Comment: Summarizing income mobility with multiple smooth quantiles instead of parameterized means*

Ian Lundberg† Brandon M. Stewart‡

2 May 2020

Abstract

Studies of economic mobility summarize the distribution of offspring incomes for each level of parent income. Mitnik and Grusky (2020) highlight that the conventional intergenerational elasticity (IGE) targets the geometric mean and propose a parametric strategy for estimating the arithmetic mean. We decompose the IGE and their proposal into two choices: (1) the summary statistic for the conditional distribution and (2) the functional form. These choices lead us to a different strategy—visualizing several quantiles of the offspring income distribution as smooth functions of parent income. Our proposal solves the problems Mitnik and Grusky highlight with geometric means, avoids the sensitivity of arithmetic means to top incomes, and provides more information than is possible with any single number. Our proposal has broader implications: the default summary (the mean) used in many regressions is sensitive to the tail of the distribution in ways that may be substantively undesirable.

Keywords: economic mobility, intergenerational mobility, intergenerational elasticity, stratification

*Manuscript prepared as a comment on Mitnik and Grusky (2020) for *Sociological Methodology*. Replication code is available on GitHub [link] and will be posted upon publication on Dataverse at `https://doi.org/10.7910/DVN/JERJ0C`. For useful comments on earlier versions, we thank Yu Xie, Matthew Salganik, Christopher Felton, and Dalton Conley. We also thank Mike Vuolo for excellent feedback during the editing process. All errors are our own. Research reported in this publication was supported by The Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P2CHD047879. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

†Ph.D. Candidate, Department of Sociology and Office of Population Research, Princeton University, `ianlundberg.org`, ilundberg@princeton.edu

‡Assistant Professor, Department of Sociology and Office of Population Research, Princeton University, `brandonstewart.org`, bms4@princeton.edu. 149 Wallace Hall, Princeton University, Princeton, NJ 08540.

Single-number summaries that capture the relationship of socioeconomic outcomes across generations are a cornerstone of economic mobility research. Studies often focus on the intergenerational elasticity (IGE) of income: the coefficient β_1 on parent log income in a model predicting offspring log income (e.g. Björklund and Jäntti 1997; Solon 2004; Aaronson and Mazumder 2008). A large β_1 is often interpreted as evidence that incomes persist to a substantial degree across generations.

$$\text{Classic approach}: \quad \underbrace{\mathbf{E}\left(\log\left(Y\right) \mid X\right)}_{\substack{\text{Mean of log}\\ \text{offspring income } Y\\ \text{given parent income } X}} = \beta_0 + \underbrace{\beta_1}_{\substack{\text{Intergenerational}\\ \text{elasticity}\\ \text{(IGE)}}} \underbrace{\log\left(X\right)}_{\substack{\text{Log parent}\\ \text{income}}} \tag{1}$$

In their thought-provoking new paper, Mitnik and Grusky (2020, hereafter MG) propose a new summary of economic mobility which they term the intergenerational elasticity *of the expectation*, defined as α_1 in Equation 2.

$$\text{MG proposal}: \quad \underbrace{\log\left(\mathbf{E}\left(Y \mid X\right)\right)}_{\substack{\text{Log of mean}\\ \text{offspring income } Y\\ \text{given parent income } X}} = \alpha_0 + \underbrace{\alpha_1}_{\substack{\text{Intergenerational}\\ \text{elasticity of the}\\ \text{expectation (IGEE)}}} \underbrace{\log\left(X\right)}_{\substack{\text{Log parent}\\ \text{income}}} \tag{2}$$

The central pivot of the MG proposal is to swap the log and the expected value on the left side of Equation 1. This comment unpacks the broader choices represented in the classic IGE and the new MG proposal.

By highlighting issues that have long been ignored, MG open the door to a new and much-needed discussion about how we ought to measure economic mobility. In particular, they identify two central problems with the classic IGE. First, there is the ubiquitous "zeros problem" faced by researchers: when some individuals have income $Y = 0$, the log income of those individuals is negative infinity, and the usual IGE β_1 is not defined.[1] The MG proposal only requires that values be non-zero for the population average income at a given value of parent income, denoted

[1] Given that mobility scholars are often conceptually interested in permanent income averaged over one's entire life, the zeros problem in empirical work arguably arises only due to measurement error since many (perhaps all) people have non-zero lifetime incomes. To highlight other issues, the present comment sets this concern aside and accepts MG's contention that the variables of interest truly have some zero values that are not simply measurement error. We encourage further work that considers the underlying measurement construct and its relationship to the measured outcome.

$\mathbf{E}(Y \mid X)$, thereby resolving the non-applicability of the classical IGE to settings with zeros. Second, MG identify a persistent error: the interpretation of the classic estimand as an arithmetic mean when it captures the geometric mean (Appendix A provides a proof). They argue that their new proposal brings the estimand "into correspondence with the field's prevailing interpretation."

A casual observer might see the proposal—swapping the expected value and the log operator on the left-hand side of Equation 1—as a minor change, but it represents a major shift in the theoretical quantity that the authors target. While we agree that MG provide methods to study mean incomes, in this comment we argue that mean incomes were not generally the optimal target of inquiry to begin with (the informal language of prior research notwithstanding). Our comment is neither a defense of business as usual nor an endorsement of MG's alternative. Instead, we build on the best part of the original contribution—the call for renewed discussion about how to summarize income mobility—and highlight some considerations at play in that choice.

Any summary of an intergenerational association involves two core choices. First, there exists an entire distribution of offspring incomes at each value of parent income (vertical slices represented by gray shapes in Figure 1). It is common to collapse this distribution to one or more *summary statistics* (points in Figure 1), such as the arithmetic or geometric mean. No amount of data can tell a researcher which summary statistic(s) to report—it is a theoretical choice. Because the conditional income distribution is right-skewed, the pivot by MG from the geometric to the arithmetic mean requires a conceptual commitment to a summary that is sensitive to the upper tail. Second, it is common to assume or learn a *functional form* (curves in Figure 1) to approximate the trend in that summary statistic across values of the predictor (parent income). The classic IGE and the MG proposal choose one summary statistic (geometric and arithmetic mean, respectively) and a log-linear functional form (Figure 1). We propose that researchers could provide several summary statistics (quantiles) estimated by a smooth but flexible functional form. Our proposal provides more information about the complete distributions of interest than either the original IGE or the MG proposal. We provide details and an accompanying visualization in Section 3.

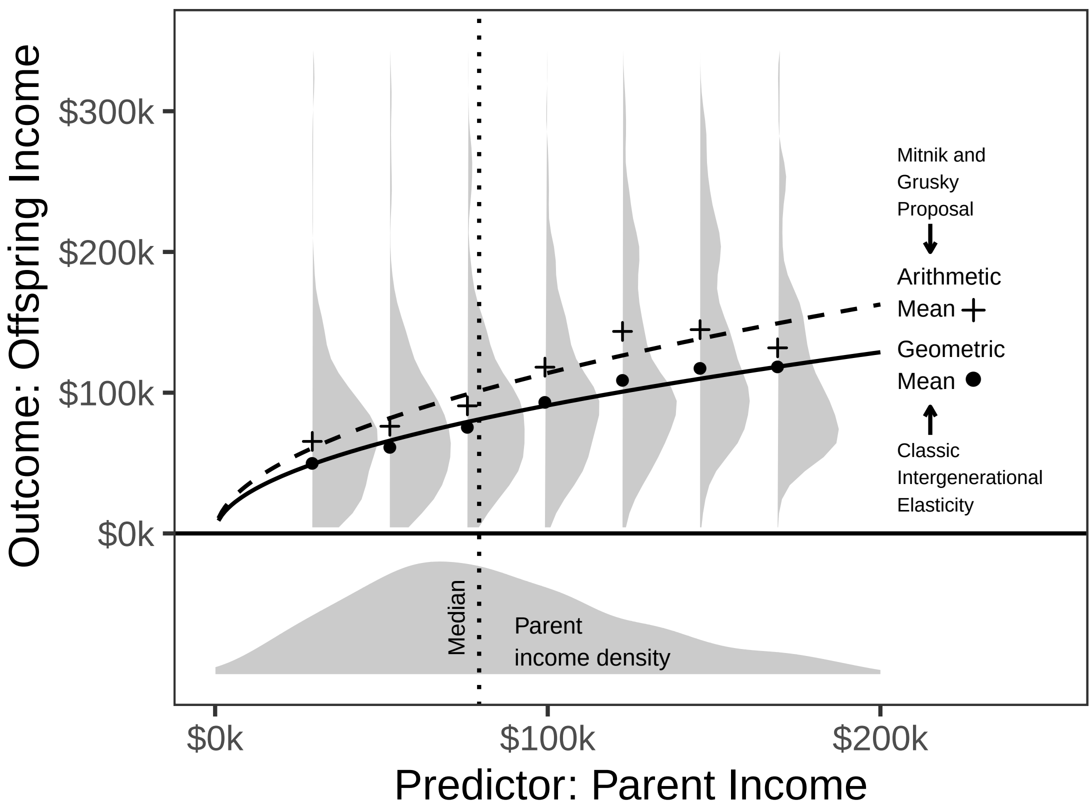

Figure 1: Previous single-number summaries involve parameterized means. Two choices in these approaches sacrifice information in order to arrive at a concise summary. (1) Within each value of parent income (x-axis), previous proposals collapse an entire distribution (vertical gray density) to a summary statistic (point). The MG proposal changes this summary from the geometric mean (•) to the arithmetic mean (+), which tends to be higher because the conditional income distribution is right-skewed. (2) The proposals approximate the trend in the summary across parent incomes x by a log-linear functional form (curve), thus enabling a single coefficient to capture the trend. This choice is unchanged in the MG proposal. Data are from the Panel Study of Income Dynamics (PSID, 2019) and are restricted to those with parent income (the predictor) from \$100 to \$200,000 (details of the analytic sample appear in Appendix D). The points in the graph do not assume a log-linear functional form (choice 2) and are estimated by a smooth curve using the `gam` function in the `mgcv` package in `R` (Wood, 2017). To produce conditional density estimates (vertical gray shapes), we estimate the joint density using the `kde2d` function in the `MASS` package in `R` (Venables and Ripley, 2013) and normalize this density within vertical slices. The y-axis of the plot is truncated at \$350k, which truncates between 0.1% of the leftmost density and 5.5% of the rightmost density.

1 Choice 1: Summary statistics for the conditional distribution

Among those whose parents have a given income level (say \$50k), not all offspring attain the same income levels. There is an entire distribution $(Y \mid X)$ of offspring incomes Y among those with parent income X = \$50k. For a given value of X, a density plot can depict this distribution (see Figure 1, where the conditional distribution is shown for 7 values of X in each panel). Reducing the distribution to a single-number summary, such as the mean or the geometric mean, sacrifices a substantial amount of information to be concise.

For those committed to a single-number summary, the choice of a summary statistic requires conceptual argument about the aspect of the distribution that should be the focus of our attention. The MG proposal directs attention to mean incomes. The mean is a straightforward summary: total all incomes and divide by the number of people. If we add a dollar to total income, the mean increases by the same amount regardless of whose pocket contains the additional dollar. Adding \$1 to the earnings of someone making \$1 million per year changes the mean by the same amount as adding \$1 to the earnings of someone making \$10k per year. In a society where the income distribution includes a few individuals with extremely high earnings, the choice becomes particularly consequential. A decision to summarize the income distribution by the mean is a decision to place substantial importance on the incomes at the top of the distribution.

The geometric mean—the summary implicit in the classic IGE—appears on the surface to be a complicated measure of central tendency: $\exp(\mathrm{E}(\log(Y \mid X)))$. Yet this summary has some desirable properties. First, the log reduces the influence of outliers, so the mean of log income is relatively *insensitive* to differences in top incomes. Instead, mean log income is sensitive to differences at the bottom of the income distribution (especially near 0). This may be desirable if we seek to direct attention to the bottom of the income distribution. Second, the geometric mean equals the median if income is log-normally distributed (we offer a proof in Appendix B). Third, the geometric mean can be justified as capturing average utility when utility is measured as log income. In this setting, the log utility function captures the diminishing marginal utility of income (Appendix C gives a broader discussion of the utility justification).

The arithmetic and geometric means are each valid summaries that emphasize different parts of the distribution. In their reading of the prior literature, MG focus on the use of the words "expected" and "average" to argue that prior authors were interested in characterizing the arithmetic mean of the distribution, and incorrectly summarized the geometric mean. This language notwithstanding, we believe a sound argument can be made in settings with strictly positive incomes that the geometric mean is actually closer to what prior authors have wanted: a summary that is less sensitive to upper-tail outliers, is near the median, and corresponds to a diminishing marginal utility function.

Both arithmetic and geometric means leave a substantial amount of information on the table by reducing the conditional distribution to a single number. We propose to take the core idea of MG—revisiting the workhorse measure of economic mobility—in a more radical direction. Rather than replacing one single summary statistic with another, we propose a series of summary statistics: quantiles of the offspring income distribution. Quantiles provide a more complete depiction of the distribution, thus moving the tradeoff between conciseness and information slightly in the direction of more information. Section 3 returns to this proposal in more detail.

2 Choice 2: Assume or learn a functional form

After defining a summary of the conditional outcome distribution (offspring income given parent income), one can explore the functional form by which that summary changes across the distribution of the predictor variable (parent income). Choice 1 defines the summary of the conditional distribution (the points in Figure 1). Choice 2 defines the curve that passes through the points. This function could be assumed or learned by a flexible smoother.

Traditional approaches and the proposed IGE assume a functional form involving two parameters, thereby allowing a single slope parameter to summarize the dependence of Y on X. In some settings, this approximation is quite good: the flexible smooth point estimates in Figure 1 are close to the parametric curve. Yet it can be difficult to know a priori whether any particular parametric form will be a good approximation. A better solution would be to avoid assuming a functional form and instead estimate the curve via machine learning. We do not discuss this issue in great depth and direct the reader to the literature on flexible smoothers (Wood, 2017). In fact, the authors of the original paper already employ smooth function estimation in other work (Mitnik et al., 2018, 2019).

3 Our proposal: Visualize a sequence of quantiles (choice 1) with a learned functional form (choice 2)

Statistical summaries involve a tradeoff between parsimony and information. The IGE favors parsimony: one number that can be stated in a sentence. The cost is a loss of information: many distinct mobility regimes could have the same IGE. We provide more information through two pivots away from both the classic IGE and the MG proposal. First, instead of relying on a single number (the arithmetic or geometric mean) to summarize the offspring income distribution at each value of parent income, we summarize by a series of quantiles. Quantiles provide a more complete picture of the distribution, are insensitive to outliers, and are well-defined when the outcome can be zero. Second, instead of assuming a log-linear functional form for the pattern of these summaries across values of parent income, we estimate smooth curves using generalized additive quantile regression (Koenker and Bassett, 1978; Fasiolo et al., 2017; Wood, 2017). The information in the resulting summary cannot be reduced to a single-number summary without a loss of information, but it can be conveyed to readers through a visualization. Figure 2 provides a proof of concept.

Our visualization provides a richer depiction of the dynamics of intergenerational mobility than any single-number summary can provide. At a given value of parent income (for example $79k, the median parent income), each line provides the model's prediction of the quantile of the conditional distribution. These each have a direct interpretation as the proportion of offspring below the predicted value. For example, at parent income of $79k, ninety percent of the offspring incomes are below $171k and fifty percent are below $82k. The quantiles reveal patterns hidden

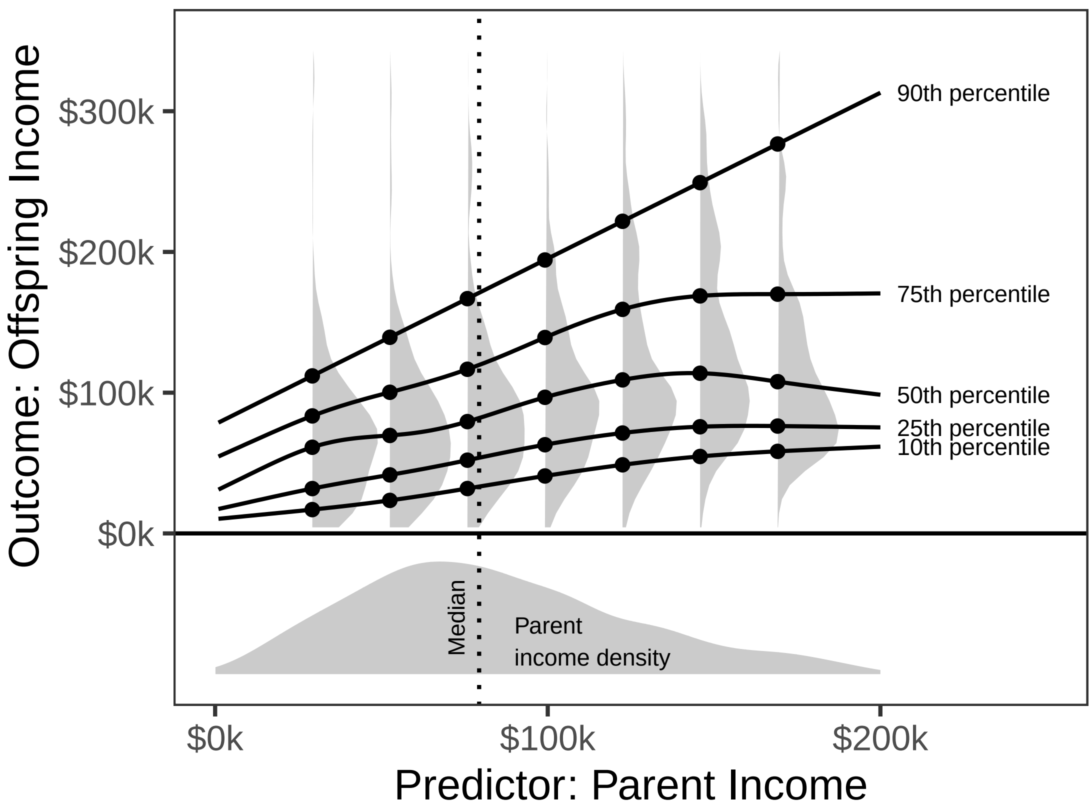

Figure 2: We propose a series of quantiles with smooth functional forms. Our proposal provides more information at the cost that it does not reduce mobility to a single number. Instead of collapsing the full conditional distribution (vertical density) to a single summary such as the mean, a series of quantiles (points) more fully captures the complete distribution of offspring incomes. A flexible smoother removes the need to approximate by a log-linear functional form. This proposal conveys more information than would be possible in a single number. For instance, it reveals that the 90th percentile of offspring incomes increases more across parent incomes than does the 10th percentile. If the conditional distributions are lognormal, then the 50th percentile is equivalent to the geometric mean, the summary at the core of the classic intergenerational elasticity (Appendix B). Data are from the Panel Study of Income Dynamics (PSID, 2019) and are restricted to those with parent income (the predictor) from $100 to $200,000 (Appendix D). We estimate the curves by the `mqgam` function in the `qgam` package in `R` (Fasiolo et al., 2017). To produce conditional density estimates (vertical gray shapes), we estimate the joint density using the `kde2d` function in the `MASS` package in `R` (Venables and Ripley, 2013) and normalize this density within vertical slices. The y-axis of the plot is truncated at $350k, which truncates between 0.1% of the leftmost density and 5.5% of the rightmost density.

by single number summaries. The steeper line for the 90th percentile indicates that the probability of having a particularly high offspring income is quite sensitive to small changes in parent income. Parents of all incomes might have offspring that attain low incomes, but only high-income parents have a large proportion of extremely high-income offspring.

In producing Figure 2, we faced a tradeoff between visual simplicity and the amount of information conveyed. We have criticized single-number summaries, but five-quantile summaries still sacrifice potentially relevant information. For example, the figure does not provide information about how the 99th percentile of offspring income relates to parent income. At the same time, the figure certainly provides more information than would be available in any single number alone. The optimal balance will differ based on the argument the author is making, and we propose Figure 2 as one example of a general class of possible visualizations that may strike a useful balance in this tradeoff.

A key selling point of the conventional IGE and the MG proposal is that they offer single-number summaries. These can be helpful for making comparisons across countries or time. If one wants such a summary and is comfortable with the ensuing information loss, it is possible to generate one by summarizing the fitted smooth curve. For instance, we can produce a first difference for (e.g.) the 50th percentile: (1) calculate the 50th percentile of offspring incomes at each observed parent income, (2) calculate the 50th percentile when $10k is added to each parent income, and then (3) report the population-average difference between (2) and (1). Median offspring income is (on average) $4k higher when parent income is $10k higher. Conducting the same first difference at the 90th percentile produces a higher first difference of $12k, reflecting the steeper slope of the 90th percentile curve. Converting a smooth curve to a one-number summary necessarily loses information; just as many distinct conditional distributions could collapse to the same single-number summary (choice 1), likewise many distinct curves could collapse to the same single-number summary of the slope (choice 2).

4 Conclusion

Economic mobility is a topic that calls for transparent and informative summaries to guide scientific understanding, public opinion, and policy. MG make an important contribution by reshaping our collective understanding of the standard indicator—the intergenerational elasticity—and providing a much-needed call for discussion of new indicators. This comment illustrates that the indicator proposed by MG shifts attention away from the bottom of the conditional income distribution and toward the top of the distribution. There is nothing inherently wrong with this shift, but it runs counter to the longstanding tendency of those who study incomes to avoid sensitivity to the upper tail. If an applied researcher would hesitate to summarize a skewed income distribution by its mean, then that applied researcher should also hesitate to use the MG proposal. Neither the IGE nor the MG proposal is better than the other—they are just different summaries. All single-number summaries sacrifice information about economic mobility in favor of parsimony. Our strongest recommendation is that researchers make this tradeoff consciously. By visualizing a series of offspring income quantiles as smooth functions of parent income, researchers can bring back some of the richness available in the data under minimal assumptions while maintaining enough simplicity for a reader to understand. Quantiles address the concerns raised by the MG proposal—they solve the "zeros problem" and estimate a known, interpretable quantity—while still maintaining a commitment to summarize the distribution in ways that are not overwhelmed by the skew common in conditional income distributions.

While our comment has primarily focused on the problem of summarizing economic mobility, our points are broader. Any regression analysis summarizes the distribution of Y given values of X with a single summary statistic such as a mean. The two choices we raise—distributional summary and functional form—are applicable in any regression context. Researchers often equate the research goal with the coefficient of a regression model but we advocate a more conscious choice of estimand (Lundberg et al., 2020). Research constrained to the study of parameterized means risks obscuring important sources of evidence. We would encourage the use of a broader set of tools in other contexts. In the study of mobility, alternate choices are particularly tractable because it is a bivariate setting. They are also particularly necessary because the outcome variable is so severely skewed.

References

Aaronson, D. and B. Mazumder 2008. Intergenerational economic mobility in the United States, 1940 to 2000. *Journal of Human Resources*, 43(1):139–172.

Björklund, A. and M. Jäntti 1997. Intergenerational income mobility in Sweden compared to the United States. *The American Economic Review*, 87(5):1009–1018.

Fasiolo, M., Y. Goude, R. Nedellec, and S. N. Wood 2017. Fast calibrated additive quantile regression. *arXiv preprint arXiv:1707.03307*.

Koenker, R. and G. Bassett 1978. Regression quantiles. *Econometrica*, 46(1):33–50.

Lundberg, I., R. Johnson, and B. M. Stewart 2020. Setting the target: Precise estimands and the gap between theory and empirics. *SocArXiv Preprint*.

Mitnik, P., V. Bryant, and D. Grusky 2018. A very uneven playing field: Economic mobility in the United States. *Stanford Center on Poverty and Inequality Working Paper*.

Mitnik, P. and D. Grusky 2020. The intergenerational elasticity of what? The case for redefining the workhorse measure of economic mobility. *Sociological Methodology*, 50:XX–XX.

Mitnik, P. A., V. Bryant, and M. Weber 2019. The intergenerational transmission of family-income advantages in the United States. *Sociological Science*, 6:380–415.

PSID 2019. *Panel Study of Income Dynamics, Public Use Dataset*. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.

Solon, G. 2004. A model of intergenerational mobility variation over time and place. In *Generational Income Mobility in North America and Europe*, M. Corak, ed., Pp. 38–47. Cambridge.

Venables, W. N. and B. D. Ripley 2013. *Modern Applied Statistics with S-PLUS*. Springer Science & Business Media.

Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.

APPENDIX

A The IGE summarizes offspring income by the geometric mean

MG show that the usual IGE can be interpreted with respect to the geometric mean of offspring income given parent income.[2] To make it possible to read this comment independently of the original paper, we show the short proof of this fact below, using $i = 1, \ldots, N$ to index the full population. First, we show that the exponentiated mean of the log equals the geometric mean.

$$\exp(\mathbf{E}(\log(Y))) = \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(Y_i)\right) \quad (3)$$

$$= \exp\left(\frac{1}{N}\log(\prod_{i=1}^{N}Y_i)\right) \quad (4)$$

$$= \exp\left(\log\left(\left(\prod_{i=1}^{N}Y_i\right)^{\frac{1}{N}}\right)\right) \quad (5)$$

$$= \left(\prod_{i=1}^{N}Y_i\right)^{\frac{1}{N}} \quad (6)$$

$$= \text{GM}(Y) \quad (7)$$

By exponentiating both sides of the IGE proposal, we show that it assumes a functional form for the geometric mean of offspring income Y given parent income X.

$$\mathbf{E}\left(\log(Y \mid X)\right) = \beta_0 + \beta_1\log(X) \quad (8)$$

$$\exp\left(\mathbf{E}\left(\log(Y \mid X)\right)\right) = e^{\beta_0}X^{\beta_1} \quad (9)$$

$$\text{GM}(Y \mid X) = e^{\beta_0}X^{\beta_1} \quad (10)$$

The IGE therefore corresponds to a particular functional form to summarize the geometric mean of offspring income at each value of parent income.

[2]As acknowledged by Mitnik and Grusky (2020), previous authors have highlighted the geometric mean interpretation (e.g. Manning 1998, Petersen 2017), but this statistical result has largely been ignored in the mobility literature.

B The usual IGE is interpretable in terms of medians if the conditional distribution of the log outcome is symmetric

Suppose the distribution of $\log(Y) \mid X$ is symmetric in the sense that it has finite mean equal to its median. This is true, for instance, if the conditional distribution of $Y \mid X$ is lognormal. In that case, the usual IGE has an interpretation with respect to the median of the distribution.

$$\text{E}\left(\log\left(Y\right) \mid X\right) = \beta_0 + \beta_1 \log\left(X\right) \tag{11}$$

Median equals mean if distribution is symmetric

$$\text{Median}\left(\log\left(Y\right) \mid X\right) = \beta_0 + \beta_1 \log\left(X\right) \tag{12}$$

Median is invariant to monotone transformations

$$\log\left(\text{Median}\left(Y \mid X\right)\right) = \beta_0 + \beta_1 \log\left(X\right) \tag{13}$$

$$\text{Median}\left(Y \mid X\right) = \exp\left(\beta_0 + \beta_1 \log\left(X\right)\right) \tag{14}$$

$$= \exp(\beta_0) X^{\beta_1} \tag{15}$$

To directly interpret β_1 requires first-order Taylor approximations.

$$\frac{\text{Median}(Y \mid X = (1+\delta)x)}{\text{Median}(Y \mid X = x)} = \left(\frac{(1+\delta)x}{x}\right)^{\beta_1} \qquad \text{by Equation 15} \tag{16}$$

$$= (1+\delta)^{\beta_1} \tag{17}$$

$$\approx 1 + \beta_1 \log(1+\delta) \qquad \text{by approximation for } \beta_1 \approx 0 \tag{18}$$

$$\approx 1 + \beta_1 \delta \qquad \text{by approximation for } \delta \approx 0 \tag{19}$$

This shows that multiplying X by a factor of $1+\delta$ is associated with an increase in the conditional median of Y by a factor of $1+\beta_1\delta$. In the percentage language used in prior research, a δ = 1% increase in X is associated with a β_1% increase in the conditional median of Y.

The interpretations of prior research are thus valid as long as (1) log income follows a symmetric (e.g. normal) distribution and (2) we replace all references to means with references to medians in prior interpretations. Nonetheless, future researchers might be advised to directly estimate conditional medians via quantile regression if that is the goal (Koenker and Bassett, 1978), rather than relying on a distributional approximation.

C Utility functions aid reasoning about means of transformed income

Both the classic IGE and the MG proposal summarize offspring incomes by a mean of some transformation of income. This section presents utility functions as one way to reason about a mean-based summary of a distribution.[3] A utility function explicitly transforms a distribution of offspring incomes $Y \mid X$ into a single-number summary of how "good" that distribution is. To do so, it transforms each person's income into some assumed utility U_i, such as income $U_i = Y_i$ or log income $U_i = \log(Y_i)$. Then, the mean sums each person-specific U_i over the population and divides by the number of people. Utilities are a natural motivation for means because of their long history as something we would like to maximize when summed over a population.

The choice of a utility function (and by extension a summary statistic) is a choice that places value on different parts of the income distribution. A linear utility function equal to income $U_i = Y_i$ involves a commitment to value an extra dollar in anyone's pocket equally. The log, as a concave increasing utility function $U_i = \log(Y_i)$, values a dollar in the pocket of a low earner more than a dollar added to the pocket of a high earner. An increase in income from $1k to $2k is extremely beneficial, whereas an increase from $101k to $102k is less so. If we accept log income as a good measure of utility or well-being, then expected log income may be a reasonable choice as an indicator of the utility one can expect given the income of one's parents.

The utility perspective illustrates how the log is a terrible choice of utility function in one

[3] A utility perspective that problematizes untransformed arithmetic means appears as early as Bernoulli ([1738] 1954). Muliere and Parmigiani (1993) provide an intellecutal history. The utility perspective continues to inform ongoing economic research (e.g. Gorajek 2019).

setting: when some individuals' incomes are zero. As highlighted by MG, the mean of log incomes is undefined (or is negative infinity) if anyone in the population has an income of zero. This is because a log utility implies a belief that utility drops an infinite amount as income drops to zero. The first penny earned adds an infinite amount to utility. The log is therefore a nonsensical utility function when some incomes are zero.

The problem with the MG proposal is that it replaces one utility function (log) with another utility function (linear) that is equally indefensible. Summarizing utility by the mean of untransformed income solves the zeros problem but sacrifices the concave shape of the log utility function. By giving up on the idea that utility increases at a decreasing rate with income, the proposal ignores the key reason for using the log (diminishing returns) to address the indefensible behavior at zero. We emphasize that a linear utility is not the only alternative to the log (Figure C1). The inverse hyperbolic sine function, $\log(\text{Income} + \sqrt{\text{Income}^2 + 1})$, is a well-known option that is similar to the log except that it is well-defined at incomes of zero (Johnson, 1949; Burbidge et al., 1988). The inverse hyperbolic sine still assumes that utility increases rapidly at low incomes and slowly after that. If one believed that the slope of utility as a function of income was somewhat closer to linear, one could define a different transformation such as $\log(\text{Income} + \$5,000)$ which would exhibit this behavior. Overall, the conclusions one draws by summarizing the mean of some transformation of income are likely to differ by the particular transformation one assumes; we simply note that a linear transformation is not the only alternative when the data include zeros.

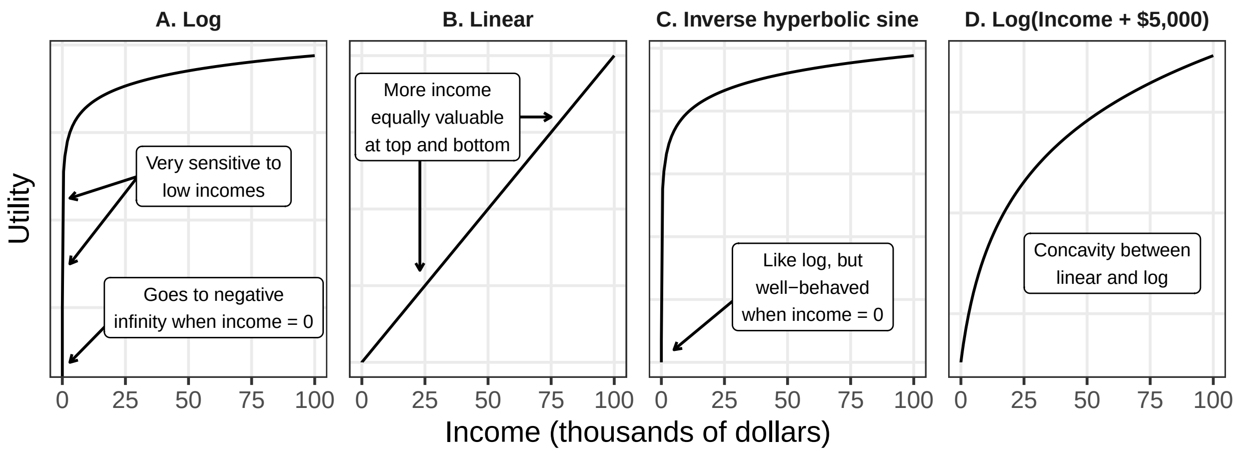

Figure C1: The mean of transformed income can be interpreted as expected utility. A transformation (e.g. log) represents a choice about the utility of offspring income which researchers could defend on theoretical grounds. The plot shows the shape of various choices. The scale of the y-axis differs across the plots. The log (A) assumes well-being increases at a decreasing rate with income, but it can be very sensitive to low values because it assumes utility drops to $-\infty$ at zero income. The MG proposal replaces the log with a linear utility function (B): each additional dollar adds an equal amount to well-being. The inverse hyperbolic sine (C, $\log(\text{Income} + \sqrt{\text{Income}^2 + 1})$) is similar to the log but is well-behaved when income is 0 (taking the value 0 instead of $-\infty$). Finally, one could define some new transformation (D) such as $\log(\text{Income} + \$5,000)$, which would maintain the concavity of the log but would be well-defined at zero and would exhibit less concavity than the log. The amount of concavity in D would depend on the value of the constant added to income. Researchers choosing among these options would need to consider substantive aspects of the application, and results may be sensitive to different choices. The MG proposal (B) is not the only way to solve the zeros problem in (A), which could be solved by alternative concave utilities such as (C) and (D).

D Details of the analytic sample

Our empirical illustration draws on data from the Survey Research Center (SRC) sample of the Panel Study of Income Dynamics (PSID, 2019). The study drew an equal-probability sample of U.S. households in 1968 and followed all sample members and their descendants over subsequent generations. We restrict to those born in 1954–1972 who were observed at least once at ages 13–17 (990 persons) and at least once at ages 35–45 (700 persons). Parent family income is measured as total family money income (including transfers), averaged over all observations when the focal individual was ages 13–17 (mean = 3.6 observations). Offspring family income is the same variable averaged over all observations when the focal individual was ages 35–45 (mean = 5.0 observations). To direct attention to the region of the predictor variable that is densely populated on both the linear and the log scales, we restrict to those with parent family incomes no greater than $200,000 (667 persons) and at least $100 (664 persons). This restriction is generous to the IGE and the MG proposals because it reduces the risk that a global log-linear parametric form could be heavily influenced by a few outlying values of the predictor variable. All estimates in the paper refer to the analytic sample of 664 persons. Because the SRC is an equal-probability sample, all analyses are unweighted and require an assumption of ignorable sample attrition for population inference.

Appendix references

Bernoulli, D. [1738] 1954. Exposition of a new theory on the measurement of risk. Translated by L. Sommer. *Econometrica*, 22(1):23–36.

Burbidge, J. B., L. Magee, and A. L. Robb 1988. Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401):123–127.

Gorajek, A. 2019. The well-meaning economist. *Reserve Bank of Australia Research Discussion Paper 2019-08*.

Johnson, N. L. 1949. Systems of frequency curves generated by methods of translation. *Biometrika*, 36(1):149–176.

Koenker, R. and G. Bassett 1978. Regression quantiles. *Econometrica*, 46(1):33–50.

Manning, W. G. 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics*, 17(3):283–295.

Mitnik, P. and D. Grusky 2020. The intergenerational elasticity of what? The case for redefining the workhorse measure of economic mobility. *Sociological Methodology*, 50:XX–XX.

Muliere, P. and G. Parmigiani 1993. Utility and means in the 1930s. *Statistical Science*, 8(4):421–432.

Petersen, T. 2017. Multiplicative models for continuous dependent variables: Estimation on unlogged versus logged form. *Sociological Methodology*, 47:113–164.

PSID 2019. *Panel Study of Income Dynamics, Public Use Dataset*. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.