HCEO WORKING PAPER SERIES

The Intergenerational Elasticity of What? The Case for Redefining the Workhorse Measure of Economic Mobility

Pablo Mitnik
David Grusky

Working Paper

2018-043

07/2018

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

THE INTERGENERATIONAL ELASTICITY OF WHAT? THE CASE FOR REDEFINING THE WORKHORSE MEASURE OF ECONOMIC MOBILITY

Pablo A. Mitnik
Stanford Center on Poverty and Inequality
(pmitnik@stanford.edu)

David B. Grusky
Stanford Center on Poverty and Inequality
(grusky@stanford.edu)

September, 2017

Abstract

The intergenerational elasticity (IGE) has been assumed to refer to the expectation of children's income when in fact it pertains to the geometric mean of children's income. We show that mobility analyses based on the conventional IGE have been widely misinterpreted, are subject to selection bias, and cannot disentangle the "channels" underlying intergenerational persistence. The solution to these problems—estimating the IGE of *expected* income or earnings—returns the field to what it has long meant to estimate. Under this approach, persistence is found to be substantially higher, thus raising the possibility that the field's stock results are misleading.

Acknowledgements. This research was supported by the U.S. Department of Health and Human Services, the Russell Sage Foundation, the Pew Charitable Trusts, and the Canadian Institute for Advanced Research. The authors gratefully acknowledge the helpful comments provided by Oscar Mitnik and Joao Santos Silva on a previous version of this paper. The Stanford Center on Poverty and Inequality is a program of the Institute for Research in the Social Sciences at Stanford University.

I. Introduction

The intergenerational elasticity (IGE) of earnings and income holds a central place in the study of intergenerational economic mobility. It has long been a workhorse measure used to characterize mobility at the national level (e.g., Mazumder 2005), to compare mobility across countries (e.g., Björklund and Jäntti 2011) and regions (e.g., Mayer and Lopoo 2008), to examine trends in mobility (e.g., Aaronson and Mazumder 2008), and to study differences in mobility by gender (e.g., Chadwick and Solon 2002) and race (e.g., Hertz 2005). The well-known "Great Gatsby Curve," which shows a negative cross-sectional relationship between income inequality and economic mobility (e.g., Corak 2013), rests famously on the IGE. The most cited trend analyses in the U.S. have, at least until recently, been based on the IGE (e.g., Lee and Solon 2009). In their very influential study, Chetty et al. (2014) turned to a different mobility measure (i.e., the rank-rank slope), but even so they felt obliged to lead off their research with analyses of the IGE. There is simply no other measure of economic mobility that comes close to the IGE in popularity.

When a measure has become the "industry standard," we often come to rely on stock interpretations of it and may mistakenly take for granted that these interpretations, long assumed to be correct, have a solid rationale for them. This is the fate of the conventionally-estimated IGE. Although it is attractive in conception, it has been specified in a way that is inconsistent with this conception and all associated interpretations, including the archetypal interpretation as a measure of regression to the arithmetic mean. The key problem is simple: The IGE is typically assumed to refer to the expectation of the children's earnings or income conditional on their parents' income when in fact it pertains to the conditional geometric mean of the children's earnings or income. We will show that, because this problem is so fundamental, there is no alternative but to replace the conventionally-estimated IGE with one that is consistent with the intended definition. This new IGE, when estimated in three high-quality datasets, provides estimates of intergenerational persistence in the U.S. substantially larger than those obtained with the conventional IGE.

The IGE is not just widely misinterpreted but also inadequate to the tasks to which it has been applied. If the children's long-run income or earnings measure includes zero in its support (which is very likely), then the conventional estimand is undefined. Even if the long-run measure of income or earnings is positive rather than nonnegative, the same is not the case for the short-

run measures (e.g., sons' earnings in a particular year) to which the mobility field routinely resorts as proxies for their unavailable long-run counterparts (e.g., sons' average lifetime earnings). These short-run measures typically have a substantial probability mass at zero.

Confronted with the resulting "zeros problem," and unaware of its roots in the adoption of a population regression function (PRF) that is inconsistent with their interpretation of the IGE, mobility scholars have responded in one of two ways, neither of which is attractive. In some recent analyses, mobility scholars have abandoned the IGE altogether in favor of the rank-rank slope or other measures, an approach that has yielded important results but of course only makes sense when the new measures answer the key questions in play. In most analyses, the problem has simply been ignored and the conventional IGE estimated via the expedient of dropping cases without earnings or income, thereby solving what is perceived as the purely practical issue of the logarithm of zero being undefined. We show that serious selection biases are generated by this practice and that all available approaches to addressing it are unattractive due to a combination of methodological and pragmatic reasons.

There are also important research situations in which it is not sensible, even as an approximation, to posit that the children's long-run measure is positive rather than nonnegative. This "approximation defense" does not work, most obviously, when the focus shifts from sons to daughters, as there are typically many daughters with zero long-run earnings. The IGE pertaining to women with positive earnings is well defined, but this IGE is not typically the one of conceptual interest. If one nonetheless estimates it when conducting cross-country comparisons or studying historical trends, misleading conclusions can easily be drawn.

The "no-zero assumption" is also indefensible when the goal is to study the role of marriage in the intergenerational transmission of family economic status. This is because the zeros problem makes it impossible to assess the contribution of differential marriage chances (across parental-income levels) to observed family-income IGEs. There have of course been ingenious attempts to circumvent the zeros problem in assessing the role that assortative mating, conditional on marriage, plays (e.g., Chadwick and Solon 2002). We will show that these attempts fall short.

The conceptual and methodological problems we have outlined all arise directly from the field's unwitting use of a PRF that references the conditional geometric mean. The solution to these problems involves (a) replacing the standard PRF with one referencing the conditional

expectation of the child's earnings or income, and (b) using well-understood estimators to estimate it. This is a simple solution that obviates the need to resort to other measures, like the rank-rank slope, as a second-best substitute.[1] As we explain in detail below, the approach we are advocating is very advantageous when estimation is based on tax and other administrative data, an especially convenient result given that the use of these data has grown markedly in recent times and can be expected to expand further in coming decades (e.g., Einav and Levin 2014).

This paper may be understood, then, as a simple call to replace the conventionally estimated IGE (i.e., the IGE of the geometric mean) with the IGE that mobility scholars long assumed that they were estimating (i.e., the IGE of the expectation). It is merely a matter, in other words, of adjusting practice to bring it into correspondence with the field's prevailing interpretation. Although one might instead adjust the interpretation, we will show why this is not a desirable course of action.

This call is issued in the context of a very large literature on economic mobility (see reviews by Black and Deveroux 2011; Corak 2006; Solon 1999), but it is most closely related to research of Mitnik et al. (2015; 2017) and Mitnik (2017a; 2017b; 2017c). In Mitnik et al. (2015), tax and other administrative data are used to estimate the IGE of the expectation, with the objective of providing the best possible estimates of intergenerational elasticities in the United States today. The goals of Mitnik (2017a; 2017b; 2017c), by contrast, are to develop a two-sample estimator of the IGE of the expectation and to advance generalized error-in-variables models for the estimation of the IGE of the expectation with various estimators. These papers assume the virtues of estimating the IGE of the expectation, whereas our precursory purpose here is to establish why one should want to estimate that IGE. While Mitnik et al. (2017) have provided evidence of selection bias in the estimation of the conventional earnings IGE, that paper exposes only one problem with the current methodological state of affairs (i.e., the selection problem) and does not focus, as we do here, on the need for an alternative estimand to successfully address that problem and many others.[2]

The discussion provided here is also related to previous arguments regarding the estimation of constant elasticities in production functions (Goldberger 1968a) and international trade (Santos Silva and Tenreyro 2006). The crucial point that the de facto estimand of "log-log regressions" is in the general case the elasticity of the geometric mean was not made in those contributions. Although this point was made by Petersen (2017), in a paper written concurrently

with ours, here we focus on the costs of estimating the IGE of the geometric mean within the field of mobility research.[3] These field-specific costs, which have not been appreciated in prior mobility research, will be shown to be substantial. We will show that (a) the field's main interpretations of the conventional IGE (e.g., regression to the mean) are invalid, (b) the standard justification for using short-run proxy variables to estimate the conventional IGE is not defensible, (c) a meaningful earnings IGE for women cannot be estimated (using conventional approaches), and (d) the role of marriage dynamics in the intergenerational transmission of advantage cannot be examined with the conventional IGE. We will lay out the conceptual and methodological advantages of switching to the IGE of the expectation in light of these field-specific problems.

We begin our analysis by showing that the conventional IGE has been widely misinterpreted. Next, we show that the conventional IGE is not just routinely misinterpreted but also leads to serious methodological problems, including (a) very consequential selection biases, and (b) the inability to examine the roles of gender and marriage in the intergenerational transmission of economic advantage. The following section introduces the IGE of the expectation and shows that it's consistent with the interpretations that were (wrongly) thought to apply to the conventional IGE. We then introduce a suite of estimators for the IGE of the expectation and show that they allow mobility scholars to estimate this elasticity in all contexts in which the conventional IGE has been estimated. We conclude by carrying out three illustrative analyses that show that intergenerational persistence is substantially higher when examined with our new estimand than with the conventional estimand.

II. The conceptual case against the conventional IGE

We begin, then, by discussing how the IGE has conventionally been interpreted and showing that these interpretations are incorrect. The following quotations, which offer various overlapping and non-exclusive interpretations by prominent scholars of economic mobility, well represent the range of ways in which the IGE has been discussed.

> [The IGE] measures the percentage differential in the son's expected income with respect to a marginal percentage differential in the income of the father (Björklund and Jäntti 2011:497).

> [The intergenerational elasticity of earnings] measures the percentage difference in expected child earnings that is associated with a one percent difference in parental earnings (Hertz 2006:2).

> [The IGE] represents the fraction of economic advantage that is on average transmitted across the generations…. When β [i.e., the IGE] is greater than zero but less than one there is some generational mobility of income, so that parents with incomes above (or below) the average will have children who grow up to have incomes above (or below) the average. However, the deviation from the average will not be as great in the children's generation.... Expected mobility is greater the lower the value of β, that is the more rapid regression to the mean. In the extreme with $\beta = 0$ there is no relationship between parent and child outcomes and the expected outcome of a child is just the average income for all children regardless of parental income (Corak 2006:4-5).

> [The IGE provides] a parametric answer to questions like, if the parents' long-run earnings are 50% above the average in their generation, what percent above the average should we predict the child's long-run earnings to be in her or his generation? (Solon 1999:1777).

> Now consider a family of four with two children whose income is right at the poverty threshold, roughly 75% below the national average. If the IGE is 0.6, then on average, it will take the descendants of the family 5 to 6 generations (125 to 150 years) before their income would be within 5% of the national average (Mazumder 2005:235; one note omitted).

> [The IGE is] the fraction of differences among parents that is typically observed among their adult children. I measure intergenerational persistence in logarithmic or percentage terms. Consider two sets of parents whose incomes differ by 50 percent. If, on average, the incomes of their children differ by 20 percent, then the intergenerational persistence of income is said to be 0.4 or 40 percent. … I refer to 'regression to the mean' in logarithmic or percentage terms. Because of economic growth, regression to the mean in percentage terms is more interesting than, say, regression to the mean in absolute terms. Economic growth alone tends to multiply everyone's income, which produces regression away from the mean in absolute terms but not in percentage terms" (Mulligan 1997:24-25: author's italics omitted).

> [D]epending upon the degree of inequality in parental incomes, even small values of β can confer substantial advantages to the children of the well off. For example, … in 1999, [U.S.] households with children under the age of 18 at the top income quintile had 12 times as much money ... as those at the bottom quintile. The generational income elasticity directly translates this ratio into the economic advantage a child from the higher-income family can expect to have in the next generation over one from the lower-income family.... With a generational elasticity as high as 0.6, children born to the higher-income parents will earn, when no other influences are at work …, almost four and a half times as much as children born to lower-income parents (Corak 2004:11-12).

These quotations make it clear that the IGE is either directly or indirectly represented as the elasticity of the expectation of the children's income or earnings with respect to their fathers' earnings or their parental income. The foregoing quotations also reveal other common (and closely related) interpretations that have been offered up. The IGE has been used to comment upon (a) the extent to which there is regression to the arithmetic mean, in percent terms, from one generation to the next, (b) the number of generations needed to regress to the arithmetic mean, (c) the share of inequality between parents that is transmitted to their children, and (d) the

economic advantage that a child from a better-off family may expect (relative to a child from a worse-off family). We examine each of these interpretations in turn.

The elasticity of the expectation. The IGE, as conventionally specified, has been interpreted as the elasticity of the expectation of the children's income or earnings. It is straightforward to show that this interpretation is incorrect. The standard PRF posited in the literature assumes that the elasticity is constant across levels of parental income:

$$E(\ln Y\,|x) = \beta_0 + \beta_1 \ln x, \qquad [1]$$

where Y is the son's or daughter's long-run income or earnings, X is long-run parental income or father's earnings, β_1 is the IGE as specified in the literature, and expressions like "Z|w" are used as a shorthand for "$Z|W = w$." The key point here is that β_1 is not, in the general case, the elasticity of the conditional expectation of the child's income. This would hold as a general result only if $E(\ln Y|x) = \ln E(Y|x)$. But of course the latter is not the case (due to Jensen's inequality). Instead, as $E(\ln Y\,|x) = \ln \exp E(\ln Y\,|x)$, and $GM(Y|x) = \exp E(\ln Y\,|x)$, Equation [1] is equivalent to

$$\ln GM(Y|x) = \beta_0 + \beta_1 \ln x, \qquad [2]$$

where GM denotes the geometric mean operator. Therefore, β_1 is the elasticity of the conditional geometric mean, meaning that it's the percentage differential in the geometric mean of children's long-run income with respect to a marginal percentage differential in parental long-run income.[4] We cannot rule out that this elasticity, were it estimated robustly and without bias (a point to which we will return), might be of interest under some circumstances. But a case for estimating it has not, to our knowledge, been made. We know of no statement to the effect that the conventional approach to estimating the IGE has been undertaken because of some genuine interest in recovering the elasticity of the conditional geometric mean of long-run income or earnings.

Regression to the arithmetic mean. We next consider the frequently-made claim that the IGE reveals the extent to which children's income regresses to the arithmetic mean (in percent terms). This claim can be evaluated by taking the expectation in Equation [1] with respect to the population distribution of X and then using the resulting equality and Equation [1] to yield:

$$E(\ln Y|x) - E(\ln Y) = \ln GM(Y|x) - \ln GM(Y) = \beta_1(\ln x - \ln GM\,(X)). \qquad [3]$$

It follows that β_1, which is assumed to lie in the open unit interval, measures the degree of regression to the *geometric mean* (in percent terms). It measures, more precisely, the degree to

which the conditional geometric mean regresses to the economy-wide geometric mean. According to Equation [3], the conditional expectation of the *logarithm* of children's income does regress to the economy-wide mean, but this cannot be the implied interpretation. If indeed the transformed income variable was presumed to be regressing to the mean, then the usual interpretation in terms of *percent* differences would not be offered (see the quotations above). Is the regression to the geometric mean of interest? Although we cannot rule out the possibility that it is (again assuming that it's estimated robustly and without bias), we are not aware of any statements in the literature suggesting an intent to estimate an IGE allowing for this interpretation.

The number of generations needed to regress to the arithmetic mean. It is also commonly claimed that β_1 can be used to assess how many generations it will take the average descendant of a family at a certain income level (e.g., the poverty level) to come to within *K* percent of a country's average income (with *K* being a small number). This is of course an alternative way of assessing the speed with which incomes regress to the mean. But β_1 does not allow for this alternative assessment either. Under the strong assumption of a first-order Markov process (see Eide and Showalter 2000; Stuhler 2012), β_1 can only tell us how many generations it will take for the geometric mean of the descendants' income to be close to the overall geometric mean of income in their generation. By using Equations [1] and [3] (the latter adapted to the multigenerational context), the Markov process assumption, and the assumption that the IGE does not change across generations, we obtain the following expression (see Appendix A):

$$n = \ln\left|\frac{k}{\ln y_0 - \ln GM(Y_0)}\right| (\ln \beta_1)^{-1}\,, \qquad [4]$$

where the subscript 0 denotes the current generation, *n* refers to the nth generation after the current one, and $k = K/100 \ll |\ln y_0 - \ln GM(Y_0)|$ is the threshold value used to stipulate that full regression to the geometric mean has occurred (i.e., full regression is stipulated to occur when $|\ln GM(Y_n|y_0) - \ln GM(Y_n)| \leq k$).

The share of inequality that is transmitted to the next generation. It is also common to claim that the IGE pertains to the "share" of percent differences in income that is transmitted across generations. This claim can be interpreted in two different ways because of an ambiguity in the specification of the share's numerator. It is clear that the denominator of the share in question is a percent difference in parental income. It is unclear, however, whether the numerator

is (a) the expected percent difference in the children's incomes, or (b) the percent difference in the children's expected incomes (see Mulligan 1997:24; see also, e.g., Corak 2004:10-11; Hertz 2006:3).[5] We consider each possibility in turn. Under the first interpretation of the numerator, the share interpretation would draw on Equation [1] to conclude that $E(\ln Y|x_2 - \ln Y|x_1) = \beta_1(\ln x_2 - \ln x_1)$. If, for any values of $Y|x_1$ and $Y|x_2$, $\ln x_2 - \ln x_1$ provides a good approximation to $(x_2 - x_1)(x_1)^{-1}$ (i.e., the "parental approximation") and $\ln Y|x_2 - \ln Y|x_1$ provides a good approximation to $(Y|x_2 - Y|x_1)(Y|x_1)^{-1}$ (i.e., the "child approximation"), it would then be the case that:

$$\beta_1 \cong E\left(\frac{Y|x_2 - Y|x_1}{Y|x_1}\right)\left(\frac{x_2 - x_1}{x_1}\right)^{-1}.$$

This is precisely as assumed under the first interpretation. But this conclusion is only as good as the approximations that underlie it. Although the parental approximation is valid when the parents in question have incomes that do not differ much in percent terms, most of the approximations for the children will not be valid given the wide range of actual incomes. Indeed, because the expected percent difference is a function of the percent differences between all values in the support of two full income distributions, we can expect $(Y|x_2 - Y|x_1)(Y|x_1)^{-1}$ to be, for most values of $Y|x_1$ and $Y|x_2$, much larger (often orders of magnitude larger) than the largest percent difference for which a difference in logarithms provides a minimally reasonable approximation.

Under the second understanding of the numerator, the share interpretation is also incorrect. This interpretation posits that:

$$\beta_1 = \frac{\ln E(Y|x_1) - \ln E(Y|x_2)}{\ln x_1 - \ln x_2}.$$

But from Equation [2] it follows that:

$$\beta_1 = \frac{\ln GM(Y|x_1) - \ln GM(Y|x_2)}{\ln x_1 - \ln x_2}. \qquad [5]$$

We can conclude that, insofar as one wants to advance a share interpretation of the IGE, the numerator for the share has to pertain to the percent difference between the geometric means of the children's conditional incomes.[6] We are unaware of any argument indicating that Equation [5] can be used to interpret β_1 as "the fraction of [percent] differences among parents that is typically observed among their adult children" (Mulligan 1997:24).

Economic advantage in terms of expectations. The IGE has also been used to translate a ratio between parental incomes into "the economic advantage a child from the higher-income family can expect to have in the next generation over one from the lower-income family" (Corak 2004:11-12; full quotation reproduced above). This usage is based on the claim that one can derive the following expression from Equation [1]:

$$\frac{E(Y|x_2)}{E(Y|x_1)} = \left(\frac{x_2}{x_1}\right)^{\beta_1}.$$

The foregoing implies that the ratio between parental incomes, when raised to the power β_1, equals the corresponding ratio of the children's expected incomes.[7] This, however, is incorrect. What can be derived from Equation [2] is the following:

$$GM\left(\frac{Y|x_2}{Y|x_1}\right) = \frac{GM(Y|x_2)}{GM(Y|x_1)} = \left(\frac{x_2}{x_1}\right)^{\beta_1}. \qquad [6]$$

This equation means that the IGE maps percent income differences between parents into the geometric mean of percent income differences between their children.

We have established, then, that the conventional IGE is not what it's long been made out to be. It's not the elasticity of the conditional expectation of the child's income; it doesn't measure the degree of regression to the arithmetic mean that occurs in one generation; it doesn't allow one to calculate the number of generations needed to regress to the arithmetic mean; it doesn't speak to the share of inequality that's transmitted to the next generation; and it can't be interpreted as the economic advantage that a child from a better-off family may expect to have in adulthood.

III. The "good summary measure" defense and sample selection bias

The foregoing makes it clear that there's no conceptual rationale for estimating the conventional IGE. Might we instead defend it as a "good summary measure" of economic persistence? This defense implies that mobility scholars need only change the way in which they interpret IGE estimates by everywhere substituting "geometric mean" for "mean" or "expectation." If our interpretations are brought into line in this way, we might then continue to use the conventional IGE and continue to build on its long history of use. The purpose of this section is to show that, although continuity with past practice of course has some value, the conventional IGE is such a flawed summary measure that the benefits of continuity are overwhelmed by the costs.

The root of the conventional IGE's problems is that, unlike other commonly used measures of central tendency (e.g., expectation, median, mode), the geometric mean is undefined, or necessarily zero, when the random variable includes zero in its support.[8] Unfortunately, zero most likely *is* in the support of all measures of long-run economic status, given that some children face severe lifetime disability, mental illness, or imprisonment. These children may live with their parents or in group quarters for their entire lives and have no long-run earnings or income. How problematic are these zeros? Because there are relatively few such children, it might be thought that the problem is minor and ignorable. We discuss in this section the methodological problems that arise even in this context that might, at first blush, seem unproblematic. We then turn to the problems that emerge in contexts where the no-zero assumption is untenable even as an approximation.

The main reason why methodological problems emerge even when there are relatively few long-run zeros is that long-run measures are almost never available. As a result, Equation [1] is not estimated directly, and instead short-run proxy variables are substituted for *X* and *Y*. This results in measurement error in both of the long-run variables. The problems that arise in estimating the conventional IGE when there is left-side measurement error (i.e., measurement error in the children's income or earnings variable) can be usefully examined in the context of the generalized error-in-variables (GEiV) model advanced by Haider and Solon (2006).[9] This model supplies the standard justification for estimating the conventional IGE by Ordinary Least Squares (OLS) using proxy variables that satisfy some conditions (Nybon and Stuhler 2016).

The GEiV model lets $\ln Z_t = \lambda_{0t} + \lambda_{1t} \ln Y + V_t$ be a linear projection of $\ln Z_t$ on $\ln Y$, where $Z_t > 0$ is children's income or earnings at age *t*, $Y > 0$, and λ_{1t} captures left-side lifecycle bias and varies with *t*. If $\lambda_{1t} = 1$, the measurement error is $\lambda_{0t} + V_t$. The (left-side) empirical assumption of the GEiV model is that $Cov(V_t, \ln X) = 0$. It follows that, if we estimate a PRF in which $\ln Z$ is substituted for $\ln Y$ (in Equation [1]), the probability limit of the slope coefficient is $\lambda_1 \beta_1$.[10] This leads to the conclusion that left-side measurement error is unproblematic as long as the measurement age pertains to the "correct years" of the child's lifecycle (i.e., years when $\lambda_1 \cong 1$). The available evidence indicates that it's best to estimate the conventional IGE with the children's measurements pertaining to approximately age 40 (Haider and Solon 2006; Böhlmark and Lindquist 2006; Mazumder 2001; Nybon and Stuhler 2016).

This model, which has been the foundation of IGE-based mobility research since the

publication of Haider and Solon (2006), does not take into account that selection bias may arise when short-run measures of children's economic status are used. The Haider-Solon approach is appropriate when (a) the short-run measures do not include zero in their support (and hence there's no selection), or (b) the short-run measures do include zero in their support but dropping children with zero income or earnings from the analysis is an ignorable selection mechanism. We consider each of these two possibilities in turn.

A. Do short-run measures include zero in their support?

The assumption that zero is not in the support of the short-run children's measures typically employed for estimation *may* be defensible in analyses of family income within a few countries with very highly developed welfare states. It is not, however, a good assumption for analyses of family income in the vast bulk of countries with less generous welfare regimes. For instance, Chetty et al. (2014: Online Appendix Table IV) report that 5.4 to 8.0 percent of 29 to 32 year-olds in the U.S. have zero family income (depending on the data set), while 9.2 to 12.6 percent have zero earned family income (again depending on the data set).[11]

When we turn to analyses of individual earnings (rather than family income), the assumption becomes even more problematic because the prevalence of zeros is no longer reduced by (a) the pooling of income across family members, and (b) the inclusion of unearned income. The zeros problem looms especially large in countries with high rates of unemployment, a large informal sector (insofar as unreported income is treated as "zero formal income"), or low rates of labor force participation. This is a troubling conclusion because the IGE of men's earnings is, by far, the most-frequently estimated measure of economic mobility.

The reason why short-run income and earnings measures with a substantial share of zeros are a concern is, of course, that the geometric mean is undefined in the presence of zeros. As a result, the "analogy principle"—which presumes the existence of a sample statistic having the same property in the sample that the parameter has in the population (Goldberger 1968b; Manski 1988)—becomes inapplicable for the estimation of the PRF of Equation [1]. This is the case even if the children's long-run income or earnings are strictly positive variables.

It has not been appreciated, however, that using proxy variables with a substantial share of zeros to estimate the PRF of Equation [1] is problematic because it makes analog estimation unfeasible. Rather, researchers have understood the zeros problem as the wholly practical one that the logarithm of zero is undefined, and they have proceeded by simply dropping from their

samples children without income or earnings (and those with low income or earnings are also typically dropped). We are not aware of any attempt to justify this practice.

B. Is Dropping Children with Zero Income or Earnings an Ignorable Selection Mechanism?

We next examine whether dropping children with zero income or earnings is an ignorable selection mechanism. We assume that the short-run (typically annual) child's measure of income or earnings pertains to an age in which $\lambda_1 = 1$, as this is the case of most interest. The GEiV model assumes $Z > 0$ when estimation is carried out with the short-run measure Z instead of the long-run measure Y. What happens, then, when $Z \geq 0$ (instead of $Z > 0$) and estimation is carried out by dropping observations whenever $Z = 0$? The estimated PRF is $\mathrm{E}(\ln Z \,|x.\, Z > 0) = \gamma_0 + \gamma_1 \ln x$. Given that the GEiV model effectively assumes that $Cov(\ln X\,, V|Z > 0) = 0$, the slope of that PRF is:

$$\gamma_1 = \frac{Cov(\ln Z\,, \ln X\,|Z > 0)}{Var(\ln X\,|Z > 0)} = \frac{Cov(\ln Y\,, \ln X\,|Z > 0)}{Var(\ln X\,|Z > 0)}. \qquad [7]$$

If Equation [1] is the PRF for the full population, then the long-run PRF when the population excludes those for whom $Z = 0$ is:

$$\mathrm{E}(\ln Y\,|x.\, Z > 0) = \beta_0 + \beta_1 \ln x + E(\Upsilon|x.\, Z > 0), \qquad [8]$$

where $\Upsilon = \ln Y - \beta_0 - \beta_1 \ln X$ (see, e.g., Heckman 2008). It then follows from Equations [7] and [8], and the standard analysis of omitted-variable bias, that γ_1 is equal to β_1 only if $Cov\big(\ln X\,, E(\Upsilon|X.\, Z > 0)\big) = 0$.

This condition is almost certainly violated because (a) the probability of having zero earnings or income any year is negatively correlated with parental income (Chetty et al. 2014: Fig. 1; Drewianka and Mercan, n.d.:19-20; Gregg et al. 2016:14-15; Mitnik et al. 2015:49, 63), and (b) those who have zero earnings or income in the measurement year tend to have lower long-run income or earnings than others with the same parental income. The latter conclusion is based in part on the wholly mechanical relationship between earnings or income at any age and the value of the corresponding long-run measures, but it is also very relevant that being unable to work or to find work is correlated across years (e.g., Kosanovich and Theodossiou Sherman 2015), that unemployment and lower wages often result from reduced human or social capital or some other common cause (e.g., Bernhardt et al. 2001), and that unemployment has a negative causal effect on later wages (e.g., Gangl 2006). It is further likely that children who are single at midlife (and therefore more likely to have zero family income) are also more frequently single

other years and therefore have lower long-run family income. It follows that using annual measures, after dropping those with no (or low) earnings or income, should lead to larger overestimates of expected log earnings or income at lower levels of parental income. We should therefore expect a negative selection bias (i.e., $\gamma_1 < \beta_1$).

Is the bias large enough to be worrying? Because long-run earnings and income measures are nearly always unavailable, it is quite difficult to estimate the magnitude of this bias. Nevertheless, the available evidence suggests that it is substantial, especially for the IGE of men's earnings. Using earnings reported in W2 forms, Mitnik et al. (2015:46-51) provided a range of IGE estimates for men's earnings, under the assumption that the mean log earnings of those without reported earnings is low and does not vary with parental income. The downward bias was nearly 30 percent even under their lower-bound estimate. Relying on the same data, but using copula-based selection models (Smith 2003) with very flexible income distributions, Mitnik et al. (2017) find an even larger downward bias. Similarly, Drewianka and Mercan (n.d.) find that estimates of the IGE of men's earnings increase very substantially when the long-term unemployed are included in the analysis, while Gregg et al. (2016) find evidence of negative selection bias in analyses in which spells out of work are ignored in computing earnings IGEs.

C. Addressing the selection problem

The foregoing indicates, then, that there is a serious selection problem and that the standard practice of simply ignoring it is ill-advised. Can the problem be adequately addressed through statistical modelling or some type of "data fix?" We conclude this section by briefly considering approaches based on (a) imputing nonzero values, (b) estimating sample selection models, or (c) using multiyear averages as the children's short-run measure. As we will show, these three approaches are all unattractive, meaning that the conventional IGE is not easily salvaged (see Appendix B for more details on each of these points.)

The first approach involves retaining in the sample those children with zero income or earnings, either by substituting all zeros with a small positive value (e.g., Couch and Lillard 1998), or by adding a small positive amount to the earnings or income of all children (e.g., Drewianka and Mercan n.d.). This approach is unattractive because even small changes in imputed values lead to disturbingly large changes in estimates (Mitnik et al. 2015:46-51; see also Chetty et al. 2014:Table 1; Minicozzi 2003:ftn. 27). The estimates are sensitive because the IGE of the geometric mean is dominated by small absolute differences at the lowest quantiles of the

child's conditional distributions (see Appendix B).

The second possibility is to instead estimate the conventional IGE with the help of a sample selection model. Consider the following PRF:

$$E(\ln Z \,|x.Z > 0) = \eta_0 + \eta_1 \ln x + E(\zeta|x.Z > 0), \qquad [9]$$

where $\zeta = \ln Z - \eta_0 - \eta_1 \ln X$, and $E(\zeta|\mathrm{x}) = 0$. If $\lambda_1 = 1$, replacing $\ln Z$ in Equation [9] yields

$$E(\ln Y \,|x.Z > 0) = (\eta_0 - \lambda_0) + \eta_1 \ln x + E(\zeta - \mathrm{V}|x.Z > 0). \qquad [10]$$

It follows that $\eta_1 = \beta_1$ (see Equation [8]).

Although estimating Equation [10] might then seem a promising alternative, the main problem with doing so is that it is extremely difficult to find a plausible exclusion restriction with the data regularly available to mobility scholars. This makes semi-parametric estimation of Equation [10] unfeasible (e.g., Vella 1998; Manski 1989; Moffit 1999; Lee 2003) and leaves us with parametric estimation based either on standard bivariate distributions (e.g., the Heckman selection model) or on copula-based selection models. The well-known worry with the first approach is its high sensitivity to the violation of distributional assumptions (e.g., Manski 1989; Moffit 1999). At the same time, reducing the dependence on distributional assumptions by using copula-based selection models rests on a complex and labor-intensive estimation approach (see Mitnik et al. 2017), which still is not robust enough to provide a viable foundation for general mobility analysis.

The third approach to salvaging the conventional IGE entails using multiyear averages as the children's short-run measure. This will of course reduce the number of zero values. We might, for example, replace the typical single-year measure of the offspring's earnings with one that averages over as many years as necessary to make the zeros problem inconsequential. Although attractive in principle, this approach is unattractive it practice because it (a) ramps up the data demands so substantially that the IGE could no longer be estimated in many countries (in particular those countries in which only cross-sectional data are available and estimation thus relies on a two-sample estimator), and (b) imposes a much longer "waiting time" before the IGE for any given cohort can be estimated. It follows that the third approach would greatly hinder the comparative study of mobility across countries and times.

The upshot is that there is no attractive work-around to the problem of selection bias when the IGE of the geometric mean is estimated. The available approaches to addressing this bias would either lead to nonrobust estimates or would limit IGE analyses to those few countries

and periods that have the data required by the averaging strategy. Neither of these choices is attractive.

IV. Women, marriage, and the impossible task of estimating a nonexistent estimand

In the previous section, we focused on long-run income measures for which the no-zero assumption, although false, may still be a reasonable approximation. When we turn to the case of women's earnings, the no-zero assumption is much more problematic because, for a great many women, lifetime earnings *are* zero. As a result, it's not just a "data problem" that's in play, as is the case for men's earnings. The no-zero assumption is, to the contrary, simply inconsistent with the gender-based division of labor that we find in all national and historical contexts for which IGEs have been estimated.

There is, then, an irresolvable conceptual mismatch between an outcome variable—women's long-run earnings—that has a very substantial probability mass at zero and the implicit use of a measure of central tendency that is undefined when this is the case. The no-zero assumption is particularly detrimental for trend analyses of women's earnings. Because of the dramatic changes in women's labor force participation, any attempt to estimate trends in the IGE after excluding women with zero earnings will be particularly compromised. The resulting estimates will reflect not just changes in the intergenerational transmission of economic advantage in the labor market but also changes in the processes by which women have been selected into the labor force (see Nicoletti and Ermisch 2007:6). A similar argument applies to cross-country comparisons.[12]

It is perhaps no surprise in this context that relatively few scholars have provided estimates of intergenerational economic persistence in women's own earnings. Although this state of affairs may partly be the result of "unconscious sexism" (Chadwick and Solon 2002: 335), it surely doesn't help matters that the conventional IGE is very ill-suited to that task.

The case for replacing the conventional IGE might therefore seem particularly strong when the focus turns to women's own earnings. There is one remaining complication, however, that must be addressed before we can reach that conclusion. We are referring here to the claim that, because women often rely on the income of their partners, their own earnings may be "an unreliable indicator of their economic status" (Chadwick and Solon 2002:335). By this logic, the failure to attend to women's own earnings stems from the recognition that the conventional IGE of women's earnings, even if it were well-defined, wouldn't be a very useful measure. This line

of reasoning further suggests that scholars should instead estimate the IGE of family income (e.g., Hertz 2007; Lee and Solon 2009; Mayer and Loopo 2004). Given that prime-age employment among men is now declining in some countries (especially the U.S.), it would also follow, as a corollary to this argument, that analyses of men's own earnings are likewise of diminishing utility.

We find this argument unpersuasive given the importance of understanding how intergenerational processes play out through different *channels* to produce different levels of family-income persistence. If we cannot identify the channels underlying economic persistence, it is difficult to identify its causes and possible policy responses. Under a simple "accounting perspective," we can distinguish four channels through which a higher-income upbringing can benefit a child: (a) parents can directly pass on economic assets to their child (via inheritance and inter vivos gifts); (b) parents can invest in or otherwise affect their child's human, social, or cultural capital in ways that boost the child's own earnings (e.g., buying childcare, socialization, genetic transmission); (b) parents can affect their child's chances of marrying and staying married through direct investments, socialization, and other indirect processes (thereby opening up the opportunity of securing income indirectly through a spouse); and (d) parents can likewise increase their child's chances of marrying a higher-income spouse (conditional on being married). We cannot understand the transmission process and develop sensibly targeted policy without understanding the relative importance of these four channels. This in turn requires a modelling framework that isn't undone by the zeros that loom so large for many of these channels.

The channels pertaining to income secured via marriage have been understudied because they cannot be properly conceptualized (even at the level of long-run measures) without running squarely into the zeros problem. The IGE of long-run earnings from spouses (with respect to the children's parental income) is undefined, given that there are many children who never marry and thus will, by definition, have zero earnings from their (nonexistent) spouses. Even if the analysis is restricted to those who are married (as is sometimes the case), the IGE of the geometric mean is still undefined because spouses may not enter the labor force or otherwise contribute any earnings over their lifetimes.

The latter point seems to have been missed by Chadwick and Solon (2002) in their influential study of the role that assortative mating plays in the intergenerational persistence of

economic status (see also Blanden 2005; Ermisch et al. 2006). When confronted with the zeros problem at the level of short-run proxy variables, they note that estimating the IGE of women's earnings is "awkward" due to the "frequency with which daughter's earnings are zero" (2002:340), and they thus propose an indirect approach for estimating the IGE and for assessing the role of assortative mating. This indirect approach, although clearly ingenious, nonetheless wrongly assumes that the IGE of women's long-run earnings is a well-defined estimand.

Moreover, even if we counterfactually assume that the estimand is well defined, Chadwick and Solon's indirect approach does not work. The key problem with their approach is best revealed by examining their empirical analysis for married women. This analysis relies on the following equation:

$$\beta_1 = S\beta_{1h} + (1 - S)\beta_{1w} \qquad [11]$$

where β_1 is the elasticity of family earnings with respect to the women's parental income; β_{1w} and β_{1h} are the respective elasticities of women's and their spouses' earnings with respect to the women's parental income; and *S* is characterized as "the typical share of husband's earnings in combined earnings" (Chadwick and Solon 2002:340).

Within the context of this analysis, Chadwick and Solon attempt to address the (short-run) zeros problem by first estimating, via OLS, the family-earnings IGE and the IGE of the husband's earnings. These estimates, when combined with empirical information on the "typical earnings share," can then be plugged into Equation [11] to (a) infer the IGE of the wife's earnings, and (b) assess how much each of the spouses' IGEs contributes to the total-earnings IGE (thereby assessing the impact of assortative mating).

The problem with this analysis is that Equation [11] does not hold for the context in which Chadwick and Solon need it to hold. If the analysis involved mathematical or deterministic rather than random variables, it would of course be true that the elasticity of family earnings is, as Chadwick and Solon (2002:337) put it, the "share-weighted average of the separate elasticities of the daughter's own earnings and her husband's." In this deterministic context, *S* is simply the ratio between the husbands' earnings and the total family earnings. However, their analysis is meant to apply to random variables, hence they implicitly assume that Equation [11] is also valid for random variables. This is not the case. In particular, note that by defining the "typical earnings share" in the random-variable context either as $S = E(Y_h)/[E(Y_h) + E(Y_w)]$) or as $S_x = E(Y_h|x)/\,[E(Y_h|x) + E(Y_w|x)]$, it doesn't follow that $\beta_1 =$

$\frac{\partial E(\ln Y|x)}{\partial \ln x}$ is the weighted average of $\beta_{1h} = \frac{\partial E(\ln Y_h|x)}{\partial \ln x}$ and $\beta_{1w} = \frac{\partial E(\ln Y_w|x)}{\partial \ln x}$ described by that equation (with S replaced by S_x when relevant).[13] We can conclude that Chadwick and Solon's indirect approach does not achieve its goals even if we make the dubious assumption that the IGE of the geometric mean of long-run earnings is defined for women.

The simple conclusion here: We cannot study the role of marriage in the transmission of economic advantage across generations if we insist on relying on the IGE of the geometric mean. Despite ingenious attempts at circumventing its limitations, the conventional IGE cannot be used to assess the role that marriage and assortative mating play in the transmission process, nor to study gender differences in the relative importance of labor and marriage markets. If we want to understand the channels through which economic advantage is passed from parents to children, there is no alternative to developing an approach that can authentically accommodate zero earnings or income.

V. The case for a new IGE

We have shown that the conventional IGE is affected by very serious conceptual and methodological shortcomings and that it's not enough, therefore, to simply repair the language with which it's described. In making this case, we have sought to lay out these shortcomings quite exhaustively, as we appreciate that, given the long tradition behind the conventional IGE, the decision to abandon it cannot be made lightly. If one accepts that the problems are indeed too fundamental to continue to ignore, this of course leaves us in need of an alternative measure. It might be argued that, because of the many problems with the conventional IGE, we have no choice but to move to an altogether new class of measures, such as the rank-rank slope or measures based on mobility tables (e.g., mobility rates across income quintiles).[14] While there are surely research situations in which these measures provide answers to the key questions at stake, there are also many situations in which the concreteness of a dollar-based measure, and its easy embedding within theoretical models of intergenerational processes, is invaluable. For these research situations, we need a new IGE that restores the conventional interpretations wrongly attached to the old IGE, handles zeros without any problems, solves the selection bias problem, and allows us to study the channels through which economic advantage is transmitted. We will introduce a new IGE—the IGE of the expectation— that does all of this. The best course of action, we will argue, is to keep the conventional language for describing and interpreting the

IGE, but now treat the IGE of the expectation—rather than the IGE of the geometric mean—as the intergenerational elasticity of reference.

A. The IGE of the expectation and its interpretations

We begin, then, by demonstrating that the conventional interpretations can be retained when the IGE of the expectation is estimated. If the assumption of constant elasticity is maintained, the PRF of Equation [1] can be substituted with the following PRF:

$$\ln E(Y|x) = \alpha_0 + \alpha_1 \ln x, \qquad [12]$$

where $Y \geq 0$, $X > 0$, and $\alpha_1 = \frac{d \ln E(Y|x)}{d \ln x}$ is the percentage differential in the expectation of children's long-run income with respect to a marginal percentage differential in parental long-run income. This new PRF thus allows us to estimate an elasticity that matches the core interpretation that has long been misapplied to the conventional IGE. Because the expectation is well defined even when some children have zero income or earnings, the IGE of the expectation is immune to the selection bias affecting the conventional IGE.

It can also be shown that, in addition to salvaging the core interpretation, the IGE of Equation [12] retains all other interpretations incorrectly applied to the conventional estimand (with a few small variations discussed below). The following equations and inequalities are all derivable from Equation [12]:

$$\alpha_1 = \frac{\ln E(Y|x_2) - \ln E(Y|x_1)}{\ln x_2 - \ln x_1} \qquad [13]$$

$$\frac{E(Y|x_2)}{E(Y|x_1)} = \left(\frac{x_2}{x_1}\right)^{\alpha_1} \qquad [14]$$

$$\ln E(Y|x) - \ln E(Y|E(X)) = \alpha_1(\ln x - \ln E(X)) \qquad [15]$$

$$\ln E(Y|x) - \ln E(Y) \cong \alpha_1(\ln x - \ln E(X)) \text{ if } CV(X) < 1 \qquad [16]$$

$$\ln E(Y|X > x) - \ln E(Y) < \ln E(X|X > x) - \ln E(X) \text{ if } x \geq E(X) \qquad [17a]$$

$$\ln E(Y|X < x) - \ln E(Y) > \ln E(X|X < x) - \ln E(X) \text{ if } x \leq E(X) \qquad [17b]$$

$$n = \ln \left| \frac{k}{\ln y_0 - \ln E(Y_0)} \right| (\ln \alpha_1)^{-1}. \qquad [18]$$

The share interpretation of the IGE is expressed by Equation [13] (which follows directly from Equation [12]). It states that, for any two different values of parental income, α_1 is the percent difference between the expectations of the children's income expressed as a share of the percent difference in the corresponding parental incomes. Or, more compactly, α_1 is the share of

percent differences in parents' income found between the expected incomes of their children.[15]

We earlier referred to a version of Equation [14] in which β_1 replaced α_1. We then showed that this expression rested on the erroneous assumption that the conventional IGE pertains to the conditional mean (Corak 2004:11-12). If the PRF defined by Equation [12] is instead used, Equation [14] follows immediately. Moreover, when it is also assumed that $Y|x_1$ and $Y|x_2$ are independent for $x_1 \neq x_2$, it can be shown that $E(Y|x_2[Y|x_1]^{-1}) \geq (x_2[x_1]^{-1})^{\alpha_1}$ (see Appendix A). This expression provides a lower bound for the expected percent income advantage (expressed as a ratio) of a child with parental income x_2 over a child with parental income x_1 (for $x_2 > x_1$).[16]

The next four equations and inequalities (i.e., [15] - [17b]) express the sense in which a "regression to the mean" interpretation can be recovered. The key backdrop to this discussion is that, insofar as there is a constant-elasticity relationship between X and $E(Y|X)$ (where the elasticity is positive but smaller than one), the locus of all $\big(x, E(Y|x)\big)$ points will not pass through $\big(E(X), E(Y)\big)$. This means that the conditional expectation of the children's income under constant elasticity will not regress to the unconditional mean. Although the PRF of Equation [12] is subject to this stricture, a "regression to the mean" interpretation *is* available in the form of regression to the expected income of children born to mean-income parents (in percent terms). Moreover, when the parental-income variable is not too dispersed, this form of regression approximates regression to the unconditional mean. No matter how much dispersion there is, the PRF of Equation [12] also entails "generalized regression to the mean in percent terms," a type of regression to the mean that is resonant with the standard notion of regression to the mean in percent terms.

These three conclusions, which are expressed in Equations [15] and [16] and Inequalities [17a] and [17b], thus form the backbone of the "regression to the mean" interpretation available under the PRF of Equation [12]. The first of our three conclusions, as expressed in Equation [15], follows immediately from Equation [12]. It implies that there is regression to the expected income of the children born to mean-income parents whenever $0 \leq \alpha_1 < 1$ (where α_1 determines how rapid that regression is). Similarly, Equation [16] implies that there is approximated regression to the unconditional mean when the coefficient of variation of the parental-income distribution, $CV(X)$, is smaller than one (see Appendix A for the derivation of this equation). The latter condition typically holds in the survey data employed by mobility

scholars, but may not hold with administrative data (given the long right tails characteristic of such data). If $\alpha_1 = 1$, Equations [15] and [16] imply that any percent difference from the mean in the parental generation is found again in the next generation among the corresponding children. This condition implies, in other words, that children do not regress to the unconditional mean or to the expected income of children born to mean-income parents. At the other extreme, $\alpha_1 = 0$ implies that there is full regression in just one generation, regardless of the size of the percent difference (from the mean) in the initial generation. If $0 < \alpha_1 < 1$, there is some regression to the expected income of the children born to mean-income parents (and to the unconditional mean), but it is not complete.

Inequalities [17a] and [17b] refer to what might be called "generalized regression to the mean in percent terms."[17] This notion, which is closely related to Samuels' (1991) "reversion to the mean," can also be labelled "reversion to the mean in percent terms." Like the usual notion of regression to the mean, generalized regression to the mean implies that the conditional expected income gets closer to the unconditional mean after one generation, but this holds only on average for children raised in all possible tails of the parental income distribution (rather than for each individual child).[18] When $0 \leq \alpha_1 < 1$, there is generalized regression to the mean, with the value of α_1 determining how rapid that regression is (see Appendix A for proofs and for additional details).

By positing a first-order Markov process, Equation [18] can be obtained when one also assumes that parental income is independent of the error term (as opposed to the weaker assumption of zero conditional mean error).[19] The resulting equation expresses, as a function of α_1, the number of generations (*n*) that it will take the average descendant of a family at a certain income level (e.g., poverty) to come to within *K* percent of the expected income of the children born to mean-income parents (with Y_0 denoting the income of the current generation). When the coefficient of variation of the parental-income distribution is smaller than one, *n* is also the approximate number of generations that it will take the average descendant of a family at a certain income level to come to within *K* percent of a country's average income. The arbitrary threshold value used to stipulate that full regression has occurred is $k = K/100 \ll |\ln y_0 - \ln E(Y_0)|$ (see Appendix A for the derivation of this equation).

The IGE of the expectation can thus be used to characterize (a) the share of inequality between parents that is transmitted to their children, (b) the economic advantage that a child from

a better-off family may expect, (c) the extent to which there is generalized regression to the mean, (d) the number of generations needed to regress to the expected income of children born to mean-income parents, and (e) the number of generations needed to regress to the arithmetic mean (when the parental income distribution is not too dispersed). We can conclude that estimating the IGE of the expectation allows us to retain the many interpretations that have long made the IGE attractive to mobility scholars.[20]

B. Studying transmission channels

The interpretative and methodological advantages discussed so far are not the only reasons why the IGE of the expectation is an attractive choice for mobility analysis. It is also attractive because conditional-expectation models, which are the models of choice throughout the sciences, allow us to exploit the wide-ranging virtues of a linear operator. Although other operators (e.g., the median) may well be attractive as measures of central tendency, they can't be easily used for analytic decompositions. This key property of a conditional-expectation model matters when one wants to study the channels (e.g., labor market, marriage market) through which the intergenerational transmission of advantage occurs. It is useful, then, to conclude this section by introducing a decomposition that allows us to assess the relative importance of such "transmission channels."

Because the IGE of the expectation is well defined in the presence of zeros, it is obvious that the IGE of women's earnings can now be unproblematically estimated. By relying on the IGE of the expectation, we can also estimate the relative importance of the "direct pathway" (via own income) and the "indirect pathway" (via marriage) through which inequality in family income is transmitted across generations. It is straightforward to derive an expression showing how the family-income elasticity depends on (a) the elasticity of the expectation of the child's own income, (b) the elasticity of the expectation of the spouse's income conditional on marriage, and (c) the elasticity of the probability of marriage. This expression, which does *not* assume that the IGE is constant across levels of parental income, is as follows:

$$E\big(\mathrm{K}(Y|x)\big) = E(S_x)\,E\big(\mathrm{K}(Y_o|x)\big) + (1 - E(S_x))\left[E(\mathrm{K}(M|x)) + E\big(\mathrm{K}(Y_s|x, M = 1)\big)\right] + C, \qquad [19]$$

where K is the IGE of the expectation operator; Y, Y_o, and Y_s are family total income, the child's own income, and her or his spouse's income, respectively (all of which may be zero); S_x is the ratio of the child's expected own income to expected family income when parental income is x;

M is a dummy variable indicating whether the child is married or not; and $C = Cov\left(S_x, \mathrm{K}(Y_o|x) - \mathrm{K}(M|x) - \mathrm{K}(Y_s|x, M)\right)$. The latter covariance has been found to be very close to zero in empirical applications. Moreover, if the elasticities are all constant across levels of parental income, then Equation [17] reduces to:

$$\mathrm{K(Y)} = S\,\mathrm{K}(Y_o) + (1 - S)\,[\mathrm{K}(M) + \mathrm{K}(Y_s|M = 1)].^{21} \qquad [20]$$

Unlike Chadwick and Solon's (2002) decomposition of the conventional IGE, Equations [19] and [20] are formulated in terms of random variables and are therefore applicable in empirical research.

Why is this decomposition important? It's important because it allows us to (a) retain all children for the analysis even when they are single or have spouses without income, (b) estimate the relative importance of the direct and indirect pathways through which inequality in family income is transmitted, and (c) determine whether there are gender differences in the processes by which economic status is transmitted across generations (see Mitnik et al. 2015:36-39 and 64-68). The IGE of the geometric mean cannot, by contrast, deliver on any of these objectives. This is a major handicap insofar as we'd like to examine the effects on intergenerational transmission of the strengthening association between family income and marriage, the growth of assortative mating, or the closing of the gender gap in labor force participation.

VI. Estimation of the IGE of the expectation

The conventional IGE has typically been estimated with the OLS estimator, two-stage least squares (TSLS) estimator, or two-sample two-stage least squares (TSTSLS) estimator. If the IGE of the expectation is to become a new workhorse measure of economic persistence, it is necessary to identify estimators for it that can play the same roles that those estimators have played for the conventional IGE. We briefly introduce a suite of such estimators here. For more details on estimation, including the rationale for preferring the estimators we discuss relative to other possible estimators, see Mitnik (2017a; 2017b; 2017c; 2017e).

Estimation with multi-year averages of parental income. The Poisson pseudo maximum likelihood (PPML) estimator (Santos Silva and Tenreyro 2006; 2011) is very well suited for playing the role that the OLS estimator has played for the conventional IGE. After substituting short-run for long-run income measures in Equation [12], the IGE of the expectation can be estimated with the PPML estimator in any context in which the conventional IGE can be estimated with the OLS estimator. Moreover, just as averaging multiple years of parental income

has been shown to reduce attenuation bias with the OLS estimator of the conventional IGE, so too Mitnik (2017a) has used a generalized error-in-variables model to show that the same averaging will reduce attenuation bias with the PPML estimator of the new IGE. At the same time, Mitnik (2017a) shows that both left-side and right-side lifecycle biases disappear when the short-run income measures pertain to the "correct ages" of parents and children, again a result that matches that which obtains in the conventional context. This analysis thus replicates Haider and Solon's (2006) analysis of OLS estimation of the conventional IGE.[22]

Estimation with instrumental variables. The conventional IGE is sometimes estimated by using parental education or occupation as instruments for the short-run parental income measures (e.g., Ng 2007; Mulligan 1997; Solon 1992; Zimmerman 1992). These instruments are, however, invalid insofar as they are correlated with the error term of the PRF of interest. In an important analysis by Solon (1992: Appendix), it was shown that under plausible empirical assumptions IV estimation with these instruments is upward inconsistent, meaning that the resulting IGE estimates can be interpreted as upper-bound estimates (see also Mitnik 2017b). The main question for our purposes is whether similar conclusions hold for the IV estimation of the IGE of the expectation. The additive-error version of the generalized method of moments (GMM) IV estimator of the Poisson or exponential regression model (Mullahy 1997; Windmeijer and Santos Silva 1997) can be used to estimate the IGE of the expectation in any context in which the TSLS estimator (or any other linear IV estimator) can be used to estimate the conventional IGE (Mitnik 2017b). We will refer to that estimator as the GMM-IVP estimator. Using generalized error-in-variables models, Mitnik (2017b) shows that this GMM-IVP estimator can be expected to produce upper-bound estimates with the instruments typically available to mobility scholars. This conclusion holds for the same conditions (regarding the instruments and the ages at which incomes are measured) under which the TSLS estimator may be expected to produce upper-bound estimates of the conventional IGE. It follows that a comparable set of estimators and methodological results are indeed available for the IV estimation of both IGEs.

Set estimation. For the conventional IGE, the OLS estimator is downward inconsistent (unless many years of parental information are available), while the TSLS estimator is upward inconsistent (with the instruments typically available). This means that the "the probability limits of the two estimators bracket the true value" of that IGE (Solon 1992:400). In an argument that nicely anticipated the current interest in the estimation of partially identified parameters, Solon

(1992) then suggested combining these estimators to produce what we may call—using terminology that is standard today—a set estimate of the conventional IGE. As Mitnik (2017b) shows, the same approach can be used for the IGE of the expectation, although in this case the set estimate is obtained by combining estimates generated with the PPML and GMM-IVP estimators. The confidence intervals for the partially identified IGEs can be constructed, as Mitnik (2017b) lays out, using Imbens and Manski's (2004) approach.

Estimation in the two-sample context: Because some countries don't have samples with information on children-parent pairs, the conventional IGE is often estimated with the TSTSLS estimator using short-run income measures drawn from two independent samples (see Jerrim et al. 2016 for a review). The resulting estimates are then interpreted as upper-bound estimates (e.g., Björklund and Jäntti 1997) or as upward-biased estimates that can be adjusted ex post to make them comparable to OLS estimates (e.g., Corak 2006: Appendix). The IGE of the expectation can be estimated under these same data constraints by using the two-sample GMM estimator of the exponential regression model (see Mitnik 2017c). This estimator produces upper-bound estimates under the same conditions (regarding the instruments and the ages at which incomes are measured) that apply in the case of the TSTSLS estimator of the conventional IGE (see Mitnik 2017c).

The foregoing indicates that (a) the IGE of the expectation can be estimated in all contexts in which the conventional IGE has been (or might be) estimated, and (b) the necessary methodological foundation for estimating the former IGE with short-run variables is available. It is especially important that lifecycle and attenuation biases, which are of course central worries in the field, can be readily addressed when estimating the IGE of the expectation. The estimators discussed in this section are, moreover, available in statistical packages widely used by social scientists.[23] It follows that one's choice of IGE can be based exclusively on the merits of the case.

There is but one remaining estimation concern. Are estimates of the IGE of the expectation sensitive to assumptions about the income of children with missing income reports? It may be recalled that analysts of administrative data often lack information on income (e.g., tax nonfilers) or earnings (e.g., those working in the informal economy).[24] This missing-data problem has led some mobility scholars to turn away from the (conventional) IGE in favor of the rank-rank slope (e.g., Chetty et al. 2014; see also Dahl and DeLeire 2008). In the very influential

study of Chetty et al. (2014), the rank-rank slope was initially introduced as a fallback that became necessary because, when different assumptions were made about the income of children with missing data, the resulting IGE estimates fluctuated markedly. If the IGE of the expectation had the same shortcoming, its appeal as a mobility measure would be greatly diminished, given that missing data of this sort are ubiquitous.

Because the missing data of concern are clearly "nonignorable," it is unwise for researchers to either (a) drop children with missing data from the analysis, or (b) use mean or multiple imputation based on the "missing-at-random" assumption (e.g., Little and Rubin 2002). At the same time, the mechanisms that render the missing data nonignorable (e.g., income below the filing threshold) also suggest that children with missing data will typically have income or earnings that are very low and, moreover, are unlikely to vary much by parental income. This suggests estimating IGEs by imputing to children with missing data a low income or earnings value. If an auxiliary dataset with information on nonfiling or on work in the informal economy were available, imputation of mean values for the main socio-demographic groups becomes feasible.[25]

This leads to our key question: Are estimates of the IGE of the expectation robust to the values imputed under this approach? The conventional approach, it may be recalled, is very problematic on this matter: As Mitnik et al. (2015:46-50) show for nonfilers in 2010 U.S. tax data, and for men without W2 earnings in that same year, the estimates of the conventional IGE vary wildly across different imputations (as also reported by Chetty 2014: Table 1). This is the case because the IGE of the geometric mean is dominated by small absolute differences at the lowest quantiles of the children's conditional distributions (see Appendix B). We can straightforwardly eliminate this sensitivity to different imputation approaches and imputed values by instead estimating the IGE of the expectation (Mitnik et al. 2015: 46-50). Although Dahl and DeLeire (2008) and Chetty et al. (2014) turned away from the conventional IGE because its estimates were fragile, Mitnik et al.'s (2015) results thus show that there is no need to resort to the rank-rank slope for wholly methodological reasons. The choice of mobility measure can be made exclusively on the basis of the research questions at hand.

VII. Illustrative empirical analyses

We have argued to this point that the conventional IGE has been widely misinterpreted, is undefined in many important cases, and is poorly suited for estimation with short-run income

and earnings measures. The pragmatist might still wonder, however, whether empirical analyses based on the conventional IGE have in practice led us all that far astray (compared to the results that would have been obtained by estimating the IGE of the expectation). Are we, in other words, worrying too much? We address this question here by using three key U.S. datasets to estimate IGEs of men's and women's earnings with respect to parental income. We then consider whether our estimates of the conventional IGE are or are not trivially different from our estimates of the IGE of the expectation.

The first dataset we use is the Statistics of Income Mobility (SOI-M) Panel. Built from U.S. tax returns, W-2 forms, and other administrative sources, this panel represents all children born between 1972 and 1975 who were living in the U.S. in 1987. The analysis here, based on the sample employed by Mitnik et al. (2015), pertains to SOI-M children in 2010 (when they were 35-38 years old). We use information on the annual earnings of children in 2010, parents' average after-federal-tax income (including refundable tax credits) when the children were 15-23 years old, and parents' average age (again when the children were 15-23 years old).[26]

The second sample, drawn from the National Longitudinal Survey of Youth 1979 (NLSY79), includes children who were 14-16 years old in 1979 (and thus born between 1963 and 1965). We use information on the annual earnings of children in 2003 and 2005 (when they were 38-42 years old), parents' average total family income in 1978-1980, and parents' age in 1979. The observations in this sample are children-years (as in Lee and Solon 2009), with most of the children appearing in the sample twice, once in 2003 and then again in 2005.

The third sample, drawn from the Panel Study of Income Dynamics (PSID), pertains to household heads and wives (or partners) born between 1954 and 1966 and observed at least once when they were between 35 and 45 years old. The observations in this sample are again children-years (with a separate record for each year in which a child appears in the PSID). We measure children's annual earnings at ages 35-45 and parents' average total family income and age when the children were 13-17 years old. We have provided descriptive statistics for the three samples in Table 1 (see Appendix C for additional details on the data and variables).

Throughout our analyses, we use the OLS estimator to estimate the conventional IGE and the PPML estimator to estimate the IGE of the expectation, assuming in both cases that the elasticity is constant across levels of parental income. When estimating the conventional IGE, mobility scholars often not only drop children without earnings, but also children with low

earnings and children whose parental income is low. We reflect this convention by estimating the conventional IGE using (a) a subsample of children with positive earnings and positive average parental income, and (b) two subsamples in which children with annual earnings or average parental income lower than $600 and $1,500 (in 2010 dollars) respectively are dropped. The estimates of the IGE of the expectation are based on the full samples.

It is also customary to include polynomials on children's and parents' ages to "absorb" the effects of age at measurement on the relationship between long-run and measured income or earnings (both for parents and children). We thus include dummies for children 38-41 years old and 42-45 years old in all of our PSID-based models. In our SOI-M and NLSY79 analyses, the variation in children's ages is quite small, which makes controlling for age unnecessary. Moreover, because the age at which parents have their children is not exogenous to their income, and because parental age may affect their children's life chances, Mitnik et al. (2015:34) argued that controlling for parental age is inconsistent with the objective of measuring the gross association between parental and children's income. For this reason, we will provide estimates both with and without controls for parental age, with our controls taking the form of a quadratic polynomial in age. We use sampling weights and compute cluster-corrected robust standard errors in all of our analyses.

We report the IGE estimates in Table 2. The estimates for the conventional IGE, although varying slightly across sample restrictions and for models with and without age controls, fall within the characteristic ranges of past analyses given the number of years of parental information used (see Mitnik et al. 2015:7-15). The estimates for men range from 0.33 to 0.35 for the SOI-M sample, from 0.43 to 0.45 for the NLSY79 sample, and from 0.42 to 0.46 for the PSID sample.

The key question is of course whether the estimates for the IGE of the expectation, which are based on the full sample, are substantially different. The results of Table 2 establish very clearly that they are. We find that the estimates are especially different in the sample that has the lowest estimates for the conventional IGE (i.e., the SOI-M Panel). The estimates for the IGE of the expectation in the SOI-M Panel are 37 to 42 percent larger than the corresponding estimates for the conventional IGE, whereas the increases in estimates for the NLSY79 and PSID samples are a somewhat lower 22 to 27 percent and 11 to 19 percent respectively. We can conclude that, for men's earnings, the conventional approach substantially understates economic persistence

and the transmission of advantage across generations.

The differences in the women estimates are also large. With the SOI-M Panel, the IGE of the expectation, estimated at 0.26 or 0.27, is between 21 and 51 percent larger than the estimates of the conventional IGE. The estimates in the NLSY79 and PSID samples are 12 to 20 percent and 43 to 63 percent larger than the corresponding estimates of the conventional IGE.

These results suggest that, for women and men alike, the extent to which economic advantages are transmitted intergenerationally has been understated in U.S. mobility analyses that rely on the conventional earnings IGE. We suspect that the great many trend analyses, cross-national comparisons, and within-country subgroup comparisons using the conventional IGE may likewise be misleading. Because selective processes vary by time, country, and subgroup, there is good reason to worry that seemingly well-established comparative results may come to be called into question.

VIII. Conclusions

The intergenerational elasticity has long been the workhorse measure of economic mobility. When any measure becomes the convention, its stock interpretations can become entrenched and unquestioned, even though they may have rested on rough-and-ready justifications. This is the case with the conventional IGE. Although its archetypal interpretation as a measure of regression to the arithmetic mean is certainly attractive, the conventional IGE has not been specified in a way that justifies that interpretation or, for that matter, any of its other frequently-rehearsed interpretations.

It might be possible to develop a conceptual rationale for the conventional IGE that is both theoretically appealing and consistent with its de facto definition. This rationale, even if it could be developed, would only get us halfway to salvaging the conventional IGE. We would still have to try to solve the deep methodological problems with which any measure based on the geometric mean is going to be saddled. It is unclear why one would want to invest in developing a compelling conceptual rationale for a measure that is so methodologically unsound.

We have sought to stress just how profound these methodological problems are. The core problem: The conventional estimand is always undefined because even long-run measures of children's income or earnings always include zero in their support. Worse yet, the probability mass at zero becomes much larger when long-run measures are proxied with short-run ones, as is almost always necessary. When faced with this problem, mobility scholars have typically

resorted to the expedient of dropping children without earnings or income, a response that generates substantial selection biases. The conventional IGE is yet less attractive when analyzing the transmission of labor market advantages among daughters or the indirect transmission of advantage (via the marriage market) among sons or daughters. For these analyses, even the long-run measures have especially substantial probability masses at zero, thus making it impossible to assess the contribution of differential marriage chances (across parental-income levels) to observed family-income IGEs.

The conceptual and methodological problems we have discussed all arise directly from the mobility field's unwitting use of a PRF that references the conditional geometric mean. The solution to these problems involves replacing this PRF with one referencing the conditional expectation of the child's earnings or income and then using well-understood estimators to estimate it. This simple fix not only recovers all interpretations wrongly ascribed to the conventional IGE (or small variations thereof), but also solves all methodological problems discussed in this paper.

It follows that researchers can now estimate the IGE whenever the research questions at hand dictate doing so. Although the methodological problems with the conventional IGE have led some analysts to fall back on the rank-rank slope or other measures, the choice of a mobility measure should be made exclusively on the basis of its capacity to answer the research questions of interest. It is important for many research questions to retain a measure that (a) rests concretely on dollars rather than a transformation of dollars into ranks, and (b) can be easily embedded within theoretical models of intergenerational processes. These two objectives can be met by redefining the IGE as we have proposed.

The methodological problems that we have discussed cannot be addressed by simply adjusting how we interpret the conventional IGE. We have provided evidence, to the contrary, that conventional estimates of earnings IGEs understate—often substantially—the extent to which economic advantages are transmitted across generations. This raises the possibility that the field's stock results on mobility trends, cross-national variability, and subgroup variability may likewise be misleading. Although we have not sought to redo all of the many important analyses based on the conventional IGE, the evidence provided here suggests that there is no alternative but to turn now to that massive task.

Appendices

A. Mathematical proofs

We refer to equations presented in the main text using the equation numbers employed there. When equations from the main text are reproduced in the appendices, we rely on their original numbers in the main text.

IGE of the geometric mean of the descendants' income in the nth generation

We have claimed in the main text that β_1 cannot be used to assess how many generations it will take the average descendant of a family at a certain income level (e.g., the poverty level) to come to within *K* percent of a country's average income (where *K* is a small number). Under the (strong) assumption of a first-order Markov process, we have further claimed that β_1 can only tell us how many generations it will take for the geometric mean of the descendants' income to be close to the overall geometric mean of income in their generation.

To derive this conclusion, we can rewrite Equation [1] in error form, as

$$\ln Y_n = \beta_{n,0} + \beta_{n,1} \ln Y_{n-1} + \Upsilon_n,$$

where Υ_n is an additive error term, $E(\Upsilon_n | y_{n-1}) = 0$, and the added subscripts $(n, n-1)$ denote the number of generations after the "current generation" (with zero denoting the current generation itself). Assuming $\beta_{n,1} = \beta_{n-1,1} = \ldots = \beta_{0,1} = \beta_{.,1}$, (i.e., the IGE is stationary), it follows from the Markov assumption that $(\beta_{.,1})^n$ is the elasticity of the geometric mean of the descendants' income in the nth generation with respect to family income in the current generation. Indeed, from the assumption of stationarity, it immediately follows that when $n = 2$:

$$E(\ln Y_2 | y_0) = (\beta_{2,0} + \beta_{.,1}\beta_{1,0}) + (\beta_{.,1})^2 \ln y_0,$$

where $E(\Upsilon_2 | y_0) = E(\Upsilon_2) = 0$ per the Markov process assumption. This means that $(\beta_{.,1})^2$ is the elasticity of the geometric mean of the descendants' income in the second generation with respect to the family income of the current generation.

It is straightforward to generalize this result and show that, under the same assumptions, $(\beta_{.,1})^n$ is the elasticity of the geometric mean of the descendants' income in the nth generation with respect to family income in the current generation. By adapting Equation [3] to the multigenerational context, we then have:

$$(\beta_{.,1})^n = \frac{k}{|\ln y_0 - \ln GM(Y_0)|}$$

$$n = \ln\left|\frac{k}{\ln y_0 - \ln GM(Y_0)}\right| \left[\ln \beta_{.,1}\right]^{-1}, \qquad [4]$$

where $k = \frac{K}{100} \ll |\ln y_0 - \ln GM(Y_0)|$ is the arbitrary threshold value used to stipulate that full regression to the geometric mean has occurred (i.e., full regression is stipulated to occur when $|\ln GM(Y_n|y_0) - \ln GM(Y_n)| \leq k$). As indicated in the text, it is assumed that $0 < \beta_{.,1} < 1$.

Expected economic advantage as a function of the IGE of the expectation

After explaining that Equation [14] follows from Equation [12], we stated in the main text that:

$$E\left(\frac{Y|x_2}{Y|x_1}\right) \geq \left(\frac{x_2}{x_1}\right)^{\alpha_1},$$

which provides a lower bound for the expected percent income advantage (expressed as a ratio) of a child with parental income x_2 over a child with parental income x_1 (for $x_2 > x_1$). To derive this expression, we can write Equation [12] in multiplicative error form as:

$$Y = \exp(\alpha_0)\, x^{\alpha_1}\, \Omega$$

where $E(\Omega|\mathrm{x}) = 1$. The lower-bound expression can be derived by further assuming $\Omega|x_1 \perp \Omega|x_2$. This means that the population errors, conditional on any two different values of parental income, are statistically independent. Given this assumption, $Y|x_2$ and $Y|x_1$ are independent, which means that $Y|x_2$ and $Y^{-1}|x_1$ are independent as well, which in turn entails

$$E\left(\frac{Y|x_2}{Y|x_1}\right) = E(Y|x_2)E(Y^{-1}|x_1).$$

Due to Jensen's inequality, $E(Y^{-1}|x_1) \geq \frac{1}{E(Y|x_1)}$, so the inequality above follows (where the inequality is strict unless $Y|x_1$ has a degenerate distribution).

IGE of the expectation and regression to the mean

In this section, we derive Equation [16], which shows that there is approximated regression to the mean (in percent terms) when the parental-income distribution has a coefficient of variation smaller than one. We also show that, per inequalities [17a] and [17b], there is generalized regression to the mean in percent terms, regardless of the dispersion of the parental-income distribution.

To derive Equation [16], we first exponentiate both sides of Equation [12] and take expectation over the population distribution of *X*. If we then take logarithm, this yields:

$$\ln E(Y) = \alpha_0 + \ln E(X^{\alpha_1}). \qquad [A1]$$

We use the last expression and Equation [12] to yield the following:

$$\ln E(Y|x) - \ln E(Y) = \ln x^{\alpha_1} - \ln E(X^{\alpha_1}). \qquad [A2]$$

Equation [A2] indicates that, if $\alpha_1 = 1$, any percent difference from the mean in the parental generation is found again in the next generation among the corresponding children (i.e., no regression to the mean). If $\alpha_1 = 0$, there is full regression to the mean in just one generation regardless of the percent difference in the initial generation.

Equation [16] differs from Equation [A2] by substituting $[E(X)]^{\alpha_1}$ for $E(X^{\alpha_1})$. This approximation is based on the following second-order Taylor-series approximation around $E(X)$:

$$E(X^{\alpha_1}) \cong [E(X)]^{\alpha_1} - 0.5\,\alpha_1\,(1 - \alpha_1)[E(X)]^{\alpha_1}\,[CV(X)]^2,$$

where $CV(X) = \frac{[Var(X)]^{1/2}}{E(X)}$ is the coefficient of variation of X. When $0 < \alpha_1 < 1$, this approximation entails not only that $E(X^{\alpha_1}) < [E(X)]^{\alpha_1}$ (which follows from Jensen's inequality) but also that the proportional difference between $E(X^{\alpha_1})$ and $[E(X)]^{\alpha_1}$, expressed as a ratio, is:

$$1 - 0.5\,\alpha_1\,(1 - \alpha_1)\,[CV(X)]^2.$$

The term $0.5\,\alpha_1\,(1 - \alpha_1)$ is largest in absolute value when $\alpha_1 = 0.5$. Therefore, the ratio between $E(X^{\alpha_1})$ and $[E(X)]^{\alpha_1}$ will be smaller than one by a *maximum* of $0.125\,[CV(X)]^2$, which can be expected to be quite small as long as $CV(X) < 1$. This is typically the case with survey data (but not with administrative data). For instance, with the NLSY79 and PSID samples employed in Section VII, the coefficient of variation is 0.67 and 0.69, respectively, so that $E(X^{\alpha_1})$ is at most six percent smaller than $[E(X)]^{\alpha_1}$. The foregoing indicates that, when $0 < \alpha_1 < 1$ and $CV(X) < 1$, there is approximate regression to the mean in percent terms:

$$\ln E(Y|x) - \ln E(Y) \cong \alpha_1(\ln x - \ln E(X)). \qquad [16]$$

As indicated in note 17, regression to the mean in percent terms requires that the following conditions hold (with $\min(X) \leq x \leq \max(X)$):

$$\ln E(Y|x) - \ln E(Y) < \ln x - \ln E(X) \text{ if } x > E(X)$$
$$\ln E(Y|x) - \ln E(Y) > \ln x - \ln E(X) \text{ if } x < E(X).$$

The term "reversion to the mean," as defined by Samuels (1991), requires that the two marginal distributions have the same mean. In the main text, we introduced the notion of "generalized regression to the mean in percent terms," which is similar to Samuel's "reversion to the mean,"

but is defined in percent terms (and does not need the assumption of equal means). For this reason, we also referred to it as "reversion to the mean in percent terms." It requires that the following conditions hold (with $\min(X) \leq x \leq \max(X)$):

$$\ln E(Y|X > x) - \ln E(Y) < \ln E(X|X > x) - \ln E(X) \text{ if } x \geq E(X) \qquad [17a]$$

$$\ln E(Y|X < x) - \ln E(Y) > \ln E(X|X < x) - \ln E(X) \text{ if } x \leq E(X). \qquad [17b]$$

We define (a) an upper tail of a distribution as any portion of the distribution between the upper bound of that distribution and a value above (or equal) to its mean, and (b) a lower tail of a distribution as any portion between the lower bound of that distribution and a value below (or equal) to its mean. If the tails pertain to the distribution of the children's conditional expected incomes, we may refer to them as conditional tails.

Under these definitions, generalized regression to the mean in percent terms means that the average income in any conditional tail is closer (in percent terms) to the unconditional mean of the children's income distribution than the average income in the corresponding parental-income tail is to the unconditional mean of the parental-income distribution. Like the usual notion of regression to the mean, generalized regression to the mean indicates that conditional expected income gets closer to the unconditional mean after one generation. This obtains, however, on average for children raised in each possible tail of the parental income distribution (rather than for each individual child). We refer to this as a "generalized form" of regression to the mean because it is implied by the usual regression to the mean (but it does not itself imply the usual regression to the mean).

We next show that generalized regression to the mean in percent terms holds whenever $0 \leq \alpha_1 < 1$ (where α_1 determines how rapid that regression is). To simplify the proof, we assume that $X > e$. This simplification is without loss of generality because it can be achieved by simply changing the monetary units used to measure income.

If we exponentiate both sides of Equation [12] and take expectation conditional on $X > x$, we have:

$$E(Y|X > x) = \exp(\alpha_0)\, E(X^{\alpha_1}|X > x). \qquad [A3]$$

Using Equations [A1] and [A3], Inequality [17a] can be written as:

$$\frac{E(Y|X > x)}{E(Y)} = \frac{E(X^{\alpha_1}|X > x)}{E(X^{\alpha_1})} < \frac{E(X|X > x)}{E(X)} \text{ if } x \geq E(X)$$

$$\frac{E(X^{\alpha_1})}{E(X^{\alpha_1}|X>x)} > \frac{E(X)}{E(X|X>x)} \text{ if } x \geq E(X). \qquad [A4]$$

A sufficient condition for the inequality expressed in [A4] is that the derivative of the left-side term with respect to α_1 is negative. If that derivative is indeed negative, α_1 then determines the rapidity of the generalized regression to the mean.

Let S be the probability that $X > x$. The left-side term of Inequality [A4] may then be written as follows:

$$\frac{E(X^{\alpha_1})}{E(X^{\alpha_1}|X>x)} = \frac{S\,E(X^{\alpha_1}|X>x) + (1-S)\,E(X^{\alpha_1}|X\leq x)}{E(X^{\alpha_1}|X>x)}$$

$$= S + (1-S)\frac{E(X^{\alpha_1}|X\leq x)}{E(X^{\alpha_1}|X>x)}.$$

The sign of the derivative of the last expression with respect to α_1 is determined by the sign of the numerator of the derivative of the ratio of expectations with respect to α_1. The sufficient condition is then:

$$E(X^{\alpha_1}|X>x)E(X^{\alpha_1}\ln X\,|X\leq x) - E(X^{\alpha_1}|X\leq x)\,E(X^{\alpha_1}\ln X\,|X>x) \;<0$$

$$\frac{E(X^{\alpha_1}|X>x)}{E(X^{\alpha_1}|X\leq x)} < \frac{E(X^{\alpha_1}\ln X\,|X>x)}{E(X^{\alpha_1}\ln X\,|X\leq x)}. \qquad [A5]$$

Consider next the following equality:

$$\frac{E(X^{\alpha_1}|X>x)}{E(X^{\alpha_1}|X\leq x)} = \frac{\ln x\,E(X^{\alpha_1}|X>x)}{\ln x\;E(X^{\alpha_1}|X\leq x)} = \frac{E(X^{\alpha_1}\ln x\,|X>x)}{E(X^{\alpha_1}\ln x\,|X\leq x)},$$

We can conclude that Inequality [A5] holds and that, therefore, Inequality [17a] holds as well.

Similarly, Inequality [17b] can be written as:

$$\frac{E(Y|X<x)}{E(Y)} = \frac{E(X^{\alpha_1}|X<x)}{E(X^{\alpha_1})} > \frac{E(X|X<x)}{E(X)} \text{ if } x \leq E(X)$$

$$\frac{E(X^{\alpha_1})}{E(X^{\alpha_1}|X<x)} < \frac{E(X)}{E(X|X<x)} \text{ if } x \leq E(X). \quad [A6]$$

A sufficient condition for the inequality expressed in [*A6*] is that the derivative of the left-side term with respect to α_1 is positive (and, as before, a positive derivative means that α_1 expresses the rapidity of the generalized regression to the mean). That term may be written as:

$$\frac{E(X^{\alpha_1})}{E(X^{\alpha_1}|X<x)} = \frac{(1-S)\,E(X^{\alpha_1}|X<x) + S\,E(X^{\alpha_1}|X\geq x)}{E(X^{\alpha_1}|X<x)}$$

$$= (1-S) + S\frac{E(X^{\alpha_1}|X\geq x)}{E(X^{\alpha_1}|X<x)}.$$

The sign of the derivative of the last expression with respect to α_1 is determined by the sign of the numerator of the derivative of the ratio of expectations with respect to α_1. The sufficient condition is then:

$$E(X^{\alpha_1}|X<x)E(X^{\alpha_1}\ln X\,|X\geq x) - E(X^{\alpha_1}|X\geq x)\,E(X^{\alpha_1}\ln X\,|X<x) > 0$$

$$\frac{E(X^{\alpha_1}|X<x)}{E(X^{\alpha_1}|X\geq x)} > \frac{E(X^{\alpha_1}\ln X\,|X<x)}{E(X^{\alpha_1}\ln X\,|X\geq x)}. \qquad [A7]$$

Consider now the following equality:

$$\frac{E(X^{\alpha_1}|X<x)}{E(X^{\alpha_1}|X\geq x)} = \frac{\ln x\,E(X^{\alpha_1}|X<x)}{\ln x\;E(X^{\alpha_1}|X\geq x)} = \frac{E(X^{\alpha_1}\ln x\,|X<x)}{E(X^{\alpha_1}\ln x\,|X\geq x)},$$

We can conclude that Inequality [A7] holds and that, therefore, Inequality [17b] holds as well.

IGE of the expectation of the descendants' income in the nth generation

We next derive Equation [15]. This equation expresses, as a function of α_1, the approximate number of generations it will take the average descendant of a family at a certain income level (e.g., poverty) to come to within *K* percent of the expected income of the children born to mean-income parents. When the coefficient of variation of the parental-income distribution is smaller than one, *n* is also the approximate number of generations that it will take the average descendant of a family at a certain income level to come to within *K* percent of a country's average income.

We start by rewriting Equation [13] in multiplicative-error form:

$$Y_n = \exp(\alpha_{n,0} + \alpha_{n,1}\ln Y_{n-1})\,\Omega_n,$$

where Ω_n is a (nonnegative) multiplicative error and the added subscripts $(n, n-1)$ denote the number of generations after the current generation (with zero denoting the current generation itself).[27] Equation [13] entails $E(\Omega_n|y_{n-1}) = 1$ but, as indicated in the main text, this assumption is not enough to derive Equation [15]. It is necessary to make the stronger assumption that Ω_n and Y_{n-1} are independent. Under this and the Markov process assumptions,

and stipulating that the IGE is stationary (i.e., $\alpha_{n,1} = \alpha_{n-1,1} = \ldots = \alpha_{0,1} = \alpha_{.,1}$), it follows that $(\alpha_{.,1})^n$ is the elasticity of the descendants' expected income in the n[th] generation with respect to family income in the current generation. For $n = 2$, we have:

$$Y_2 = \exp(\alpha_{2,0} + \alpha_{.,1}\alpha_{1,0} + (\alpha_{.,1})^2 \ln Y_0)\,(\Omega_1)^{\alpha_{.,1}}\Omega_2$$

$$E(Y_2|y_0) = \exp(\alpha_{2,0} + \alpha_{.,1}\alpha_{1,0} + (\alpha_{.,1})^2 \ln y_0)\,E((\Omega_1)^{\alpha_{.,1}}\Omega_2|y_0)$$

$$\ln E(Y_2|y_0) = (\alpha_{2,0} + \alpha_{.,1}\alpha_{1,0} + ln\ C) + (\alpha_{.,1})^2 \ln y_0,$$

where:

C is a positive constant;

$E((\Omega_1)^{\alpha_{.,1}}\Omega_2|y_0) = E((\Omega_1)^{\alpha_{.,1}}|y_0)\,E(\Omega_2|y_0)$ per the Markov assumption;

$E(\Omega_2|y_0) = E(\Omega_2) = 1$ per the Markov assumption; and

$E((\Omega_1)^{\alpha_{.,1}}|y_0) = E((\Omega_1)^{\alpha_{.,1}}) = C$ per the assumption that Ω_n and Y_{n-1} are independent.

This means that $(\alpha_{.,1})^2$ is the elasticity of the descendants' expected income in the second generation with respect to the family income of the current generation. By repeated application of the foregoing analysis, it follows that $(\alpha_{.,1})^n$ is the elasticity of the descendants' expected income in the n[th] generation with respect to family income in the current generation.

Therefore, adapting Equation [13] to the multigenerational context, we have:

$$(\alpha_{.,1})^n = \frac{k}{|\ln y_0 - \ln E(Y_0)\,|}$$

$$n = \ln\left|\frac{k}{\ln y_0 - \ln E(Y_0)}\right| [\ln \alpha_{.,1}]^{-1}, \qquad [16]$$

where $k = \frac{K}{100} \ll |\ln y_0 - \ln E(Y_0)\,|$ is now the arbitrary threshold value used to stipulate that full regression to the expected income of the children born to mean-income parents has occurred. (with the assumption that $\alpha_{.,1} > 0$).

B. Approaches for addressing selection bias in the estimation of the conventional IGE

We considered and rejected three approaches to addressing the selection bias generated by the "zeros problem." Each of these is discussed in more detail next.

Imputing income or earnings

We first considered an approach that entailed retaining children with zero income or earnings in the sample by either (a) imputing a small positive value, or (b) adding a small

positive amount to the earnings or income of all children. We argued that this approach is not viable because very small changes in imputed values lead to disturbingly large changes in estimates. We also pointed out that estimates are fragile because the IGE of the geometric mean is dominated by small absolute differences at the lowest quantiles of the conditional distributions.

The latter point can be demonstrated as follows. Let $Q_Y(x,\tau)$ denote the τ^{th} quantile of the conditional distribution of the child's long-run income when long-run parental income equals *x*. As indicated by Mitnik (2017d:7), at any value of parental income *x*, the conventional IGE can be written as:

$$\frac{d\ln E(\ln Y\,|x)}{d\ln x} = \int_0^1 \frac{x}{Q_Y(x,\tau)}\frac{dQ_Y(x,\tau)}{dx}d\tau. \qquad [B1]$$

Equation [B1] shows that the IGE of the geometric mean can be interpreted as a weighted average of all quantile-specific derivatives of the child's income with respect to parental income. It also indicates that, as we move from larger to lower quantiles, there's an increase in the weight that identical marginal differences receive in the "computation" of the IGE of the geometric mean. This differential weighting is very consequential: The lowest-quantile weights can be expected to be several orders of magnitude larger than the highest-quantile weights.[28]

An example may help here. For a low quantile, if the child's incomes corresponding to two different parental-income values are $100 and $200, the proportional difference is 100 percent. By contrast, the corresponding child's incomes for a high quantile might be $90,000 and $120,000 (for the same parental values), which counts as a proportional difference of only 30 percent even though the absolute difference is 300 times larger.[29]

Modeling selection

The main text also introduced a second approach for addressing selection bias that entailed estimating the conventional IGE with the help of a selection model (i.e., estimating Equation [10]). This approach is difficult to implement because semi-parametric estimators assume an "exclusion restriction" (e.g., Vella 1998; Manski 1989; Moffit 1999; Lee 2003). It's of course extremely hard to identify any variable in typically-used datasets that might satisfy that restriction (especially for estimation of the IGE of men's earnings). Although it has been recently shown that such a restriction is not strictly needed for identification (Escanciano, Jacho-Chávez, and Lewbel 2016), to the best of our knowledge there is no semi-parametric estimator that (a)

does not assume an exclusion restriction, (b) can be employed to estimate Equation [10] with the complex-survey data typically available for the estimation of the IGE (as opposed to data obtained via simple random sampling), (c) has an associated distribution theory that is valid in the complex-sampling context, and (d) does not rely on a tuning parameter, bandwidth, or some similar parameter that is chosen either discretionarily or using data-driven procedures that are non-equivalent (at least in finite samples).

As noted in the text, fully parametric selection models do not require an exclusion restriction, which is of course attractive in principle. However, parametric selection models based on standard bivariate distributions (e.g., the Heckman selection model) are fragile, as seemingly small misspecifications may generate large biases in estimates. This is particularly so in a case, like the one considered here, in which identification would rest exclusively on distributional assumptions (e.g., Manski 1989; Moffit 1999).

By contrast, copula-based selection models (Smith 2003; Hasebe and Vijverberg 2012) provide much more flexibility, which is certainly helpful. Using such models, Mitnik et al. (2017) estimated the conventional IGE using many combinations of copulas and income distributions, with the best models then selected using information criteria. Although copula-based selection models are useful, they involve a complex, computationally expensive, and labor-intensive procedure, which still is not robust enough to provide a viable foundation for general mobility analysis.

Use of different short-run measures

The third approach we considered entails changing the short-run measures employed to estimate the conventional IGE. For the vast majority of researchers, the dependent variable for the analysis has been (a) the logarithm of annual income or earnings (the most common case by a wide margin); (b) the average of the logarithm of income or earnings over a few years of positive income or earnings (which is the same as the logarithm of the geometric mean of children's income or earnings over those years); or (c) the logarithm of average income or earnings over a few years (but usually too few years to make the zeros problem inconsequential). The zeros problem can be addressed by changing the dependent variable to the logarithm of average children's income or earnings (including years in which they are zero) over as many years as necessary to make the zeros problem inconsequential.

There are several reasons why this is an unattractive strategy. First, it has long been

argued that one practical advantage of the IGE, as against the correlation, is that it is "not biased by classical measurement error" in the dependent variable "and so is often easier to estimate with real-world data" (Black and Devereux 2011:1490).[30] If estimation of the IGE requires that the children's short-run measures are averages of many years of data, this practical advantage disappears. Second, assuming that the longitudinal data needed to implement this strategy are available (i.e., annual measurements or, at least, measurements over fairly short intervals), the strategy imposes a long "waiting time" before an IGE for any given cohort can be estimated. The analyst has to wait, in other words, until enough years of data "around age 40" have been collected. This approach cannot, then, be used in the large number of countries that do not have longitudinal datasets with adequate waves of data around age 40.[31] Third, because many countries do not have longitudinal datasets with the required information on children-parents pairs, the IGE has been widely estimated with cross-sectional datasets, typically using the two-sample two-stage least squares (TSTSLS) estimator (Jerrim et al. 2016).[32] Even in countries in which longitudinal data are available, scholars have often resorted to cross-sectional data to study mobility trends (e.g., Aaronson and Mazumder 2008; Fortin and Lefebvre 1998). It is therefore clear that the averaging strategy could not be used in many countries and that, because of this limitation, resorting to this strategy would hinder the comparative study of mobility across countries and times.

C. Data and variables

We provide here a more detailed description of the data and variables employed in the empirical analyses of Section VII. As the PSID and the NLSY79 are well known and have been widely used in intergenerational-mobility research, we provide comparatively more information about the SOI-M Panel, a new data set first used by Mitnik et al. (2015).

The SOI-M Panel, which was built from U.S. tax returns and other administrative sources, represents all children born between 1972 and 1975 who were living in the U.S. in 1987. The backbone of the SOI-M Panel is the 1987-1996 SOI Family Panel, which is based on a stratified random sample of 1987 tax returns, with a sampling probability that increases with income. All dependents in the 1987 tax returns of the SOI Family Panel born between 1972 and 1975 are included in the SOI-M Panel. In any given year, some individuals are not required to file tax returns because they fall below the filing threshold, with the implication that the SOI Family Panel does not represent all children from the 1972-1975 birth cohorts. To address this

problem, the sample of the SOI-M Panel was augmented by including all children who were (a) born between 1972 and 1975, and (b) listed as dependents in the returns of the "refreshment segment" of the Office of Tax Analysis (OTA) Panel. This segment of the OTA Panel represents those people in the 1987 non-filing population who appeared in a return in at least one year between 1988 and 1996.[33] The resulting sample of children, all of whom were drawn either from the SOI Family Panel or the OTA Panel, was then tracked into 1998-2010 tax returns, W-2 forms, and other administrative sources for income information (see Mitnik et al. 2015 for details on these sources).

Although the SOI-M Panel covers the years 1998-2010, the analyses reported in Section VII of the paper pertain to SOI-M children in 2010 (when they were 35-38 years old). The selected records include information on (a) the annual earnings of children in 2010, and (b) parental income and age when the children were 15-23 years old. Children's earnings are measured as the sum of gross ("Medicare") wages reported in Form W-2 and 65 percent of self-employment income reported in Schedule SE. Parental income is measured as parents' average after-federal-tax income when the children were 15-23 years old. Before computing this average, parental income in each year was expressed in 2010 dollars, employing the Consumer Price Index for Urban Consumers – Research Series (CPI-U-RS). The annual income measure was computed from Form 1040, by summing pre-tax "total income" (which includes labor earnings, capital income, unemployment insurance income, and the taxable portion of pensions, annuities, and social security income) and nontaxable interest, and subtracting out net federal taxes (which include refundable credits). As in Mitnik et al. (2015), children with more than 3 years of missing parental information, nonpositive average parental income, or earnings or average parental income above $7,000,000 were not included in the sample used in our analyses.

From the NLSY79, we selected a sample of children who were 14-16 years old in 1979, thus representing the birth cohorts 1963-1965. We used information on the annual earnings of children in 2003 and 2005 (when they were 38-42 years old), parents' average total family income in 1978-1980, and parents' ages in 1979. Children's earnings are measured as the sum of "total income from wages and salary" and 65 percent of "total income from farm or business," both of which are available in the 2004 and 2006 surveys. The NLSY79 collects this income information using up to six questions (for each respondent) aimed at minimizing the rate of nonresponse. We used all information available (rather than only the income information

provided as exact figures). To measure annual family income in 1978, 1979, and 1980, we used "total net family income" in those years, which is available in the 1979, 1980, and 1981 surveys, respectively. After resorting to Pareto imputation (e.g., Fichtenbaum and Shahidi 1988) to address top-coding, we expressed the 1978-1980 incomes in 2010 dollars using the CPI-U-RS, and averaged them to obtain the parental-income measure employed in our analyses. Parental age was obtained by using information for father's and mother's age available in the surveys for 1979, 1987, and 1988. The sample selection rules required that (a) children had a positive average parental income, and (b) earnings and average parental income were no larger than $900,000 (an upper bound similar to that applied by Lee and Solon 2009). The observations in this sample are children-years.

From the PSID, we selected a sample of household heads and wives (or female cohabitors) that included those who were (a) born between 1954 and 1966, and (b) observed in the PSID at least one year when they were between 35 and 45 years old. The observations are again children-years (i.e., any year a child was in the PSID at those ages is included in the sample). We used information on children's annual earnings at ages 35-45 and on parents' average total family income and ages when the children were 13-17 years old. We measured children's annual earnings via the PSID's "income from labor" variable up to 1992. Starting in 1993, the PSID's measure of "income from labor" changed, as it no longer included the labor portion of business income and 50 percent of farm income. We thus used several variables to reconstruct the pre-1993 PSID notion of "income from labor" in 1993 and thereafter. The earnings information for 1989-2010 was obtained from the surveys for 1990-2011. To measure annual parental income, we used the PSID notion of "total family income" in the years 1967-1983 (available in the 1968-1984 surveys). The income components used to compute total family income were affected by top coding in the period 1970-1978 (i.e., top codes were not only in place but were "binding" in that period for some children). Because the PSID-computed total family income for those years was based on these top-coded values, we proceeded as follows: (a) we addressed top-coding of all income components in 1970-1978 by using Pareto imputation, and (b) we recomputed total family income for those years with these Pareto-imputed variables. After expressing all annual income variables in 2010 dollars using the CPI-U-RS, we averaged them to obtain the five-year parental-income measure employed in our analyses. The sample

selection rules required that (a) children had a positive average parental income, and (b) earnings and average parental income were no larger than $900,000.

There are many important differences among our three datasets (and they are reflected in the results reported in Section VII). The PSID and the NLSY79 data do not cover institutionalized people (i.e., people living in a correctional institution, mental institution, or an institution for the handicapped or poor), whereas the SOI-M data do cover them. The PSID data also do not cover people who are not household heads or spouses of household heads. The SOI-M data codes people working in the informal economy as having zero earnings (i.e., people whose earnings are not reported in Form W-2 and who do not have self-employment income reported in Schedule SE).

We estimate earning elasticities with respect to parental income. Although it is more common to estimate them with respect to fathers' earnings, as Corak (2006:54) and Mazumder (2005:250) point out, there are good reasons to prefer a measure of family income. This is because doing so (a) incorporates the income of mothers and thus better indexes the full complement of resources available to invest in children, (b) reflects the ability of families to draw on other income sources in response to transitory earnings shocks, and (c) avoids any selection bias that may result from omitting sons with absent fathers (as they are likely to be comparatively disadvantaged). This approach, although less common, has of course been often used (e.g., Berman and Taubman 1990; Chadwick and Solon 2002; Levine and Mazumder 2002; Mazumder 2005:250-251).

Notes

[1] This is not to gainsay the value of the rank-rank slope. Although we will introduce a new IGE that allows us to retain the very desirable interpretations long attributed to the conventional IGE, there are of course perfectly legitimate reasons to estimate the rank-rank slope (or other mobility measures) rather than intergenerational elasticities.

[2] A few other analyses have discussed some sources of selection bias in the estimation of the conventional IGE (Couch and Lillard 1998; Minicozzi 2003; Gregg et al. 2016; Mitnik et al. 2015; Drewianka and Mercan n.d.). None of these previous contributions has, however, advanced a formal analysis of how selection bias is generated or traced it back to the joint use of (a) the geometric mean as the measure of central tendency, and (b) the need to rely on short-run proxy measures of income. We do just that in our paper.

[3] We received a pre-publication version of Petersen's (2017) paper only after circulating a draft of our own paper for comments.

[4] We are grateful to Joao Santos Silva for pointing out to us that β_1 is the elasticity of the conditional geometric mean. Although Jäntti and Jenkins (2015:838) note in passing that the IGE "measures the degree of regression to the (geometric) mean in income," they do not pursue the issue further. The parameter β_1 is (also) the IGE of the expectation only when the error term satisfies very special conditions (Santos Silva and Tenreyro 2006; Petersen 2017; Wooldridge 2002:17).

[5] It is possible to find support for either understanding. For instance, a contextual reading of the relevant paragraph in Mulligan (1997:24) seems to favor the second possibility, whereas a literal reading seems to favor the first. It is important, then, to consider both interpretations.

[6] This share interpretation, as well as the two interpretations that pertain to the regression to the geometric mean, are only valid insofar as the differences in parental income are small enough for the difference-in-logarithms approximation to a percent difference to be accurate. This requirement is often ignored in the literature.

[7] In his derivation, Corak (2004:ftn. 1) ignored the error term, writing $Y_{x_2}/Y_{x_1} = (x_2/x_1)^{\beta_1}$. The most plausible interpretation is that he meant our expression in the text, but a literal reading of the sentence we quoted may suggest $E(Y|x_2[Y|x_1]^{-1}) = (x_2[x_1]^{-1})^{\beta_1}$ (i.e., "expected economic advantage" rather than "economic advantage in terms of expectations"). Neither expression can be derived from Equation [1]. A deterministic interpretation, which is another possibility, would of course be completely unfounded.

[8] All arguments hold regardless of whether the geometric mean is undefined or zero.

[9] We simplify our analysis by assuming that, for the parents, a long-run income variable is available.

[10] We will leave implicit the subscript *t* from hereon in.

[11] The datasets in question are the Current Population Survey and the American Community Survey.

[12] Although the IGE of the geometric mean for positive earnings is well defined, for reasons that will become clear later this is rarely, if ever, the parameter of interest.

[13] For Equation [11] to hold with S_x defined as suggested, $E(\ln i|x)$ would have to equal $\ln E(i|x)$, with $i = Y,\ Y_h, Y_w$. This is not the case due to Jensen's inequality. Blanden (2005:9) also noticed that Equation [11] does not apply to random variables, but she suggested that, even though "the decomposition will not be precise," it "nevertheless is a useful concept to keep in mind." Although Equation [11] may indeed play a heuristic role, it does not provide any basis for sound estimation.

[14] This was indeed the rationale that Chetty et al. (2014) provided for turning to the rank-rank slope.

[15] The caveat in note 6 also applies here and in what follows.

[16] The expression is closely related to the alternative interpretation of Corak's (2004:11-12) statements as pertaining to "expected economic advantage" rather than "economic advantage in terms of expectations." See note 7.

[17] By comparison, regression to the mean requires the following:

$$\ln E(Y|x) - \ln E(Y) < \ln x - \ln E(X) \text{ if } x > E(X)$$
$$\ln E(Y|x) - \ln E(Y) > \ln x - \ln E(X) \text{ if } x < E(X).$$

Both here and in Inequalities [17a] and [17b], $\min(X) \leq x \leq \max(X)$.

[18] In referring to a "tail of the distribution," we mean any portion of the distribution (a) between the upper bound of the distribution and a value above (or equal to) the mean, or (b) between the lower bound of the distribution and a value below (or equal to) the mean.

[19] Everywhere else we rely on the weaker assumption of zero conditional mean error. Although the independence assumption is of course strong, researchers prepared to make the Markov assumption may decide in favor of also making it.

[20] Chetty et al. (2014) argue that the conventional IGE is a "person-weighted" IGE that weights all individuals equally, whereas our proposed replacement is a "dollar-weighted" IGE that weights individuals in proportion to their income. See Mitnik (2017d) for a discussion of this characterization and how it falls short.

[21] In Equation [19] and [20], all elasticities are with respect to the child's parental income, while the expectations in Equation [19] are with respect to the distribution of parental income.

[22] Mitnik (2017a) also shows that (a) approximately 13 years of information are needed to eliminate the bulk of attenuation bias, and (b) lifecycle biases tend to vanish when the income measures pertain to parents and children who are approximately 40 years old. These results are similar to those obtained for the OLS estimation of the conventional IGE.

[23] See Mitnik (2017e) on estimating the IGE of the expectation in Stata using the three estimators discussed here. The PPML estimator has also been implemented in many other broadly used statistical packages (e.g., SAS, R, LIMDEP).

[24]Although it's conventional to supplement tax data with earnings reports (for nonfilers), doing so does not solve the problem that those without even earnings reports are often working in the informal economy (and hence don't have zero earnings or income).

[25] If auxiliary data are available, multiple imputation is also possible (see Mitnik 2015:28-29).

[26] The empirical analysis based on the SOI-M dataset was conducted as part of the Joint Statistical Research Program of the Statistics of Income Division of the Internal Revenue Service (see Mitnik et al. 2015).

[27] We use a multiplicative error here for convenience. See Santos Silva and Tenreyro (2006:644) for a discussion of the equivalence between additive- and multiplicative-error formulations of PRFs of this sort.

[28] This paragraph is partially based on a point made by Joao Santos Silva in personal communication.

[29] In order to provide a clear example, we have loosely interpreted the right-hand side of Equation [B1] as a weighted average of finite forward difference quotients.

[30] This argument holds under the GEiV model as long as $\lambda_1 = 1$.

[31] This is the case even in countries, like the U.K., with relatively rich longitudinal data. The two datasets that are used in the U.K. to study economic mobility (e.g., Dearden et al. 1997; Gregg et al. 2016) have at most three waves of data in the needed age range.

[32] As Jerrim et al. (2016) point out, this has proven to be the only feasible approach in Australia, China, France, Japan, Italy, South Africa, Spain and Switzerland. Other countries in the same situation are Argentina (Jiménez and Jiménez 2009), Brazil (Dunn 2007), Chile (Núñez and Miranda 2011), and Ecuador, Nepal, Peru, and Singapore (Grawe 2004).

[33] See Nunns et al. (2008) for detailed information on the SOI Family Panel and the OTA Panel.

References

Aaronson, David and Bhashkar Mazumder. 2008. "Intergenerational Economic Mobility in the US: 1940 to 2000. *Journal of Human Resources* 43(1): 139-172.

Behrman, Jere and Paul Taubman. 1990. "The Intergenerational Correlation between Children's Adult Earnings and their Parents' Income: Results from the Michigan Panel Survey of Income Dynamics." *Review of Income and Wealth*, 36 (2), 115-127.

Bernhardt, Annette, Martina Morris, Mark Handcock, and Marc Scott. 2001. *Divergent Paths. Economic Mobility in the New American Labor Market*. New York: Russell Sage.

Björklund, Anders and Markus Jäntti. 2011. "Intergenerational Income Mobility and the Role of Family Background". *The Oxford Handbook of Economic Inequality*, edited by B. Nolan, W. Salverda and T. Smeeding. Oxford: Oxford University Press.

Black, Sandra and Paul Devereux. 2011. "Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics*, Volume 4b, edited by David Card and Orley Ashenfelter. Amsterdam: Elsevier.

Blanden, Jo. 2005." Intergenerational Mobility and Assortative Mating in the UK." Mimeo.

Böhlmark, Anders and Matthew Lindquist. 2006. "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden," *Journal of Labor Economics*, 24(4):879–896.

Chadwick, Laura and Gary Solon. 2002. "Intergenerational Income Mobility among Daughters." *The American Economic Review* 92(1): 335-344.

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

Corak, Miles. 2004. "Generational Income Mobility in North America and Europe: An Introduction." In *Generational Income Mobility in North America in Europa*, edited by Miles Corak. Cambridge: Cambridge University Press.

Corak, Miles. 2006. "Do Poor Children Become Poor Adults? Lessons from a Cross Country Comparison of Generational Earnings Mobility." Discussion Paper 1993. Bonn: Institute for the Study of Labor.

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

Couch, Kenneth, and Dean Lillard. 1998. "Sample Selection Rules and the Intergenerational Correlation of Earnings." *Labour Economics* 5: 313-329.

Dahl, Molly and Thomas DeLeire. 2008. "The Association between Children's Earnings and Fathers' Lifetime Earnings: Estimates Using Administrative Data." Institute for Research on Poverty Discussion Paper 1342-08, University of Wisconsin-Madison.

Dearden, Lorraine, Stephen Machin and Howard Reed. 1997. "Intergenerational Mobility in Britain." *The Economic Journal* 107(440): 47-66.

Drewianka, Scott, and Murat Mercan. N.D. "Long-term Unemployment and Intergenerational Earnings Mobility." Mimeo.

Dunn, Christopher. 2007. "The Intergenerational Transmission of Lifetime Earnings: Evidence from Brazil," *The B.E. Journal of Economic Analysis & Policy* 7(2): Article 2.

Eide, Eric and Mark Showalter. 2000. "A Note on the Rate of Intergenerational Convergence of Earnings." *Journal of Population Economics* 13:159-162.

Einav, Liran and Jonathan Levin. 2014. "Economics in the Age of Big Data." *Science* 346(6210): 715-721.

Escanciano, Juan Carlos, David Jacho-Chávez, and Arthur Lewbel. 2016. "Identification and Estimation of Semiparametric Two Step Models." *Quantitative Economics* 7: 561-589.

Ermisch, John, Marco Francesconi and Thomas Siedler. 2006. "Intergenerational Mobility and Marital Sorting." *The Economic Journal* 116: 659–679.

Fichtenbaum, Rudy and Hushang Shahidi. 1998. "Truncation Bias and the Measurement of Income Inequality." *Journal of Business & Economic Statistics* 6(3):335-337.

Fortin, Nicole and Sophie Lefebvre. 1998. "Intergenerational Income Mobility in Canada." *In Labour Markets, Social Institutions, and the Future of Canada's Children,* edited by Miles Corak. Ottawa: Statistics Canada.

Gangl, Markus. 2006. "Scar Effects of Unemployment: An assessment of Institutional Complementarities." *American Sociological Review* 71 (6): 986-1013.

Goldberger, Arthur. 1968a. "The Interpretation and Estimation of Cobb-Douglas Functions." *Econometrica* 36(3/4): 464-472.

Goldberger, Arthur. 1968b. *Topics in Regression Analysis*. New York: Macmillian.

Grawe, Nathan. 2004. "Intergenerational Mobility for Whom? The Experience of High- and Low-Earnings Sons in International Perspective." In *Generational Income Mobility in North America in Europa*, edited by Miles Corak. Cambridge: Cambridge University Press.

Gregg, Paul, Lindsey Macmillan, and Claudia Vittori. 2016. "Moving Towards Estimating Sons' Lifetime Intergenerational Economic Mobility in the UK." *Oxford Bulletin of Economics and Statistics* 79(1): 79-100.

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

Hasebe, Takuya and Wim Vijverberg. 2012. "A Flexible Sample Selection Model: A GTL-Copula Approach." IZA Discussion Paper 7003. Bonn.

Heckman, James. 2008. "Selection Bias and Self Selection." In The New Palgrave Dictionary of Economics (Second Edition), edited by Steven Durlauf and Lawrence Blume. New York: Palgrave.

Hertz, Tom. 2005. "Rags, Riches and Race: The Intergenerational Economic Mobility of Black and White Families in the United States." In *Unequal Chances. Family Background and Economic Success*, edited by Samuel Bowles, Herbert Gintis, and Melissa Osborne Groves. New York, Princeton and Oxford: Russell Sage and Princeton University Press.

Hertz, Thomas. 2006. "Understanding Mobility in America." Washington D.C.: Center for American Progress.

Hertz, Tom. 2007. "Trends in the Intergenerational Elasticity of Family Income in the United States." *Industrial Relations* 46(1): 22-50.

Imbens, Guido and Charles Manski. 2004. "Confidence Intervals for Partially Identified Parameters." *Econometrica* 72(6):1845-1857.

Jäntti, Markus and Stephen Jenkins. 2015. “Income mobility.” In *Handbook of Income Distribution*, Volume 2A, edited by Anthony B. Atkinson and François Bourguignon. Elsevier.

Jerrim, John, Álvaro Choi, Rosa Simancas. 2016. “Two-Sample Two-Stage Least Squares (TSTSLS) Estimates of Earnings Mobility: How Consistent are They?” *Survey Research* Methods 10(2): 85-102.

Jiménez, Maribel and Mónica Jiménez. 2009. “La Movilidad Intergeneracional del Ingreso: Evidencia para Argentina.” CEDLAS Working Paper 84. La Plata: Universidad Nacional de la Plata.

Kosanovich, Karen and Eleni Theodossiou Sherman. 2015. “Trends in Long-Term Unemployment.” Washington, DC: Bureau of Labor Statistics.

Lee, Chul-In and Gary Solon. 2009. “Trends in Intergenerational Income Mobility." *The Review of Economics and Statistics* 91 (4): 766-772.

Lee, Lung-Fei. 2003. “Self-Selection.” In *A Companion to Theoretical Econometrics*, edited by Badi Baltagi. Blackwell.

Levine, David and Bhashkar Mazumder. 2002. “Choosing the Right Parents: Changes in the Intergenerational Transmission of Inequality – Between 1980 and the Early 1990s.” Federal Reserve Bank of Chicago Working Paper 2002-08.

Little, Roderick .J.A. and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data* (2nd edition). New York: John Wiley.

Manski, Charles. 1988. *Analog Estimation Methods in Econometrics*. New York: Chapman & Hall.

Manski, Charles. 1989. “Anatomy of the Selection Problem.” *The Journal of Human Resources*, 24(3): 343-360.

Mazumder, Bhashkar. 2001. “The Miss-measurement of Permanent Earnings: New Evidence from Social Security Earnings Data.” Federal Reserve Bank of Chicago Working Paper 2001-24.

Mazumder, Bhashkar. 2005. “Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data.” *The Review of Economics and Statistics* 87(2): 235-255.

Mayer, Susan and Leonard M. Lopoo. 2004. “What Do Trends in the Intergenerational Economic Mobility of Sons and Daughters in the United States Mean?” In *Generational Income Mobility in North America in Europa*, edited by Miles Corak. Cambridge: Cambridge University Press.

Mayer, Susan E., and Leonard Lopoo. 2008. “Government Spending and Intergenerational Mobility.” *Journal of Public Economics* 92: 139-58.

Minicozzi, Alexandra. 2003. “Estimation of Sons' Intergenerational Earnings Mobility in the Presence of Censoring.” *Journal of Applied Econometrics* 18(3): 291-314.

Mitnik, Pablo. 2017a. “Estimating the Intergenerational Elasticity of Expected Income with Short-run Income Measures: A Generalized Error-In-Variables Model.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo. 2017b. “Intergenerational Income Elasticities, Instrumental Variable Estimation, and Bracketing Strategies.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo. 2017c. “Two-Sample Estimation of the Intergenerational Elasticity of Expected Income.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo. 2017d. “‘Person-Weighted’ versus ‘Dollar-Weighted’: A Flawed Characterization of Two Intergenerational Income Elasticities.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo. 2017e. “Estimators of the Intergenerational Elasticity of Expected Income.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo, Victoria Bryant, Michael Weber and David Grusky. 2015. “New Estimates of Intergenerational Mobility Using Administrative Data.” SOI Working Paper, Statistics of Income Division, Internal Revenue Service.

Mitnik, Pablo, Bryant Victoria, and Michael Weber. 2017. “Parental Income and Labor Market Advantages: Upholding the Received View.” Mimeo.

Moffitt, Robert. 1999. “New Developments in Econometric Methods for Labor Market Analysis.” In *Handbook of Labor Economics*, Volume 3, edited by Orley C. Ashenfelter and David Card. Amsterdam: Elsevier.

Mullahy, John. 1997. “Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior.” *Review of Economics and Statistics* 79(4): 586-593.

Mulligan, Casey. 1997. *Parental Priorities and Economic Inequality*. Chicago: The University of Chicago Press.

Ng, Irene. 2007. “Intergenerational Income Mobility in Singapore,” *The B.E. Journal of Economic Analysis & Policy* 7(2): Article 3.

Nicoletti, Cheti and John Ermisch. 2007. “Intergenerational Earnings Mobility: Changes across Cohorts in Britain.” *The B.E. Journal of Economic Analysis & Policy* 7(2): Article 9.

Nunns, James, Deena Ackerman, James Cilke, Julie-Anne Cronin, Janet Holtzblatt, Gillian Hunter, Emily Lin and Janet McCubbin. 2008. “Treasury's Panel Model for Tax Analysis.” Working Paper 3. Washington, D.C., Department of the Treasury.

Núñez, Javier and Leslie Miranda. 2011. “Intergenerational Income and Educational Mobility in Urban Chile.” *Estudios de Economía* 38(1): 195-221.

Nybom, Martin and Jan Stuhler. 2016. “Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation.” *The Journal of Human Resources* 15(1): 239-268.

Petersen, Trond. 2017. “Multiplicative Models for Continuous Dependent Variables: Estimation on Unlogged versus Logged Form.” *Sociological Methodology* 47:113-164.

Samuels, Myra. 1991. “Statistical Reversion Toward the Mean: More Universal Than Regression Toward the Mean.” *The American Statistician* 45(4): 344-346.

Santos Silva, J. M. C. and S. Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88(4): 641-658.

Santos Silva, J. M. C. and Silvana Tenreyro. 2011. “Further Simulation Evidence on the Performance of the Poisson Pseudo-maximum Likelihood Estimator.” *Economics Letters* 112: 220-222.

Smith, Murray. 2003. “Modelling Sample Selection Using Archimedean Copulas.” *Econometrics Journal* 6(1):99-123.

Solon, Gary. 1992. “Intergenerational Income Mobility in the United States.” *American Economic Review* 82: 393-408.

Solon, Gary. 1999. “Intergenerational Mobility in the Labor Market.” *Handbook of Labor Economics*, Volume 3A, edited by Orley C. Ashenfelter and David Card. Amsterdam: Elsevier.

Stuhler, Jan. 2012. “Mobility across Multiple Generations: The Iterated Regression Fallacy.” Discussion Paper 7072. Bonn: Institute for the Study of Labor.

Vella, Francis. 1998. “Estimating Models with Sample Selection Bias: A Survey.” *The Journal of Human Resources* 33(1): 127-169.

Windmeijer, F. A. G. and J. M. C. Santos Silva. 1997. “Endogeneity in Count Data Models: An Applications to Demand for Health Care.” *Journal of Applied Econometrics* 12: 281-294.

Wooldridge, Jeffrey. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass: The MIT Press.

Zimmerman, David J. 1992. “Regression toward Mediocrity in Economic Stature.” *American Economic Review* 82 (3): 409-429.

Table 1: Descriptive Statistics (weighted values)

	SOI-M	NLSY79	PSID
Child's gender (% female)	48.9	48.9	49.7
Child's age			
Mean	36.5	40.1	39.4
Standard deviation	1.1	1.3	3.1
Child's earnings			
Mean	36,547	46,744	52,039
Standard deviation	56,438	53,829	60,000
Average parental age			
Mean	45.4	43.2	43.8
Standard deviation	6.2	6.5	6.8
Average parental income			
Mean	64,183	63,156	88,207
Standard deviation	91,709	42,495	60,468
Number of observations	12,872	5,188	13,564
Number of children	12,872	2,763	2,424

Note: Monetary values in 2010 dollars (adjusted by inflation using the Consumer Price Index for Urban Consumers - Research Series)

Table 2: IGEs of earnings

	SOI-M		NLSY79		PSID	
	Exc. par. age	Inc. par. age	Exc. par. age	Inc. par. age	Exc. par. age	Inc. par. age
Men						
IGE of geometric mean						
Children's earnings > $ 0	**0.35** (0.035)	**0.33** (0.035)	**0.44** (0.047)	**0.43** (0.048)	**0.44** (0.047)	**0.46** (0.049)
Children's earnings ≥ $ 600	**0.34** (0.030)	**0.33** (0.030)	**0.45** (0.044)	**0.44** (0.045)	**0.44** (0.046)	**0.45** (0.048)
Children's earnings ≥ $ 1,500	**0.34** (0.029)	**0.34** (0.030)	**0.45** (0.041)	**0.44** (0.042)	**0.42** (0.045)	**0.43** (0.047)
IGE of expectation	**0.49** (0.028)	**0.46** (0.030)	**0.55** (0.058)	**0.54** (0.059)	**0.50** (0.060)	**0.51** (0.063)
Women						
IGE of geometric mean						
Children's earnings > $ 0	**0.18** (0.035)	**0.17** (0.036)	**0.25** (0.048)	**0.25** (0.048)	**0.26** (0.058)	**0.23** (0.061)
Children's earnings ≥ $ 600	**0.22** (0.029)	**0.21** (0.029)	**0.26** (0.045)	**0.25** (0.045)	**0.26** (0.053)	**0.24** (0.056)
Children's earnings ≥ $ 1,500	**0.21** (0.026)	**0.20** (0.026)	**0.25** (0.039)	**0.25** (0.040)	**0.24** (0.049)	**0.22** (0.053)
IGE of expectation	**0.27** (0.027)	**0.26** (0.028)	**0.29** (0.048)	**0.29** (0.048)	**0.38** (0.063)	**0.37** (0.067)

Note: Point estimates are in bold, standard errors are in parentheses.