M. (2009). 'Intergenerational Income Mobility and the Role of Family Background'. In B. Nolan, W. Salverda, and T. M. Smeeding (eds), *The Oxford Handbook of Economic Inequality* (pp. 491–521). https://doi.org/10.1093/oxfordhb/9780199606061.013.0020
Björklund, A., Lindahl, M., and Plug, E. (2006). 'The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data'. *The Quarterly Journal of Economics*, 121(3): 999–1028. https://doi.org/10.1162/qjec.121.3.999
Black, D., Sanders, S., and Taylor, L. (2003). 'Measurement of Higher Education in the Census and Current Population Survey'. *Journal of the American Statistical Association*, 98(463): 545–54. https://doi.org/10.1198/016214503000000369
Black, S. E., and Devereux, P. J. (2011). 'Recent Developments in Intergenerational Mobility'. In D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics* (Vol. 4, pp. 1487–541). Amsterdam: Elsevier.
Card, D., Domnisoru, C., and Taylor, L. (2018). 'The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility'. NBER Working Paper 25000. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w25000
Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). 'Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States'. *The Quarterly Journal of Economics*, 129(4): 1553–623. https://doi.org/10.1093/qje/qju022
Emran, M. S., Greene, W., and Shilpi, F. (2018). 'When Measure Matters: Coresidency, Truncation Bias, and Intergenerational Mobility in Developing Countries'. *Journal of Human Resources*, 53(3): 589–607. https://doi.org/10.3368/jhr.53.3.0216-7737R1
Emran, M. S., and Shilpi, F. (2019). 'Economic Approach to Intergenerational Mobility: Measures, Methods, and Challenges in Developing Countries'. Working Paper 98/2019. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2019/734-7
Eriksen, J., and Munk, M. D. (2020). 'The Geography of Intergenerational Mobility — Danish Evidence'. *Economics Letters*, 189(–): 109024. https://doi.org/10.1016/j.econlet.2020.109024
Fox, L., Santibanez, L., Nguyen, V., and Andre, P. (2012). *Education Reform in Mozambique : Lessons and Challenges.* Washington, DC. https://doi.org/10.1596/978-0-8213-8975-1
Gibbons, C. E., Serrato, J. C. S., and Urbancic, M. B. (2019). 'Broken or Fixed Effects?' *Journal of Econometric Methods*, 8(1): 20170002. https://doi.org/10.1515/jem-2017-0002
Glewwe, P. (2012). 'How Much of Observed Economic Mobility Is Measurement Error? IV Methods to Reduce Measurement Error Bias, with an Application to Vietnam'. *The World Bank Economic Review*, 26(2): 236–64. https://doi.org/10.1093/wber/lhr040
Glewwe, P., and Muralidharan, K. (2016). 'Chapter 10 – Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications'. In E. A. Hanushek, S. Machin, and L. Woessmann (eds), *Handbook of the Economics of Education* (Vol. 5, pp. 653–743). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-444-63459-7.00010-5
Handcock, M. S., and Morris, M. (1998). 'Relative Distribution Methods'. *Sociological Methodology*, 28(1): 53–97. https://doi.org/10.1111/0081-1750.00042
Harwell, M. (2019). 'A Strategy for Using Bias and RMSE as Outcomes in Monte Carlo Studies in Statistics'. *Journal of Modern Applied Statistical Methods*, 17(2): eP2938. https://doi.org/10.22237/jmasm/1551907966
Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2007). 'The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends'. *The B.E. Journal of Economic Analysis & Policy*, 7(2): –. https://doi.org/10.2202/1935-1682.1775
Karlson, K., and Landersø, R. (2021). 'The Making and Unmaking of Opportunity: Educational Mobility in 20th Century-Denmark'. SSRN Scholarly Paper No. ID 3794081. Rochester, NY: Social Science Research Network.
Landersø, R., and Heckman, J. J. (2016). 'The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the U.S.'. Working Paper 22465. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w22465
Mann, H. (1868). *Life and Works of Horace Mann* (Vol. 3). Boston: Walker, Fuller.

Morris, M., Bernhardt, A. D., and Handcock, M. S. (1994). 'Economic Inequality: New Methods for New Trends'. *American Sociological Review*, 59(2): 205–19. https://doi.org/10.2307/2096227

Nybom, M., and Stuhler, J. (2017). 'Biases in Standard Measures of Intergenerational Income Dependence'. *Journal of Human Resources*, 52(3): 800–25. https://doi.org/10.3368/jhr.52.3.0715-7290R

Solon, G. (1992). 'Intergenerational Income Mobility in the United States'. *The American Economic Review*, 82(3): 393–408.

Appendix A: Additional derivations

As stated in the main text, there are at least three generic challenges when estimating IGM. To capture the implications of these challenges, analytical characterization of the expected bias in our mobility measures is helpful. To do so, we extend the generalized errors-in-variables approach suggested by Nybom and Stuhler (2017), which applies to a wide range of data transformations, taking into account the additional complexities associated with education data.

The structure presented in equation (7) is shown for data expressed in levels. It also applies naturally to the two-rank transforms; however, the error parameters will not be identical when moving from one transform to another (with the same data). Therefore, using the system in equation (3d), we now provide expressions for the bias of the heritability estimator and the out-performance estimator under various transformations.

Data in levels:

$$\hat{\beta}-\beta=\beta\frac{(1+\lambda_c)}{(1+\lambda_p)}\left(1-\frac{\text{var}(\mu_i)}{\text{var}(p_i)}\right)-\delta(1+\lambda_p)\frac{\text{cov}(m_i^*,p_i^*)}{\text{var}(p_i)}+\frac{\text{cov}(\nu_i,\mu_i)}{\text{var}(p_i)}-\beta$$

$$\begin{aligned}\hat{\Delta}_{p_{25}}-\Delta_{p_{25}^*}=&[\text{E}(c_i)-\text{E}(c_i^*)]-(\hat{\beta}-\beta)\,(\text{E}(p_i)-p_{25})\\&+\beta(\text{E}(p_i^*)-\text{E}(p_i))+(1-\beta)(p_{25}^*-p_{25})\end{aligned}$$

Reference distribution:

$$(\hat{\beta}^d-\beta^d)/12=\beta\frac{(1+\lambda_c^d)}{(1+\lambda_p^d)}\left(1/12-\text{var}(\mu_i^d)\right)-\delta(1+\lambda_p^d)\text{cov}(m_i^{*d},p_i^{*d})+\text{cov}(\nu_i^d,\mu_i^d)-\beta^d/12$$

$$\hat{\Delta}_{p_{25}}^d-\Delta_{p_{25}^*}^d=[\text{E}(c_i^d)-\text{E}(c_i^{*d})]-0.25(\hat{\beta}^d-\beta^d)$$

Rank-rank:

$$(\hat{\beta}^r-\beta^r)/12=\beta^r\frac{(1+\lambda_c^r)}{(1+\lambda_p^r)}\left(1/12-\text{var}(\mu_i^d)\right)-\delta(1+\lambda_p^r)\text{cov}(m_i^{*r},p_i^{*r})+\text{cov}(\nu_i^r,\mu_i^r)-\beta^r/12$$

$$\hat{\Delta}_{p_{25}}^r-\Delta_{p_{25}^*}^r=-0.25(\hat{\beta}^r-\beta^r)$$

Appendix B: Additional tables and figures

B1 Simulation tables

Table B1: Summary of selected mobility measures

Mobility dimension		Exp.	Transformation		
			None (levels)	Reference distribution	Rank-rank
Rel.	Heritability	$\beta^{\cdot}$	$\text{Cov}(c^*, p^*)/\text{Var}(p^*)$	$12 \cdot \text{Cov}(c^{*d}, p^{*d})$	$12 \cdot \text{Cov}(c^{*r}, p^{*r})$
Abs.	Out-perform.	$\Delta^{\cdot}_{p_{25}}$	$\text{E}(c^*) - \beta\left(\text{E}(p^*) - p^*_{25}\right) - p^*_{25}$	$\text{E}(c^{*d}) - 0.25(1+\beta^d)$	$0.25(1-\beta^r)$
	Upward mob.	$\text{up}^{\cdot}$	$\text{E}(c^*_i > p^*_i \mid p^*_i \leq p_{50})$	$\text{E}(c^{*d}_i > p^{*d}_i \mid p^{*d}_i \leq 0.5)$	$\text{E}(c^{*q}_i > p^{*q}_i \mid p^{*q}_i \leq 0.5)$

Source: authors' elaboration.

Table B2: Stylized scenarios—low-income country

Scenario	a	b	α	β	f_0	Δ_α	Δ_β	δ	c_0	p_0	λ_1	λ_2	γ	θ_1	θ_2
Classical ME	1	3.2	2	0.45	0.5	0	0	0	0	0	0	0	0	0.3	0.3
Correlated additive MEs	1	3.2	2	0.45	0.5	0	0	0	0	0	0	0	0.4	0.3	0.3
Multiplicative ME, parent	1	3.2	2	0.45	0.5	0	0	0	0	0	0	0.1	0	0	0
Multiplicative ME, child	1	3.2	2	0.45	0.5	0	0	0	0	0	0.1	0	0	0	0
Multiplicative ME, both	1	3.2	2	0.45	0.5	0	0	0	0	0	0.1	0.1	0	0	0
Incomplete education	1	3.2	2	0.45	0.5	0	0	1	0	0	0	0	0	0	0
Truncation bias	1	3.2	2	0.45	0.5	-0.5	0.1	0	0	0	0	0	0	0	0
Combined	1	3.2	2	0.45	0.5	-0.5	0.1	1	0	0	0.1	0.1	0.3	0.3	0.3

Note: the scenarios calibrate to a typical low-income country with low average levels of parental education. For each chosen scenario we run 50 iterations of the simulation model.

Source: authors' elaboration.

Table B3: Stylized scenarios—middle-income country

Scenario	a	b	α	β	f_0	Δ_α	Δ_β	δ	c_0	p_0	λ_1	λ_2	γ	θ_1	θ_2
Classical ME	2	5	3.5	0.55	0	0	0	0	0	0	0	0	0	0.3	0.3
Correlated additive MEs	2	5	3.5	0.55	0	0	0	0	0	0	0	0	0.4	0.3	0.3
Multiplicative ME, parent	2	5	3.5	0.55	0	0	0	0	0	0	0	0.1	0	0	0
Multiplicative ME, child	2	5	3.5	0.55	0	0	0	0	0	0	0.1	0	0	0	0
Multiplicative ME, both	2	5	3.5	0.55	0	0	0	0	0	0	0.1	0.1	0	0	0
Incomplete education	2	5	3.5	0.55	0	0	0	1	0	0	0	0	0	0	0
Truncation bias	2	5	3.5	0.55	0.25	-0.5	0.1	0	0	0	0	0	0	0	0
Combined	2	5	3.5	0.55	0.25	-0.5	0.1	1	0	0	0.1	0.1	0.3	0.3	0.3

Note: the scenarios calibrate to a typical middle-income country with low average levels of parental education. For each chosen scenario we run 50 iterations of the simulation model.

Source: authors' elaboration.

Table B4: Assumptions about parameters' range and distributions—general simulations

Parameter	Min	Max	Distribution
a	0.5	5	Uniform
b	1	10	Uniform
α	1	7	Uniform
β	0.2	0.8	Uniform
f_0	0	0.5	Uniform
Δ_α	-1	1	Uniform
Δ_β	-0.2	0.2	Uniform
δ	0	1	Uniform
c_0	-0.1	0.1	Uniform
p_0	-0.1	0.1	Uniform
λ_c	-0.2	0.2	Uniform
λ_p	-0.2	0.2	Uniform
γ	-0.5	0.5	Uniform
θ_1	0.05	0.95	Uniform
θ_2	0.05	0.95	Uniform
λ_q	0	1	Uniform

Note: we allow for all calibrated parameters and measurement error components to vary simultaneously within a broad range. We run 500 iterations of the simulation, and for each iteration we separately apply: none of the above restrictions, each one individually (within-interval errors and dynamic age effect), and both together. For each iteration, parameters/errors are simulated within the chosen ranges using a scrambled Halton sequence, which assures broad coverage of the parameter space.

Source: authors' elaboration.

Table B5: Mean bias of alternative IGM estimators under different data transforms and simulated error structures, middle-income example

		Heritability ($\hat{\beta}$)			Out-performance ($\hat{\Delta}$)			Upwards mobility		
Scenario	‖ Transform →	Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
1	Classical ME	-7.1	-6.4	-9.0	1.1	0.6	2.2	-0.9	-0.7	1.3
		(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.2)	(0.1)	(0.2)
2	Correlated additive MEs	-4.5	-3.7	-6.1	0.7	-0.0	1.5	-1.0	-0.9	1.1
		(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
3	Multiplicative ME, parent	-6.4	0.3	-0.0	-0.6	-5.0	0.0	-3.9	-0.5	0.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)
4	Multiplicative ME, child	5.4	0.6	-0.0	3.5	5.6	0.0	3.3	3.3	0.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
5	Multiplicative ME, both	-1.6	0.0	-0.0	2.8	-0.0	0.0	0.0	0.0	0.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
6	Incomplete education	-12.9	-7.8	-6.6	-0.6	-1.3	1.6	-0.8	-0.8	2.4
		(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)
7	Truncation bias	-1.7	-0.6	-1.0	0.1	-0.0	0.0	-0.0	-0.0	0.0
		(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
8	Combined	-16.1	-10.3	-11.6	1.1	-0.9	2.7	-2.7	0.7	3.6
		(0.1)	(0.1)	(0.1)	(0.2)	(0.1)	(0.0)	(0.1)	(0.2)	(0.2)

Note: cells show the mean bias of each estimator, calculated as per equation (9a) for different estimators and transformations (as indicated in the columns); each row indicates a specific stylized set of assumptions (see Appendix B3), for which 50 separate simulations are run and the mean bias and its standard error (in parentheses) are reported.

Source: authors' estimates.

Table B6: Mean fractional bias of alternative IGM estimators under different data transforms and simulated error structures, middle-income example

Scenario ‖ Transform →		Heritability ($\hat{\beta}$)			Out-performance ($\hat{\Delta}$)			Upwards mobility		
		Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
1	Classical ME	-14.0	-13.4	-16.4	8.6	2.8	19.9	-1.3	-0.9	2.1
		(0.2)	(0.2)	(0.2)	(0.1)	(0.1)	(0.2)	(0.3)	(0.2)	(0.3)
2	Correlated additive MEs	-8.5	-7.6	-10.8	5.5	-0.1	14.0	-1.3	-1.3	1.7
		(0.2)	(0.2)	(0.1)	(0.1)	(0.1)	(0.2)	(0.2)	(0.2)	(0.2)
3	Multiplicative ME, parent	-12.5	0.6	-0.0	-5.5	-25.6	0.0	-5.4	-0.7	0.0
		(0.0)	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.1)	(0.1)	(0.0)
4	Multiplicative ME, child	9.4	1.1	-0.0	25.2	22.2	0.0	4.3	4.3	0.0
		(0.0)	(0.1)	(0.0)	(0.0)	(0.1)	(0.0)	(0.1)	(0.1)	(0.0)
5	Multiplicative ME, both	-2.9	0.0	-0.0	20.5	-0.0	0.0	0.0	0.0	0.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
6	Incomplete education	-26.5	-16.5	-11.7	-5.1	-5.9	15.0	-1.1	-1.1	3.9
		(0.2)	(0.1)	(0.1)	(0.1)	(0.1)	(0.2)	(0.0)	(0.0)	(0.1)
7	Truncation bias	-3.1	-1.2	-1.6	0.6	-0.0	0.3	-0.1	-0.1	0.0
		(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.3)	(0.1)	(0.1)	(0.1)
8	Combined	-33.2	-22.2	-21.4	7.7	-4.0	23.8	-3.7	0.9	5.8
		(0.3)	(0.3)	(0.3)	(1.8)	(0.2)	(0.4)	(0.2)	(0.2)	(0.3)

Note: cells show the mean fractional bias of each estimator, calculated as per equation (9b) for different estimators and transformations (as indicated in the columns); each row indicates a specific stylized set of assumptions (see Appendix B3), for which 50 separate simulations are run and the mean fractional bias and its standard error (in parentheses) are reported.

Source: authors' estimates.

Table B7: Mean within-estimator rank absolute fractional bias of alternative IGM estimators under different data transforms and simulated error structures, middle-income example

Scenario ‖ Transform →		Heritability ($\hat{\beta}$)			Out-performance ($\hat{\Delta}$)			Upwards mobility		
		Levels	Ref.	Rank	Levels	Ref.	Rank	Levels	Ref.	Rank
1	Classical ME	1.8	1.2	3.0	2.0	1.0	3.0	1.4	1.4	2.3
		(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
2	Correlated additive MEs	1.9	1.1	3.0	2.0	1.0	3.0	1.3	1.3	2.4
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
3	Multiplicative ME, parent	3.0	2.0	1.0	2.0	3.0	1.0	3.0	2.0	1.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
4	Multiplicative ME, child	3.0	2.0	1.0	3.0	2.0	1.0	2.0	2.0	1.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
5	Multiplicative ME, both	3.0	1.9	1.1	3.0	1.6	1.3	1.0	1.0	1.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)
6	Incomplete education	3.0	2.0	1.0	1.1	1.9	3.0	1.0	1.0	3.0
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
7	Truncation bias	3.0	1.2	1.9	1.8	1.5	2.7	1.3	1.3	2.4
		(0.0)	(0.1)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
8	Combined	3.0	1.8	1.2	2.0	1.0	3.0	2.0	1.2	2.8
		(0.0)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)

Note: cells show the rank absolute fractional bias across transformation, and its standard error (in parentheses) is reported. The closer the value to the unity, the smaller the absolute bias; each row indicates a specific stylized set of assumptions (see Appendix B3), for which 50 separate simulations are run and the absolute bias is calculated as the mean fractional bias absolute value.

Source: authors' estimates.

Table B8: Mean within-transform rank absolute fractional bias of alternative IGM estimators under different data transforms and simulated error structures, middle-income example

	Levels			Reference distribution			Rank-rank		
Scenario ‖ Transform →	$\hat{\beta}$	$\hat{\Delta}$	up	$\hat{\beta}^d$	$\hat{\Delta}^d$	up^d	$\hat{\beta}^r$	$\hat{\Delta}^r$	up^r
1 Classical ME	3.0	2.0	1.0	3.0	1.8	1.2	2.0	3.0	1.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)
2 Correlated additive MEs	3.0	2.0	1.0	3.0	1.3	1.7	2.0	3.0	1.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)
3 Multiplicative ME, parent	3.0	1.6	1.4	1.5	3.0	1.5	2.0	2.9	1.0
	(0.0)	(0.1)	(0.1)	(0.1)	(0.0)	(0.1)	(0.0)	(0.1)	(0.0)
4 Multiplicative ME, child	2.0	3.0	1.0	1.0	3.0	2.0	1.8	2.5	1.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.0)
5 Multiplicative ME, both	2.0	3.0	1.0	2.9	2.0	1.0	2.0	2.8	1.0
	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)
6 Incomplete education	3.0	2.0	1.0	3.0	2.0	1.0	2.0	3.0	1.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
7 Truncation bias	3.0	1.7	1.3	2.6	1.9	1.5	2.4	2.2	1.3
	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
8 Combined	3.0	2.0	1.0	3.0	1.9	1.1	2.1	2.9	1.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)

Note: cells show the rank absolute fractional bias across estimators, and its standard error (in parentheses) is reported. The closer the value to the unity, the smaller the absolute bias; each row indicates a specific stylized set of assumptions (see Appendix B3), for which 50 separate simulations are run and the absolute bias is calculated as the mean fractional bias absolute value.

Source: authors' estimates.

B2 Tables for Mozambique

Table B9: Child and parent median years of education by province

	Child's education			Parent's education		
Province	1997	2007	2017	1997	2007	2017
Niassa	2	3	5	2	3	6
Cabo Delgado	2	3	4	2	3	5
Nampula	2	3	5	3	4	5
Zambezia	2	3	4	2	3	5
Tete	2	3	5	2	3	5
Manica	2	4	5	2	4	6
Sofala	3	4	5	3	4	6
Inhambane	3	4	4	3	4	5
Gaza	3	4	4	3	4	4
Maputo Province	4	6	6	4	6	7
Maputo City	5	7	7	6	7	7

Source: authors' estimates.

Table B10: Relative and absolute mobility measures using a reference distribution transformation (with age restriction), at province level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Upwards mobility			Out-performance			Heritability		
Province	1997	2007	2017	1997	2007	2017	1997	2007	2017
Niassa	0.372	0.601	0.395	0.224	0.263	0.138	0.370	0.349	0.538
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.004)
Cabo Delgado	0.363	0.632	0.336	0.233	0.282	0.117	0.268	0.268	0.535
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Nampula	0.392	0.637	0.352	0.202	0.248	0.126	0.340	0.329	0.567
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Zambezia	0.420	0.682	0.351	0.252	0.292	0.152	0.342	0.295	0.548
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Tete	0.387	0.625	0.425	0.251	0.300	0.168	0.382	0.329	0.539
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Manica	0.462	0.810	0.527	0.255	0.364	0.215	0.402	0.287	0.463
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Sofala	0.429	0.766	0.523	0.210	0.316	0.209	0.474	0.352	0.506
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Inhambane	0.664	0.860	0.736	0.364	0.426	0.326	0.256	0.239	0.335
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Gaza	0.656	0.854	0.685	0.352	0.409	0.316	0.275	0.271	0.370
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Maputo Province	0.729	0.750	0.694	0.295	0.352	0.260	0.343	0.330	0.425
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)
Maputo City	0.548	0.806	0.780	0.273	0.344	0.281	0.343	0.325	0.361
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Mozambique	0.483	0.719	0.493	0.243	0.299	0.200	0.410	0.363	0.508
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)

Note: cells report mean of mobility estimators, calculated separately for each census year and province using a reference distribution. 'Mozambique' is calculated for the whole country, dynamic age effect restriction applied. See Table 8 for corresponding estimates with no dynamic age effect; standard errors in parentheses.

Source: authors' estimates.

Table B11: Intercept point at which average child's education is expected to out-perform their parents, using a reference distribution transformation (with age restriction), at province level

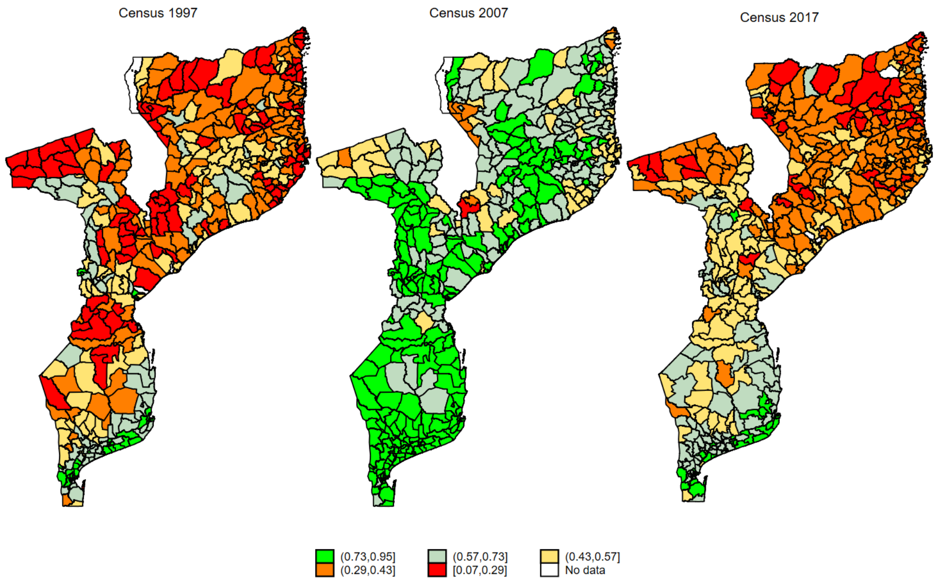
	Reference distribution			Levels		
Province	1997	2007	2017	1997	2007	2017
Niassa	0.606	0.655	0.548	3.313	5.701	6.312
	(0.002)	(0.001)	(0.002)	(0.021)	(0.021)	(0.028)
Cabo Delgado	0.568	0.636	0.502	2.706	5.223	5.399
	(0.001)	(0.001)	(0.002)	(0.013)	(0.015)	(0.025)
Nampula	0.556	0.620	0.541	3.182	5.456	6.022
	(0.001)	(0.001)	(0.001)	(0.010)	(0.010)	(0.018)
Zambezia	0.633	0.664	0.587	3.490	5.593	5.799
	(0.001)	(0.001)	(0.001)	(0.009)	(0.009)	(0.015)
Tete	0.657	0.697	0.615	3.731	5.861	6.736
	(0.001)	(0.001)	(0.002)	(0.019)	(0.016)	(0.022)
Manica	0.677	0.760	0.650	4.462	7.163	7.319
	(0.001)	(0.001)	(0.001)	(0.018)	(0.016)	(0.019)
Sofala	0.648	0.738	0.673	4.894	7.503	7.838
	(0.001)	(0.001)	(0.001)	(0.019)	(0.017)	(0.020)
Inhambane	0.740	0.810	0.741	5.046	7.596	7.576
	(0.001)	(0.001)	(0.001)	(0.014)	(0.018)	(0.019)
Gaza	0.736	0.811	0.751	4.981	7.622	7.529
	(0.001)	(0.001)	(0.001)	(0.013)	(0.018)	(0.022)
Maputo Province	0.700	0.775	0.702	6.078	8.835	8.990
	(0.001)	(0.001)	(0.001)	(0.015)	(0.017)	(0.016)
Maputo City	0.665	0.760	0.689	7.209	10.022	9.740
	(0.001)	(0.001)	(0.001)	(0.014)	(0.018)	(0.017)
Mozambique	0.662	0.720	0.656	4.693	7.058	7.458
	(0.000)	(0.000)	(0.000)	(0.005)	(0.005)	(0.006)

Note: cells report mean of mobility estimators, calculated separately for each census year and province using a reference distribution. 'Mozambique' is calculated for the whole country, dynamic age effect restriction applied; the intersection point indicates the intersection with the line of equality and, thus, the highest level of parental education below which offspring outcomes are expected to exceed those of their parents, on average. See Table 9 for corresponding estimates with no dynamic age effect; standard errors in parentheses.

Source: authors' estimates.

B3 Graphs

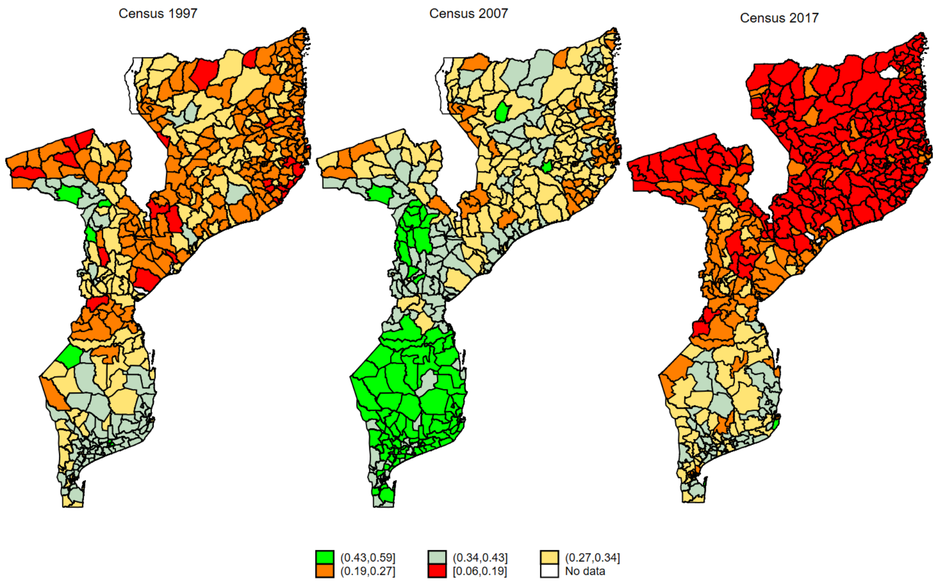
Figure B1: Upwards mobility using a reference distribution transformation (with age restriction)—Mozambique at *posto administrativo* level



Note: graphs report mean of upwards mobility estimator, calculated separately for each census year and posto administrativo using a reference distribution; dynamic age effect restriction applied. See Figure 5 for corresponding estimates with no dynamic age effect.

Source: authors' estimates

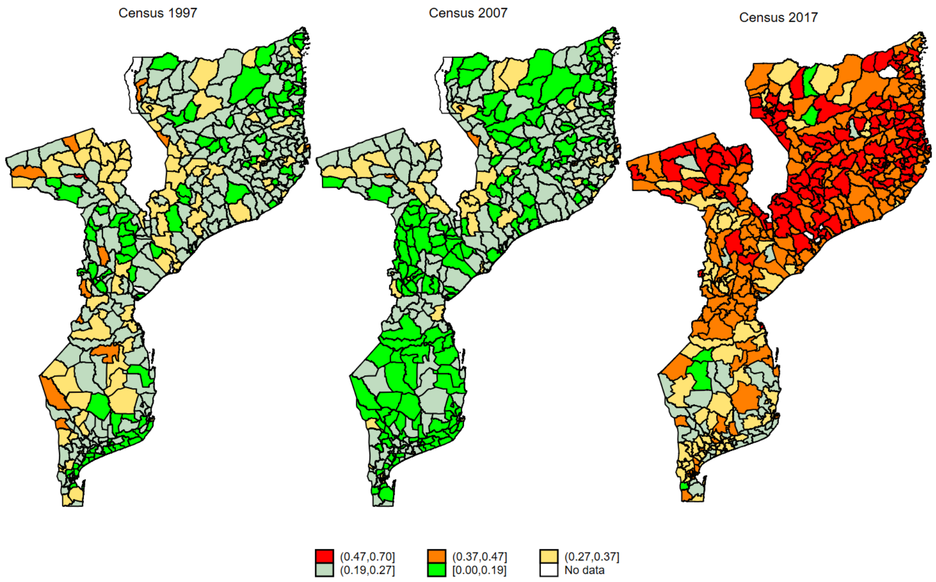
Figure B2: Out-performance using a reference distribution transformation (with age restriction)—Mozambique at *posto administrativo* level



Note: graphs report mean of out-performance mobility estimator, calculated separately for each census year and posto administrativo using a reference distribution; dynamic age effect restriction applied. See Figure 6 for corresponding estimates with no dynamic age effect.

Source: authors' estimates

Figure B3: Heritability using a reference distribution transformation (with age restriction)—Mozambique at *posto administrativo* level



Note: graphs report mean of the heritability estimator, calculated separately for each census year and posto administrativo using a reference distribution; dynamic age effect restriction applied. See Figure 7 for corresponding estimates with no dynamic age effect.

Source: authors' estimates.