American Economic Journal: Applied Economics 2024, 16(2): 66–98
https://doi.org/10.1257/app.20210686

Intergenerational Mobility in India: New Measures and Estimates across Time and Social Groups†

By Sam Asher, Paul Novosad, and Charlie Rafkin*

We study intergenerational mobility in India. We propose a new measure of upward mobility: the expected education rank of a child born to parents in the bottom half of the education distribution. This measure works well under data constraints common in developing countries and historical contexts. Intergenerational mobility in India has been constant and low since before liberalization. Among sons, we observe rising mobility for Scheduled Castes and declining mobility among Muslims. Daughters' intergenerational mobility is lower than sons', with less cross-group variation over time. A natural experiment suggests that affirmative action for Scheduled Castes has substantially improved their mobility. (*JEL* I24, J13, J15, J16, J62, O12, Z12)

There are two widely held narratives regarding access to opportunity in India. On the one hand, economic liberalization, rapid economic growth, and urbanization have vastly expanded the set of opportunities available to Indians, leading to the emergence of a large middle class. The political sphere has also opened, with the growth of a wide range of parties organized around caste, region, and ideology. Decades of affirmative action programs have targeted government benefits at historically disadvantaged groups. On the other hand, some of India's entrenched inequalities seem as persistent as ever. Marriage across religious, caste, and class lines is exceedingly rare. Elites in business, government, and civil society still largely come from the upper classes and castes. Inequality has risen, and religious cleavages may be deepening (Chancel and Piketty 2019). In this paper, we shed light on changing access to opportunity in India by studying the intergenerational transmission of economic status (Solon 1999; Black and Devereux 2011; Chetty et al. 2014a; Chetty

*Asher: Imperial College London (email: sam.asher@imperial.ac.uk); Novosad: Dartmouth College (email: paul.novosad@dartmouth.edu), corresponding author; Rafkin: Massachusetts Institute of Technology (email: crafkin@mit.edu). Seema Jayachandran was coeditor for this article. We are thankful for useful discussions with Alberto Abadie, David Autor, Emily Blanchard, Raj Chetty, Eric Edmonds, Shahe Emran, Francisco Ferreira, Amy Finkelstein, Nate Hilger, Larry Katz, David Laibson, Ethan Ligon, Erzo Luttmer, Whitney Newey, Elias Papaioannou, Nina Pavcnik, Bruce Sacerdote, Frank Schilbach, Na'ama Shenhav, Forhad Shilpi, Andrei Shleifer, Gary Solon, Bob Staiger, Doug Staiger, Chris Snyder, and Elie Tamer, among others. Annaka Balch, Ali Campion, Toby Lunt, Ryu Matsuura, and Taewan Roh provided excellent research assistance. This project received financial support from the IZA GLM-LIC program. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1122374. This paper contains some material previously contained in the retired paper "Getting Signal from Interval Data: Theory and Applications to Mortality and Intergenerational Mobility." An earlier version of this paper had the subtitle "Across Time, Space, and Communities."

†Go to https://doi.org/10.1257/app.20210686 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

et al. 2020), with a particular emphasis on the 478 million people (39 percent) who belong to India's major disadvantaged groups (Muslims, Scheduled Castes, and Scheduled Tribes).

We seek to measure the persistence of socioeconomic *rank* across generations, isolating intergenerational mobility from changes in inequality and growth, in the spirit of Solon (1999). Recent advances in the literature on intergenerational mobility have yielded a set of measures based on the conditional expectation function—henceforth, the CEF—of a child's socioeconomic rank, given a parent's rank. The most notable of these is *absolute upward mobility*, which is the income rank of a child born to a parent at the twenty-fifth percentile (Chetty et al. 2014a). These measures have several desirable features, two of which we highlight: (i) by focusing on ranks, they isolate mobility from economic growth and inequality, which is crucial in a rapidly growing country like India; (ii) they are directly interpretable as mobility measures for population subgroups, like Scheduled Castes.[1]

We cannot apply these existing measures to study intergenerational mobility in India. Linked parent-child income data are unavailable in India, as in many developing countries. If it were available, such data would be difficult to interpret in multigenerational households with joint production. For these reasons, studies on intergenerational mobility in developing countries and historical contexts have often used education as the primary measure of social status.[2] We highlight and then resolve a key challenge that arises in the use of CEF-based mobility measures with education data: education is reported in bins—often very large bins. Among parents of the 1950–1969 generations in India, more than 50 percent of fathers (and 80 percent of mothers) have bottom-coded levels of education; identifying a parent at the twenty-fifth percentile (or any other) is fraught, and naive approaches can produce incorrect results.

We gain traction on this problem by recasting it as an interval censoring problem, such that the binned education data reflect a continuous (but unobserved) latent rank distribution.[3] This framing of the problem lets us draw on tools from the partial identification literature to put bounds on the parent-child rank CEF (Manski and Tamer 2002; Novosad, Rafkin, and Asher 2022). We introduce a new measure of upward mobility, *bottom-half mobility*, which is the expected rank of a child born to a parent in the bottom half of the education distribution. Bottom-half mobility has a very similar interpretation to absolute upward mobility, but it can be bounded tightly even in contexts with severe interval censoring. Our approach further reveals that, once interval censoring is taken into account, other canonical measures like

[1] Section II explains why the earlier canonical class of linear measures, such as the regression coefficient from a regression of child education on parent education, is difficult to interpret for population subgroups.

[2] Intergenerational educational mobility is also directly of interest; canonical intergenerational mobility models like Loury (1981); Becker and Tomes (1986); and Galor and Zeira (1993) often emphasize the role of human capital investment. Recent studies of intergenerational mobility focusing on education include Black et al. (2005); Güell et al. (2015); Wantchekon et al. (2015); Card et al. (2022); Alesina et al. (2021); and Eshaghnia et al. (2022). More are summarized in Black and Devereux (2011).

[3] Section IIB discusses this assumption in detail and sketches out how it arises directly from a standard human capital model.

absolute upward mobility and the rank-rank gradient have bounds that are too wide to be meaningful.[4]

To study upward mobility in India, we use data from the 2012 India Human Development Survey (IHDS). A key feature of this dataset is that individuals were asked about their parents' education level. We thus observe parent-child links even when parents and children live in different households, and we can measure changes in mobility using older cohorts. We document trends in educational mobility from the 1950–1959 to the 1985–1989 birth cohorts. We focus on measuring mobility from fathers to sons and daughters; mobility from mothers to children cannot be bounded tightly because the bottom coding of mothers' education is so severe.

We present two main results. First, upward mobility has remained constant over the entire sample period, despite dramatic gains in average levels of education and income. An Indian son born in the bottom half of the parent education distribution in 1985–1989 (our youngest cohort) can expect to obtain percentile 38; a daughter obtains percentile 36.[5] A similar child in the United States, which has low intergenerational mobility by OECD standards, on average attains education percentile 42.[6] This suggests that India's decades of economic growth have lifted the living standards of individuals in the bottom half of the socioeconomic distribution without substantially changing their likelihood of moving to a higher socioeconomic *rank*. By contrast, a naive application of the canonical rank-rank gradient measure to our data would suggest that mobility had improved significantly.

Second, we show that upward mobility has changed substantially over time for some social groups, particularly among sons. We divide the population into Scheduled Castes (SCs), Scheduled Tribes (STs), Muslims, and Forwards/Others. Consistent with prior work (Hnatkovska et al. 2012; Emran and Shilpi 2015), we find that sons from India's constitutionally protected marginalized groups, the SCs and STs, have closed respectively 50 percent and 30 percent of the mobility gap with Forwards/Others. In contrast, upward mobility for Muslim sons has steadily declined from the 1960s to the present. The expected educational rank of a Muslim man born in the bottom half of the parent distribution has fallen from between percentiles 31 and 34 to a dismal 29.[7] Muslim sons have considerably worse upward mobility today than both SCs (38) and STs (33), a striking finding given that, compared to Muslims, STs tend to live in more rural and remote areas. Higher caste groups have experienced constant and high upward mobility over time. This result

[4] Naive application of these tools that ignore the censoring problem leads to results that are misleadingly precise, because they ignore the loss of information associated with the coarse measurement of education; they may also be biased (see Section IV).

[5] We compute educational attainment ranks for sons and daughters separately because men and women mostly compete in separate labor markets, and education plays a different role in determining their life outcomes.

[6] In a society where children's outcomes are independent of parents (i.e., perfect mobility), a child born in the bottom half of the distribution obtains the fiftieth percentile on average. In a society with no upward mobility, i.e., where all children obtain the same percentile as their parents, the same child attains the twenty-fifth percentile.

[7] The comparable figure for US Black men is 35.

contradicts a popular notion that it is increasingly difficult for higher-caste Hindus to get ahead.[8]

Our measures for father-daughter mobility are less precise, but the subgroup patterns appear to be different. Daughters from poor Muslim, SC, and ST households all have persistently lower mobility than Forwards/Others. There is minimal convergence over the sample period; the bounds on changes are consistent with women's mobility being either stable or declining slightly.

The final section of the paper examines several mechanisms for the divergence of SCs from Muslims over the last 30 years. We show that this divergence cannot be explained by differential returns to education, occupational patterns, geography, or differential fertility. However, we find suggestive evidence that the basket of affirmative action policies targeted to India's scheduled groups (but not to Muslims) has played a key role in their rising mobility. Following Cassan (2019), we study a natural experiment that added many castes to the SC lists in 1977 following a rules-based mechanism. We show that caste groups that were newly assigned to SC status on average experienced a 7–8 rank point increase in upward mobility over the next twenty years. This is the same size as the rank mobility gap that has opened between Muslims and SCs over the same period. Our findings are thus consistent with the possibility that educational quotas, government job reservations, and other affirmative action policies are fully responsible for the upward mobility gap between SCs and Muslims. However, because we are limited to birth cohort $\times$ demographic group variation and only a small sample of individuals whose SC status changed ($N = 696$), these results are suggestive and not dispositive.

This paper makes both methodological and empirical contributions. Methodologically, we show how to calculate CEF-based mobility measures using education data; our proposed measure, bottom-half mobility, may be useful in many contexts beyond India. To our knowledge, the bias caused by ignoring interval censoring in education data has not been recognized by prior work on intergenerational educational mobility. We describe the relationship between our measure and others in the literature in Section IIC.

Empirically, we present several previously unknown facts about upward mobility in India. Our findings imply that virtually all of the upward mobility gains in India over recent decades have accrued to SCs and STs. These are both groups who have constitutional protections and reservations in politics and education, and who have been specifically targeted by development policies. There is no evidence that any of these gains have come at the expense of higher-caste groups. On the other hand, mobility has declined for Muslims. We are not aware of studies of intergenerational mobility for Indian Muslims, even though they number almost 200 million people (similar to the number of people in SCs).[9]

[8] The sample size of the IHDS is too small to study geographic variation in any detail. For a preliminary analysis of the microgeographic variation in upward mobility using individuals aged 20–23 in the high resolution 2011–2012 Socioeconomic and Caste Census (SECC), please see the working paper version of this manuscript (Asher et al. 2021).

[9] Other economics papers on Indian Muslims include Khamis et al. (2012) and Bhalotra and Zamora (2010), who note poor education outcomes among Muslims. The Sachar Committee Report (2006) and Basant and Shariff (2010) summarize some recent research on Muslims on India, none of which addresses intergenerational mobility.

Our analysis of bottom-half mobility has several limitations. First, by averaging results in the bottom half of the distribution, bottom-half mobility may conceal changes within the bottom half. We cannot rule out the possibility that outcomes for the bottom 10 percent have improved substantially, while, for example, outcomes for percentiles 40 through 50 have declined.[10] Going forward, we use the terms "upward mobility" and "bottom-half mobility" interchangeably; but our results are by no means conclusive about upward mobility from the very bottom of the distribution.

Second, in principle, the observed divergence between SCs and Muslims could mechanically arise if Muslim parents are increasingly concentrated among the lowest latent ranks in the bottom half (which we would not observe directly). We present a host of tests that reject such an explanation; similar tests would be advised when using our method to study population subgroups in other contexts.

Our estimates contribute to a growing literature on intergenerational mobility.[11] Post-independence India is an important setting for studying upward mobility, both because it has undergone rapid economic change, and because of the history of the caste system, which has been thought to entrench mobility across generations. Our finding that India's growth episode had little effect on rank mobility recalls other recent studies that have found intergenerational persistence in the context of substantial social upheaval (Ager, Boustan, and Eriksson 2021; Alesina et al. 2020a). Prior work on India has: (i) emphasized absolute outcomes (such as consumption), which are rising for all groups due to India's substantial economic growth (Maitra and Sharma 2009; Hnatkovska et al. 2013); or (ii) compared subgroups using the parent-child outcome correlation or regression coefficient, which describes the outcomes of subgroup members relative to their own group, rather than to the national population (Hnatkovska et al. 2013; Emran and Shilpi 2015; Azam and Bhatt 2015).[12] Studies of affirmative action in India have found impacts on educational attainment of SC/STs (Frisancho Robles and Krishna 2016; Bagde et al. 2016; Cassan 2019; Khanna 2020), but have not examined intergenerational mobility.

More broadly, our paper relates to research on religion and economic development (McCleary and Barro 2006; Becker and Woessmann 2009). Alesina et al. (2020b) and Platas (2018) find low mobility and educational outcomes for Muslims in sub-Saharan Africa, where Islam plays a different cultural, political, and social role from in India.[13]

Stata code to calculate our mobility measures and replicate this paper is posted online.[14]

[10] Testing for changes within the bottom half requires different data on parents, since the coarse education bins in the data mean that information on relative parent status cannot be inferred from parent education.

[11] For review papers on intergenerational mobility, see Corak (2013), Black and Devereux (2011), and Roemer (2016). Recent work on intergenerational mobility has described the United States (Chetty et al. 2014b; Chetty et al. 2014a; Chetty et al. 2020), western Europe (Bratberg et al. 2017), and Africa (Alesina et al. 2021), among many other regions.

[12] Note that there is a parallel literature examining the persistence of income within an individual lifetime in India; this is sometimes described as *income or economic mobility*. That literature is focused largely on measurement error in income over the course of an individual's lifetime and is thus not directly related to our work (Azam 2016; Li, Millimet, and Roychowdhury 2019).

[13] See also Kuran (2018) for a summary of the literature on Islam and economic performance.

[14] https://github.com/devdatalab/paper-anr-mobility-india/.

I. Context on Intergenerational Mobility in India

India's rapid economic growth and pervasive caste system make it a particularly important setting for understanding intergenerational mobility for at least two reasons. First, Indian society has undergone a large transformation over the last 40 years. Economic liberalization, starting in the 1980s, dismantled many parts of India's post-independence experiment with central planning. Decades of sustained economic growth have resulted in substantial reductions in poverty and the rise of a large middle class. This setting thus lets us examine intergenerational mobility in the context of rapid economic growth in a large developing nation.

Second, India's caste system is characterized by a set of informal rules that inhibit intergenerational mobility by preventing individuals from taking up work outside of their caste's traditional occupation and from marrying outside of their caste. While some have argued that economic growth is reducing the influence of old social and economic divisions, caste and religion remain important predictors of economic status (Munshi and Rosenzweig 2006; Ito 2009; Hnatkovska et al. 2013; Mohammed 2019). Since independence in 1947, the government has systematically implemented policies intended to reduce the disadvantage of communities that are classified as SCs or STs. These groups are targeted by a range of government programs and benefit from reservations in educational and political institutions.

India's Muslims constitute a similar population share as the SCs and STs (14 percent for Muslims versus 17 percent for SCs and 14 percent for STs). While Muslim disadvantage has been widely noted, including by the well-known federal Sachar Report (2006), there are few policies in place to protect them, and there has not been an effective political mobilization in their interest. To the contrary, a large-scale social movement (the Rashtriya Swayamsevak Sangh) and several major political parties have rallied around pro-Hindu platforms and policies which arguably discriminate against Muslim religious, economic, and cultural practices. Violent anti-Muslim riots have been closely tied to political parties and political movements (Wilkinson 2006; Berenschot 2012; Blakeslee 2018).

II. Methods: Measuring Intergenerational Mobility in Developing Countries

Our goal is to study the persistence of socioeconomic rank across generations, holding constant changes in inequality and growth, following the notion of intergenerational mobility in Solon (1999) and Chetty et al. (2014a). India's rapid growth and changes in inequality have been widely studied.[15] Our goal is to understand whether India's major growth episode has affected individuals' ability to move up in the *relative* status distribution.

Like many other studies in developing countries and historical settings, we focus on education as the primary measure of relative social status (Solon 1999; Güell et al. 2013; Wantchekon et al. 2015; Alesina et al. 2021; Card et al. 2022; Derenoncourt 2022). Income mobility is difficult to measure in these

[15] See, for example, Rodrik and Subramanian (2005), Lamba and Subramanian (2020), and Chancel and Piketty (2019).

settings for three reasons. First, high-quality linked parent-child income data are rare. Second, measurement error in income (which is substantial) leads mobility estimates to be biased downward; education levels are measured more accurately and do not vary over the life cycle. Third, earnings are difficult to ascribe to individuals in multigenerational households with joint production. In addition to these measurement concerns, human capital investment is a key form of intergenerational investment, and plausibly a target of public policy. Thus, intergenerational educational mobility is important even when income is observed.

A. *Rank CEF-Based Mobility Measures and Intergenerational Educational Mobility*

Our paper builds on the seminal work of Chetty et al. (2014a), who proposed absolute upward mobility and other rank-rank CEF-based mobility measures. Absolute upward mobility (denoted as p_{25} below) is the income rank of a child born to a parent at the twenty-fifth percentile (Chetty et al. 2014a). This measure has several characteristics that make it desirable as a measure of upward mobility.

First, rank measures like p_{25} distinguish mobility from economic growth and inequality. The canonical mobility measure, the coefficient from a regression of child income on parent income (or the equivalent with education), combines information from changes in rank mobility and in inequality (Chetty et al. 2014a). Similarly, the probability that a child obtains a higher raw socioeconomic outcome than their parent (used in Chetty et al. 2017 and Alesina et al. 2021) simultaneously captures intergenerational mobility and economic growth. Indeed, in developing countries like India, most children obtain more income and education than their parents; but this in itself does not tell us whether children from the bottom have better access to the best opportunities than they did before.

Second, measures like p_{25} are easily and meaningfully calculated for population subgroups. Comparing the class of linear mobility measures (like the rank-rank gradient) across subgroups is less straightforward, because the gradient compares children's outcomes against more advantaged members only of their own group. Meaningful comparison thus requires examining both the subgroup gradient and the average outcome difference across groups (Hertz 2008; Jácome, Kuziemko, and Naidu 2021). In contrast, p_{25} has the same interpretation for a subgroup as for the full population: it describes the expected rank of a child in the subgroup who is born to a parent at the twenty-fifth population percentile.[16]

Last, p_{25} directly describes outcomes in the bottom half of the distribution. Gradient mobility measures can directly inform the analyst about the bottom half if one assumes linearity. But in fact linear rank-rank CEFs (as found in the United States) appear to be the exception rather than the rule.[17]

[16] Indeed, we are not aware of any studies that have applied the Hertz (2008) method to study educational mobility; Jácome, Kuziemko and Naidu (2021) is a particularly clear implementation using this method to study intergenerational income mobility in the United States.

[17] The rank-rank CEF is highly linear in the United States, but nonlinear CEFs are found in Canada, Europe, and, as we show below, in India (Bratsberg et al. 2007; Boserup et al. 2014; Bratberg et al. 2017; Connolly et al. 2019).

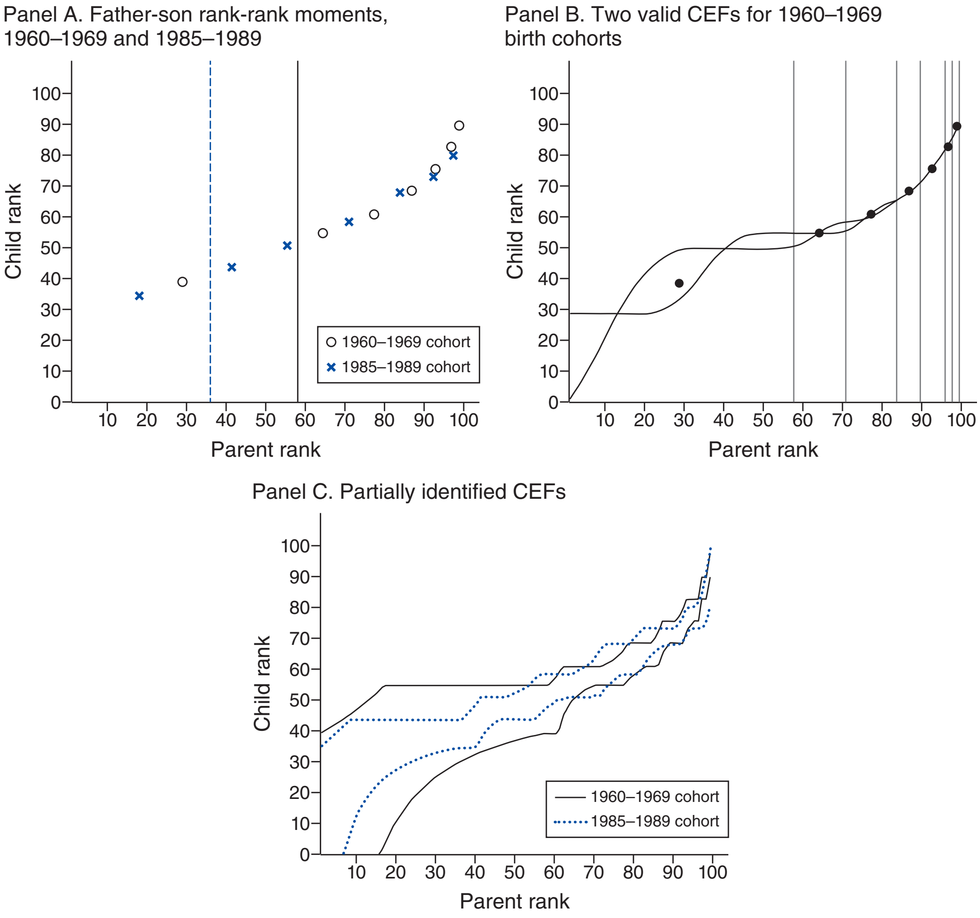

Figure 1. Father-Son Mobility: Raw Moments and Example CEFs

Notes: Panel A of the figure shows the average child education rank in each parent education rank bin for the 1960–1969 and 1985–1989 birth cohorts. The vertical lines show the boundaries for the bottom parent bin, which corresponds to less than two years of education. The solid line corresponds to the 1960–1969 birth cohort and the dashed line to the 1985–1989 birth cohort. Points are displayed at the midpoint of each parent rank bin. Panel B shows the 1960–1969 moments again, along with two simulated conditional expectation functions, which are equally good fits to the moments. Panel C shows bounds on father-son CEFs.

Source: IHDS (2012)

The key challenge to estimating p_{25} in education data is in defining what it means for a parent to be at the twenty-fifth percentile, when the twenty-fifth percentile appears in the middle of a coarse education bin. Figure 1, panel A demonstrates why this is difficult; the figure shows the average child education rank in each parent education rank bin for the 1960–1969 and 1985–1989 birth cohorts in India. In the 1960–1969 birth cohort, 57 percent of fathers report a bottom-coded education level; in the 1985–1989 cohort, the figure is 36 percent. How does one identify the expected child rank given a parent at the twenty-fifth percentile in these birth cohorts?[18]

Figure 1, panel B shows two conditional expectation functions that are both perfectly consistent with the 1960–1969 moments. The data available cannot distinguish between these two functions, but they have different implications for upward mobility: one function implies a much higher expected rank for a child born to parents at the bottom of the distribution than the other. This example highlights our methodological insight: the CEF of child rank given parent rank can at best be partially identified from education data.

Coarse data like that described here are ubiquitous in the mobility literature. In developing countries, bottom-coding rates in excess of 50 percent are widespread (Narayan and Van der Weide 2018); the same is true for older generations in richer countries (Long and Ferrie 2013). Internationally comparable censuses often report education in as few as four categories.[19] Transforming transition matrices based on education or occupation data (with arbitrary coarse rank boundaries) into quantile transition matrices faces the same barriers as point estimating the CEF above.

B. *Estimating Bounds on a CEF with Censored Education Data*

This subsection describes how we partially identify CEF-based mobility measures with coarse education data. The intuition behind our method is suggested by Figure 1, panel B. There are many possible CEFs that can fit the parent-child rank moments. The mobility measures we consider are functions of CEFs. Thus, for any given mobility measure, the bounds on the mobility measure are governed by the CEFs that minimize and maximize the measure, as long as the CEFs are consistent with the observed data. The method extends Manski and Tamer (2002) and is derived in Novosad, Rafkin, and Asher (2022). Online Appendix B formalizes the bounds and reproduces the derivation for the case of upward mobility. This section focuses on an intuitive understanding of how the function is bounded.

We require only two assumptions to meaningfully bound the CEF. The first is definitional and already implicit in the discussion above: we assume the existence of a continuous, latent education rank. This assumption falls directly out of the standard human capital model. The latent education rank is based on the education level that would be chosen from a *continuous* rather than a discrete set of education choices, if continuous choices were available. In practice, the educational choice set is lumpy, so we observe the rank only in coarse bins. The latent rank reflects how much the marginal benefit or cost of obtaining the next level of education (e.g., middle school) would need to change in order for a given individual to progress to that level. Individuals who would need only a small benefit shift to choose the next education level have the highest educational ranks within their rank bin. The latent rank

[18] The graph makes clear that p_{25} *cannot* be proxied by the expected child rank in the parent bin containing the twenty-fifth percentile. When there is less censoring, the bottom bin represents a lower average set of parent ranks.

[19] Income mobility is also often based on censored estimates; in the well-known British Cohort Study, one income bin contains more than 30 percent of the data.

is a proxy for the underlying factors that shift individuals' demand for education, which can be expected to be correlated with socioeconomic status (Card 1999).[20, 21]

Second, we assume that the conditional expectation of child education rank given parent education rank is monotonically increasing, including within education groups (where it is unobserved): having a more advantaged parent cannot make a child worse off. This assumption is motivated by the fact that average socioeconomic outcomes of children are strongly monotonic in parent socioeconomic outcomes across many socioeconomic measures and countries, as well as in every birth cohort that we study in India (online Appendix Table A1).[22] The monotonicity assumption is not necessary; online Appendix B shows how other parametric assumptions can be used to bound the mobility function (e.g., constraining the maximum curvature of the underlying conditional expectation function).

For parsimony, our framework considers a setting in which only the parent rank is interval-censored—i.e., we take the child rank variable as *not* censored. We address extensions with censored child ranks and potential bias from our approach in Section IVE.

With just these two assumptions, we can obtain sharp bounds on the CEF. Figure 1, panel C shows the bounds on the Indian CEF for the 1960–1969 and the 1985–1989 birth cohorts. The bounds will be tightest when parent education bins are small and the slope of the CEF is small: when the two neighboring bins have moments that are close in magnitude, the monotonicity restriction constrains the CEF substantially. At the bottom of the education distribution, where the bins are large and the lowest values of the CEF are constrained only by the lower rank limit of zero, the bounds on the CEF are very wide. Intuitively, because of the censoring problem, there is simply not enough information in the parent-child education distribution to bound the child rank CEF at very low parent ranks.[23]

Bottom-Half Mobility.—We propose a measure, *bottom-half mobility*, which describes the expected outcome of a child born to parents in the bottom half of the parent distribution, or $\mu_0^{50} = E\left[y \mid x \in [0,50]\right]$ for child rank y and parent rank x. Note the similarity in interpretation to absolute upward mobility, which is the expected outcome of a child born to the *median* parent in the bottom half of the

[20] Like other papers on intergenerational educational mobility, we use education strictly as a proxy for socioeconomic status. Our interpretation is meaningful even if individuals at different latent ranks within the same bin have obtained *exactly* the same number of years or hours of education. All else equal and in expectation, individuals with higher latent ranks will have socioeconomic advantages in dimensions other than years of education.

[21] Transforming education levels into ranks is analogous to transforming individual incomes into ranks, which is ubiquitous in the mobility literature even when incomes are interval censored (Rosenbaum 2000).

[22] In small samples, empirical nonmonotonicity may emerge from monotonic distributions due to sampling error. This occurs at the very top of the distribution in a minority of subgroup cohorts in our data; these nonmonotonicities do not affect our calculations of upward mobility, which only use information from bins adjacent to the bottom half of the parent distribution.

[23] Alternate estimation approaches might yield tighter predictions at the bottom of the distribution, but they do so only by imposing strong assumptions on the CEF. For example, if we estimate a linear rank-rank gradient, we can obtain a tight predicted value of the CEF at the bottom of the distribution. But the precision is misleading: the available data do not strongly motivate the linearity assumption that buys this precision.

parent distribution (i.e., $p_{25} = E[y \mid x = 25]$).[24,25] The measures p_{25} and μ_0^{50} both describe the outcomes for groups in the bottom half of the socioeconomic distribution. In both cases, a value of 50 implies convergence to the mean rank in one generation, and a value of 25 implies no convergence at all. Both measures are useful, just as the mean and median of a given distribution have advantages and disadvantages as summary statistics. The main advantage of μ_0^{50} is that, as we show below, it can often be tightly bounded in contexts where p_{25} cannot.[26]

Figure 2 provides a graphical example that walks through the calculation of μ_0^{50} for the 1960–1969 birth cohort in India. For intuition about why μ_0^{50} is tightly bounded, notice that the value of the CEF can be point estimated within ranges that correspond to bin boundaries in the data. If a given parent education bin has rank boundaries a and b, then μ_a^b is precisely equal to the expected child outcome in that rank bin. If we then seek to obtain a value that is *close* to that one, e.g., μ_a^{b+c} where c is small, it will also be bounded tightly, because that new value is just a weighted mean of μ_a^b (which is point estimated) and μ_b^{b+c}, which has a small weight.[27] In contrast, absolute mobility $p_i = E[y \mid x = i]$ cannot be point estimated at any point on the CEF. The bounds on μ_a^b will thus be tightest when a and b are close to actual rank bin boundaries and when the slope of the CEF is small.

C. *Comparison with Other Measures of Educational Mobility*

This subsection compares bottom-half mobility to other measures of intergenerational educational mobility. First, Alesina et al. (2021) study upward mobility across all of Africa and define upward mobility as the probability that a child born to a parent who has not completed primary school manages to do so. Their measure importantly captures whether poor countries now provide a minimal standard of education to households that have not been reached by the education system previously. It is a measure of absolute mobility—it can be driven by both aggregate education growth and changes in the persistence of socioeconomic rank across generations. A disadvantage is that this measure conditions on different rank groups in different times and countries; in rank terms, this measure approximately describes $E[y > 52 \mid x \in [0,76]]$ in Mozambique (where 76 percent of parents and 48 percent of children have not completed primary school) and $E[y > 18 \mid x \in [0,42]]$ in South Africa. Our approach is instead to estimate $E[y \mid x \in [0,50]]$ in all settings, thus isolating the rank persistence interpretation of intergenerational mobility, as in Solon (1999) and Chetty et al. (2014a).

[24] If the CEF is linear, $p_{25} = \mu_0^{50}$; if the CEF is concave at the bottom of the parent distribution, then $\mu_0^{50} < p_{25}$. Most CEFs in the literature are concave.

[25] Our framework can bound many other measures, such as μ_0^{20} (the expected rank of a child born in the bottom 20 percent), $E[y > 80 \mid x < 20]$ (the probability that a child born in the bottom 20 percent makes it into the top 20 percent), or the conditional *median* function of child rank given parent rank.

[26] Online Appendix B provides a formal statement of analytical bounds on μ_a^b (derived in Novosad, Rafkin, and Asher 2022) for the estimation of individual mortality risk given the same individual's education rank). It shows how we estimate arbitrary functions of the CEF, including the best linear approximator to the CEF (i.e., the rank-rank gradient).

[27] There are alternative cases where μ_a^n is not exactly a weighted mean; see Novosad, Rafkin and Asher (2022) for an example in the context of mortality.

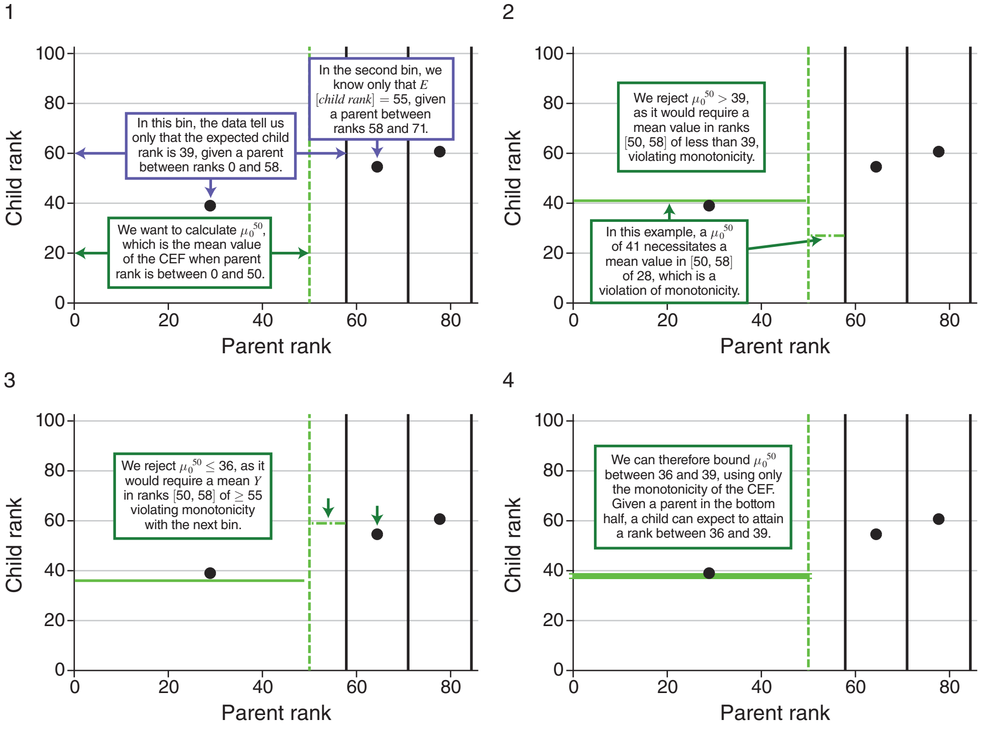

FIGURE 2. INTUITION: CALCULATING μ_0^{50} FOR 1960–1969 BIRTH COHORT

Second, when constructing transition matrices from interval data, researchers have sometimes randomly reassigned individuals across bins to create quantile bins. While this approach may seem innocuous, in fact it implicitly assumes that the CEF is a step function with zero slope between bin boundaries. This can result in biased estimates that are misleadingly precise.[28]

Third, Card, Domnisoru, and Taylor (2022) and Derenoncourt (2022) define upward mobility in the 1920s birth cohort as the share of children attaining ninth grade (i.e., passing approximately the fiftieth percentile in their cohort), given parents at a fixed education level. Their measures are similar to μ_0^{50} but are constrained to conditioning on parent percentiles with bin boundaries in the data; our work generalizes this approach, partially identifying μ_a^b for any a and b, increasing comparability across contexts.

Finally, Jácome, Kuziemko, and Naidu (2021) study Black and White intergenerational *income* mobility in the historical United States, using the decomposition method of Hertz (2008). Their core idea is that mobility differences between groups can be described with a combination of the gradient of the within-group parent-child

[28]For example, using this approach to a country like Ethiopia, where over 80 percent of parents report a bottom-coded education level, would misleadingly suggest virtually identical outcomes for children growing up in the bottom three quartiles of the parent distribution. Using bottom-half mobility in this context would yield wide bounds—describing the appropriate level of uncertainty in a context where education data provide virtually no information on relative status differences between the bottom three quartiles of the parent distribution.

CEF *and* the average outcome difference between child groups. Both components are necessary to correctly interpret subgroup mobility. In contrast, the bottom-half mobility measure lets us summarize subgroup mobility from the bottom half with a single meaningful statistic.[29]

As we have noted, a useful feature of bottom-half mobility is that it can be meaningfully compared across contexts and population subgroups. For instance, we can compare the upward mobility of Indian Muslims to Black people in the United States with a scalar measure. Naturally, this is only valid to the extent that changes in relative rank have similar meanings across contexts. We believe that it is meaningful to compare a 5-point rank change in the United States to a 5-point rank change in India, even though it would translate to different changes in education and consumption in the two locations. Rank-based mobility measures describe the extent to which a society offers its best opportunities to all of its members, whatever those opportunities are. Naturally, our measure is not a complete description of utility or well-being.

III. Data

We use linked data on parents and children from the India Human Development Survey (IHDS), a nationally representative survey of 41,554 households, with rounds in 2004–2005 and 2011–2012. The IHDS identifies religion and ST or SC status. We classify individuals who are both SC/ST and Muslim (who make up less than 2 percent of SC/STs) as Muslims.[30] About half of Muslims are Other Backward Castes (OBCs); we classify them as Muslims as well.[31]

Crucially, the IHDS records the education of parents for the majority of respondents, even if those parents have died or are not resident in the household. Mobility estimates that are based strictly on coresident parents and children can be biased if, for example, higher mobility children have a different rate of household exit; estimates from IHDS are not subject to this concern.[32]

We estimate mobility in the past by studying children from older birth cohorts, also in the 2011–2012 IHDS.[33] We pool the data into ten-year birth cohorts for

[29] We also show that the rank-rank gradient cannot be meaningfully bounded in our context due to the censoring problem.

[30] Classifying this group as SC/ST or excluding this group does not affect any of the results because overall they represent less than 0.4 percent of the population.

[31] We do not consider OBCs as a separate category in this paper because OBC status is inconsistently reported across surveys, due to misreporting, changes in OBC schedules, and inconsistency between state and federal lists. Analysis of mobility of OBCs may be feasible using subcaste-level descriptors and classifications, but is beyond the scope of this paper. We pool Christians, Sikhs, Jains, and Buddhists, who collectively make up less than 5 percent of the population, with higher-caste Hindus (*i.e.*, forward castes and OBCs); we describe this group as "Forwards/Others." We find broadly similar results if we exclude these other religions from the sample.

[32] Parent-child coresidence rates decline rapidly with child age (online Appendix Figure A1). Online Appendix Figure A2 examines the degree of bias that would arise from estimating upward mobility strictly from coresident parent-child pairs, a bias that we can measure since we observe the parent education even for individuals who do not live with their parents. The bias rises substantially for sons over age 25 and daughters over the age of 18, suggesting that earlier Indian mobility estimates based on coresident parent-child pairs with children as old as 40 should be treated with caution.

[33] To address concerns regarding recall bias, online Appendix Table A2 examines the discrepancy between individuals' reports of their parents' education, and the parents' self-report when they are in the same household. The discrepancies are small on average and uncorrelated with other household characteristics.

1950–1969, and five-year birth cohorts for 1970–1989, where we have more power. The data do not contain links for mothers or daughters for the 1950–1959 birth cohort. The oldest cohort of children that we follow was born in the 1950s and would have finished high school before the beginning of the liberalization era in the 1980s. The cohorts born in the 1980s would have completed much of their schooling during the liberalization era.[34]

We aggregated our data to the education definitions most widely used in the literature on India, resulting in seven categories: (i) illiterate with less than primary; (ii) literate with less than primary; (iii) primary; (iv) middle; (v) secondary; (vi) higher secondary; and (vii) post-secondary or higher.[35] Additional details on construction of parent-child links, coding of education categories, and other data sources can be found in online Appendix D.

Summary statistics describing socioeconomic outcomes of individuals in the top and bottom half of the education distribution are shown in online Appendix Table A3; relative to the top half, individuals in the bottom half are poorer on all dimensions, more likely to work for a wage, more rural, and more likely to be Muslim, SC, or ST.

IV. Results: Intergenerational Mobility in India

A. *Changes in Educational Attainment, 1950–1989 Birth Cohorts*

We begin by presenting summary statistics on educational attainment that motivate our efforts to measure mobility. Education levels have increased dramatically from the 1950 to 1989 birth cohorts, for all social groups (Figure 3). In the 1950–1954 birth cohort, the average woman had 2.3 years of education, while the average man had 5.1 years. In the 1985–1989 birth cohort, these numbers rise to 7.2 and 8.5 years, respectively. Muslims, SCs, and STs consistently report one to two fewer years of education than forward castes. SC men have had particularly large educational gains, their gap with forward groups having closed by almost 50 percent over the sample period. All groups have gained education over the cohorts, but the extent of catch-up with forward caste groups is variable. As has been documented elsewhere, all groups have also seen gains in income and living standards.

Our goal is to understand whether substantial gains in education have translated into changes in *relative* socioeconomic status. Has the growth in education increased the probability that children from low-status dynasties can obtain as high a level of education as the better-off members of their cohorts?

Parent-child transition matrices convey information about intergenerational mobility, but interpreting changes in the transition matrix over time is not

[34] Cohorts born in the 1990s may not have completed their education at the time that they were surveyed and are therefore excluded.

[35] While the IHDS data include education with higher granularity than these seven categories, there is still extreme interval censoring at the bottom of the distribution. Moreover, there is only a small sample of parents at most between-category levels, and the mean education of their children is thus noisily estimated. Given the reliance of our approach on monotonicity of expected child outcomes given parent outcomes, including these uncommon granular education observations adds more noise than signal to the result. We therefore use the aggregated categories for our main results, but show robustness to use of the granular education categories.

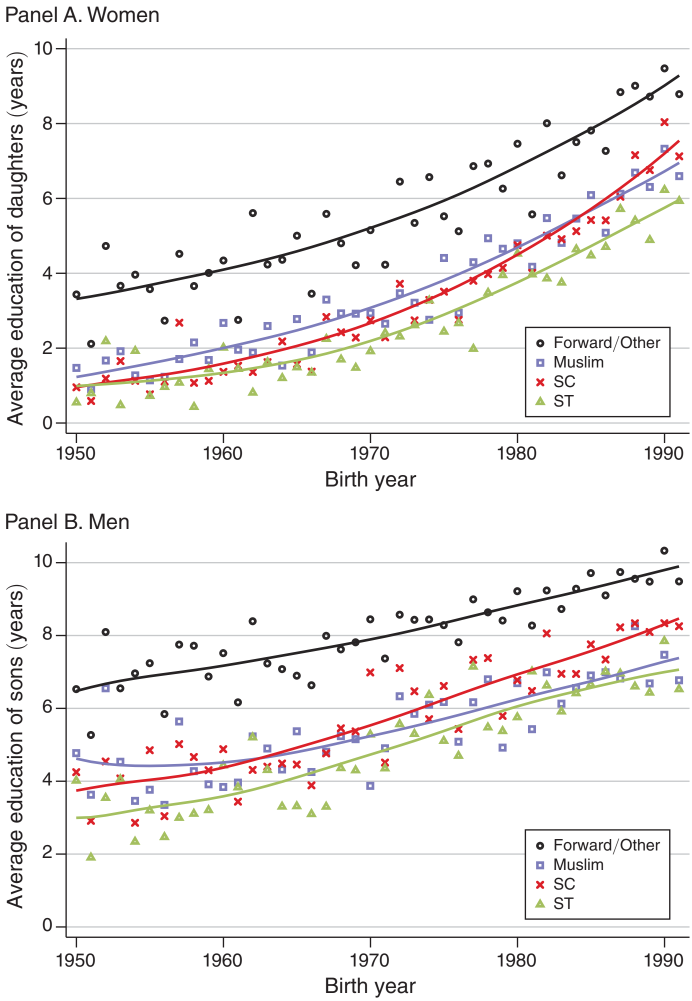

Figure 3. Changes in Educational Attainment by Birth Cohort

Notes: The points show average levels of educational attainment for each birth cohort/social group. The lines are lowess curves fit to each group's time series.

Source: IHDS (2012)

straightforward when aggregate education levels are shifting substantially. Online Appendix Table A1 shows the father-son transition matrix for each decade in our sample. The information in the transition matrix is aggregated and summarized in Panel A of online Appendix Figure A3, which shows how child ranks given parent education ranks have changed over time, but also how given parent education levels represent different parent ranks in different generations. The set of points representing each decade describes the CEF of parent rank given child rank.

Panels B through D of the same Appendix figure show the same rank-rank transition plots for father-daughter, mother-daughter, and mother-son pairs. A notable feature of the mother-child graphs is that a very large share of mothers have bottom-coded education. The relative ranking of mothers *within* these bins is unobserved; we will show below that this makes it impossible to precisely estimate upward mobility for them.[36] The nonparametric graphs in the four panels of online Appendix Figure A3 suggest that Indian intergenerational mobility has not changed much over time: the implied CEFs of the different decades appear to be largely overlapping, with the possible exception of a small reduction in the persistence of outcomes at the very top of the distribution.

B. *Introducing Bottom-Half Mobility*

We continue by calculating formal measures of upward mobility from the statistics in the previous section. Panels A through C of Figure 4 respectively show bounds on the rank-rank gradient (β), absolute upward mobility (p_{25}), and bottom-half mobility (μ_0^{50}) for the 1960–1969 and the 1985–1989 birth cohorts in India. As benchmarks, we show similar measures for United States and Denmark.[37] The bounds on the conventional measures β and p_{25} are not informative either in levels or in changes. In contrast, bottom-half mobility is bounded tightly in the 1960s and nearly point-estimated in the 1980s.

Figure 4 shows the key advantage of bottom-half mobility: it can be tightly bounded even with severely interval-censored rank data, when other measures are uninformative. Intuitively, bottom-half mobility is tightly bounded when there is a rank bin boundary close to percentile 50.

The wide bounds we calculate for β and p_{25} reflect a strength of our approach relative to prior work. When we account for interval censoring, μ_0^{50} is the only measure of the three that is informative. When rank data are highly censored, we should indeed have less certainty over the ability of individuals to move up from the bottom of the rank distribution. Conventional methods for measuring β or p_{25} in this setting can deliver precise point estimates, but only on the basis of implicit assumptions (like CEF linearity) which researchers seldom justify and may be unlikely to hold.

In fact, estimates that do not account for interval censoring at all can deliver inaccurate results. If one naively calculates the rank-rank gradient (using the mean education ranks in each bin), not only is precision overestimated, but the gradient point estimate *falls* from 0.57 in 1950–1959 to 0.51 in 1985–1989. This result erroneously suggests an improvement in mobility, when the true value of β is not precisely estimated enough to draw this conclusion. Why do the data yield this result? The nonparametric CEF plot (online Appendix Figure 1) suggests that this difference is driven entirely by the very top of the distribution. The best-fit line producing β, taking the bin means as given and without accounting for interval censoring, changes slope to fit the top of the distribution. But other best-fit lines could accurately match the data.

The rank-rank gradient, absolute mobility, and bottom-half mobility are all scalar statistics that capture different characteristics of the intergenerational persistence of

[36] A stylized example: if 95 percent of mothers had bottom-coded education, observing the joint distribution of mother-daughter education outcomes would provide little information on the relative likelihood of emerging from the bottom of the latent distribution.

[37] The rank-rank gradient is benchmarked against educational mobility estimates from Hertz et al. (2007). For p_{25} and μ_0^{50}, we use income mobility estimates from Chetty et al. (2014a).

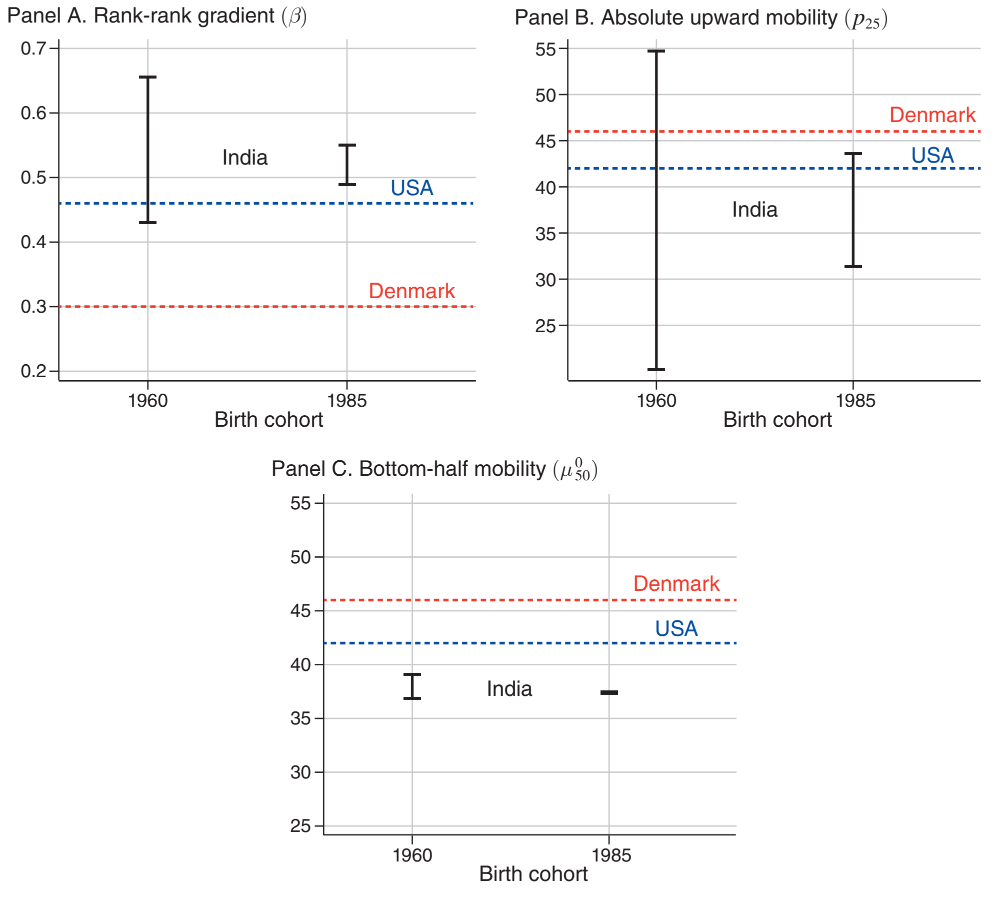

Figure 4. Bounds on Father-Son Mobility Measures in India: 1960–1969 and 1985–1989 Birth Cohorts

Notes: The figure shows bounds on three mobility statistics for the 1960–1969 and 1985–1989 birth cohorts, estimated on father-son pairs in India. For reference, we display estimates of similar statistics from United States and Denmark. Data on rank-rank education gradients for United States and Denmark are from Hertz (2008). For p_{25} and μ_0^{50}, the United States and Denmark references are income mobility estimates from Chetty et al. (2014a). The Indian measures are all based on education data. The rank-rank gradient is the slope coefficient from a regression of son education rank on father education rank. p_{25} is absolute upward mobility, which is the expected rank of a son born to a father at the twenty-fifth percentile. μ_0^{50} is bottom-half mobility, which is the expected rank of a son born to a father below the fiftieth percentile.

Source: IHDS (2012)

rank, and they may all be of independent policy interest. However, only bottom-half mobility can be measured informatively given the type of education data typically available in developing countries.

C. *Changes in National Upward Mobility*, *1950–1959 to 1985–1989*

Our main measure of upward mobility—bottom-half mobility—has been largely flat over the sample period (Figure 5). Upward mobility for men moved from $[36.6, 39.0]$ for the 1960–1969 birth cohort to $[37.5, 37.9]$ for the 1980–1985 birth

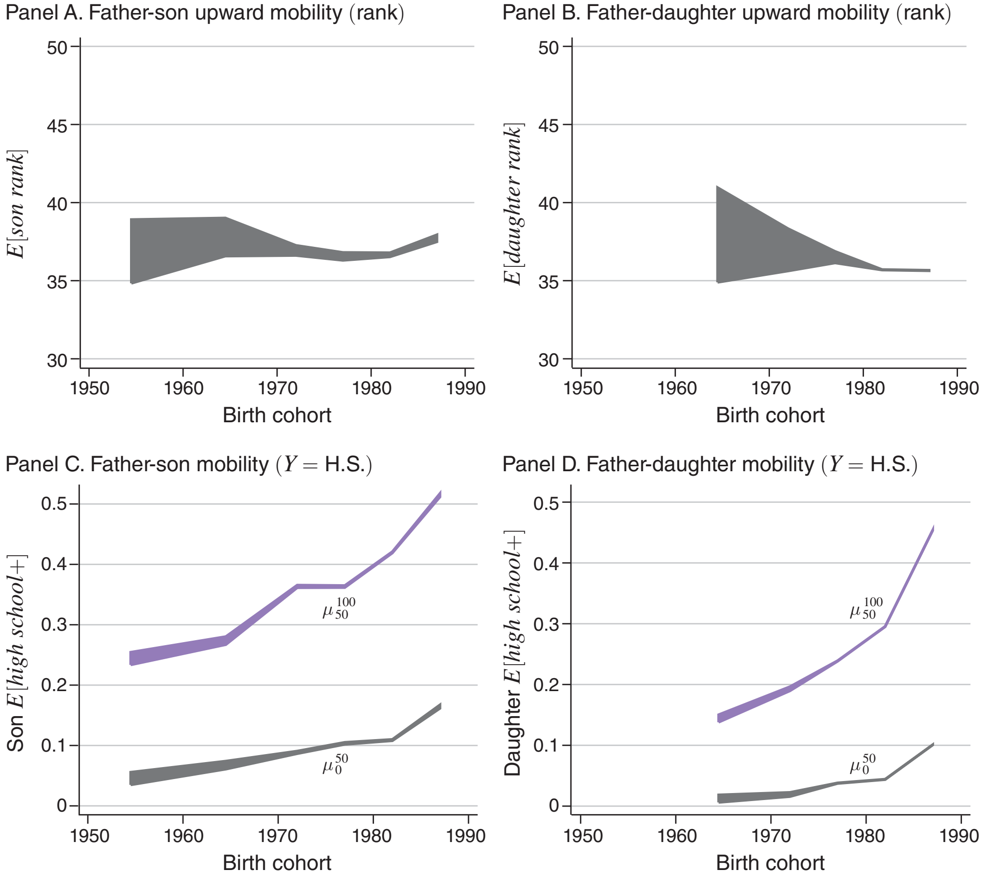

Figure 5. Bottom-Half Mobility, 1950–1989 Birth Cohorts

Notes: Figure 5 presents bounds on national intergenerational mobility, using cohorts born from 1950 through 1989. Panels A and B show bottom-half mobility ($\mu_0^{50} = E[y|x \in [0;50]]$), where x is parent rank and y is child rank. This is the average rank attained by children born to parents who are in the bottom half of the education distribution, respectively for sons and daughters. Panels C and D show an analogous measure, $E[HS|x \in [0;50]]$ (gray) and $E[HS|x \in [50;100]]$ (purple). The first (gray) is the share of children completing high school, conditional on having parents in the bottom half of the education distribution. The second (purple) is the share of children completing high school, conditional on having parents in the top half of the parent distribution.

Source: IHDS (2012)

cohort (Figure 5, panel A).[38] For comparison, this measure in the United States, which has low intergenerational mobility by OECD standards, is 41.7.[39] The mobility bounds for the 1950–1959 birth cohorts are wider, leaving open the possibility of gains from the 1950s to the 1960s birth cohorts.[40]

[38] When the distance between upper and lower bounds is less than 0.3, we report the midpoint as a point estimate.

[39] Source: Our calculations using μ_0^{50}, based on data from Chetty et al. (2020).

[40] Online Appendix Figure A4 shows that these results are unlikely to be affected by survivorship bias. We estimate upward mobility for the same birth cohorts using the IHDS 2004–2005; if mobility estimates for older cohorts were affected by differential mortality of high mobility groups, we would find different estimates from the earlier data, but the bounds are highly similar and equally static.

Among daughters, we cannot reject a broadly similar pattern to the father-son results, though the wider bounds leave open the possibility of mobility losses over the period (Figure 5, panel B). In the youngest birth cohort, father-daughter mobility is 35.6, about two rank points lower than father-son mobility. Daughters are thus less likely to escape low relative socioeconomic status than sons.

Obtaining informative mobility estimates for mother-child relationships is more difficult, because mothers are much more likely to be in bottom-coded education categories.[41] Under such severe censoring, we cannot estimate μ_0^{50} with any precision. Even in the most recent 1985–1989 birth cohort, we estimate bottom-half mobility to be $[37.5, 41.4]$ for mother-son pairs and $[33.8, 39.1]$ for mother-daughter pairs.[42] We thus focus on estimates of mobility based on fathers.

We can also calculate μ_0^{50} with child education *levels* as the *y* variable. For example, $E\big[\textit{child years} \geq 12 \,|\, x \in (0,50)\big]$ describes the likelihood that a child attains high school or greater, conditional on having a parent in the bottom half. Panels C and D of Figure 5 show this measure for father-son and father-daughter links, respectively. The graphs also show $E\big[\textit{child years} \geq 12 \,|\, x \in (50,100)\big]$, the likelihood of high school attainment given a parent in the *top* half of the education distribution. These graphs show the secular increase in high school attainment over time for children from privileged and underprivileged backgrounds. Daughters from bottom-half families have experienced the smallest gains, while daughters born in the top half of the distribution have almost closed the gap with well-off sons. For both sons and daughters, gains in high school attainment have accrued almost entirely for children from the top half of the distribution, a reflection of the stagnant overall upward rank mobility seen in panels A and B. The estimates in panels C and D reflect both changes in intergenerational mobility and changes in aggregate years of education earned; for this reason, we continue to focus on μ_0^{50}, which isolates changes in relative status, going forward.

The flat trend in μ_0^{50} does not rule out countervailing trends in different parts of the bottom half. For instance, it is possible that μ_0^{25} has gone up during the sample period, while μ_{25}^{50} has gone down. While changes like these would be substantive and important, given the censoring of latent education ranks, there is simply not enough information in the parent education distribution to precisely estimate μ_0^{25} in the early years of our sample. Calculating upward mobility from the very bottom of the distribution is perhaps possible with data on parent occupation or incomes, but is beyond the scope of this paper.

To summarize, children born to less privileged families in post-liberalization India have very similar prospects for moving up in the rank distribution as they did in the pre-liberalization era. To be clear, living standards have improved for individuals across the rank distribution; it is the probability of making progress in rank terms that is unchanged and low by international standards. This result stands in contrast to the narrative of India becoming a land of greater mobility in terms of relative social status.

[41] Among mothers of the 1960s birth cohort, 82 percent had less than two years of education. For the 1985–1989 birth cohort, this number was 65 percent.

[42] Online Appendix Figure A4 shows the admittedly uninformative graph of this measure over time.

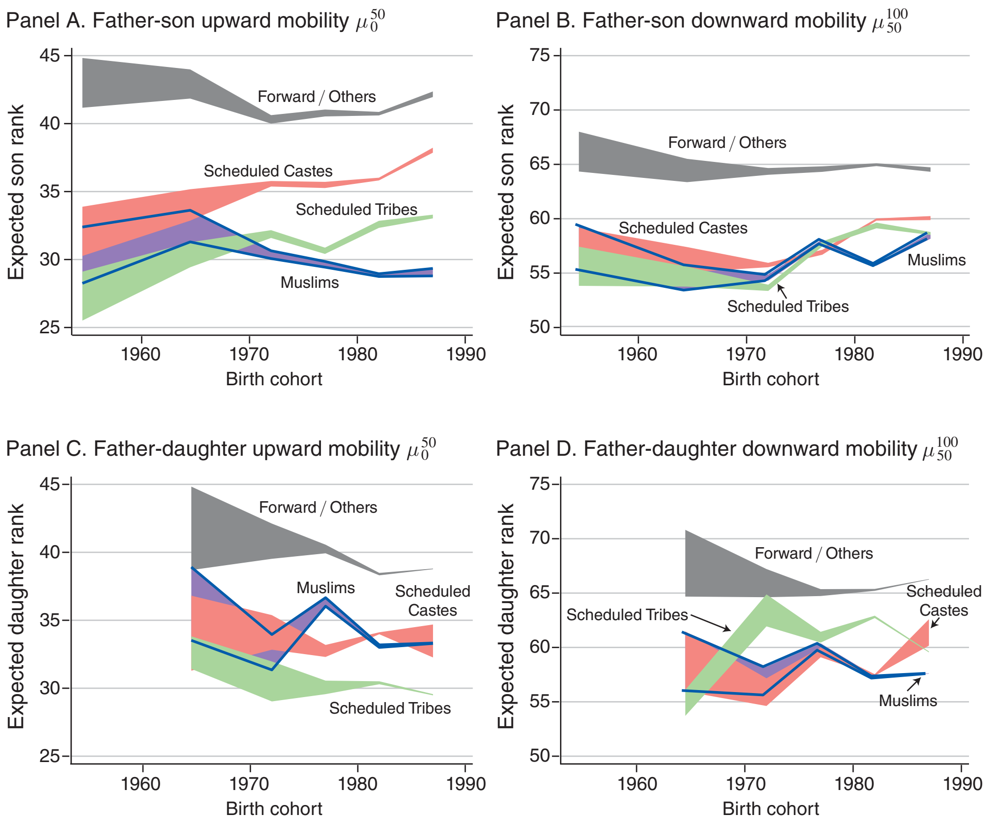

FIGURE 6. TRENDS IN MOBILITY BY SUBGROUP, 1950–1989 BIRTH COHORTS

Notes: Figure 6 presents bounds on trends in intergenerational mobility, stratified by four prominent social groups in India: Scheduled Castes, Scheduled Tribes, Muslims, and Forward Castes/Others. The mobility measure in Panels A and C is bottom-half mobility (μ_0^{50}), or the average rank among children born to fathers in the bottom half of the father education distribution. The measure in Panels B and D is top-half mobility (μ_{50}^{100}), or the average rank among children born to fathers in top half of the father education distribution. Linked father-daughter education data are not available for the 1950–1959 birth cohort.

Source: IHDS (2012)

D. *Changing Mobility across Social Groups*

We next examine how these levels and trends differ across social groups. Figure 6 presents bottom- and top-half mobility for Muslims, SCs, STs, and Forwards/Others. Panel A shows father-son pairs, revealing substantial trend differences across groups. As noted by other researchers, upward mobility for STs, and especially for SCs, has improved substantially (Hnatkovska et al. 2012; Emran and Shilpi 2015). The expected rank for SC children born in the bottom half of the parent distribution has risen from $[33, 35]$ in 1960–1969 to 38 in 1985–1989, closing half of the mobility gap with upper castes. Upward mobility for members of STs rose from $[29, 31]$ to 33 over the same period.

In contrast, Muslim upward mobility declined substantially, falling from $[31,34]$ to 29 in the same period. This change not only constitutes a major decline in mobility, but makes Muslim men the least upwardly mobile group in modern India. Muslim mobility also appears to be low in contrast with marginalized groups elsewhere in the world: the comparable figure for US Black men is 35.[43] Mobility for Muslim sons is lower even than for ST sons—even though many STs continue to live in rural areas that are geographically distant from India's growth centers. The fact that a Muslim man from a bottom-half family can expect to obtain only the twenty-ninth percentile implies that Muslims born into low status are very likely to remain low status. Finally, the Forward/Others group, predominantly higher-caste Hindus, shows little change, with mobility shifting from $[42,44]$ to 42. The flat trend in upward mobility for sons can therefore be decomposed into gains for SCs and STs and losses for Muslims.

Panel B shows downward mobility $\left(\mu_{50}^{100}\right)$ for father-son links over the same period; this measure reflects the persistence of high status for members of each group. We see a small amount of convergence between the three marginalized groups and the Forward/Others group, chiefly from the 1970s to the 1980s birth cohort. But there is no sign of the dramatic divergence between SCs and Muslims that we find for upward mobility.[44]

To interpret these results, note that one rank point is associated with about 0.15 years of education on average in 1985. This suggests that while bottom-half Muslim and SC men from the 1950–1959 birth cohort attained similar levels of education, by the 1980–1985 birth cohort, SCs could expect to attain 1.4 years more than Muslims, on a base of 6.5 years of education.

Panels C and D of Figure 6 show the same results for father-daughter pairs. Among daughters, with the exception of recent minor gains for SCs from top-half families, none of the marginalized groups have made substantial gains relative to Forwards/Others. There is also little sign of the divergence between SCs and Muslims that we observe among sons.

To summarize, we observe a sharp divergence between upward mobility for sons from SC and Muslim groups. Muslim sons from poor families have declining mobility and very little opportunity to improve their relative social status.[45]

It is important to clarify the extent to which these outcomes are zero-sum. Ranks are inherently zero-sum, but rank improvements in the bottom half can come from either other groups in the bottom half or from the top half of the distribution.[46] The rise of bottom-half mobility for SCs implies a decline in bottom-half mobility for

[43] This figure was calculated using the methodology in this paper and education data from Chetty et al. (2020). Bottom-half *income* mobility for US Black men is 39 (Chetty et al. 2020).

[44] Online Appendix Figure A shows analogous results to Figure 6, but with education levels (at least primary, and at least high school) as outcomes, rather than education ranks. The advantage of these graphs is that they present outcomes for children that are not subject to interval censoring; the parent variable remains interval censored. The results are consistent with the rank-based estimates, confirming that the separation between SC and Muslim sons is not driven by unobserved changes in interval-censored ranks for *children* from these groups.

[45] This low mobility may also adversely affect female Muslims, since marriage is nearly universal in India and almost entirely within social group, and female labor force participation is very low. Understanding how marriage ties interact with the upward mobility of sons and daughters is of interest but beyond the scope of the present paper.

[46] Note that lower expected ranks in the top half of the distribution are mechanically tied to rising equality of opportunity—they mean that individual opportunity to achieve the top ranks is being broadened.

TABLE 1—CHANGES IN UPWARD MOBILITY OVER TIME

	All groups	Forward/Others	Muslims	SCs	STs
Panel A. Father-son pairs					
1960–1969	[36.6, 39.0]	[41.8, 44.0]	[31.3, 33.6]	[32.9, 35.2]	[29.4, 31.3]
	{35.7, 39.8}	{40.6, 45.3}	{29.4, 35.6}	{31.3, 36.8}	{27.0, 33.7}
1980–1989	[37.1, 37.2]	[41.3, 41.3]	[28.9, 29.0]	[36.9, 37.0]	[33.1, 33.1]
	{36.4, 37.8}	{40.3, 42.3}	{27.7, 30.2}	{35.5, 38.5}	{31.2, 35.1}
Change over time	[−1.9, 0.6]	[−2.7, −0.5]	[−4.7, −2.3]	[1.8, 4.1]	[1.8, 3.7]
	{−3.0, 1.7}	{−4.3, 1.1}	{−7.0, −0.0}	{−0.4, 6.3}	{−1.3, 6.8}
Fraction overlapping bounds	0.818	0.310	0.054	0.090	0.119
Panel B. Father-daughter pairs					
1960–1969	[34.9, 41.0]	[38.7, 44.8]	[33.5, 38.9]	[31.3, 36.8]	[31.4, 33.8]
	{34.1, 41.8}	{37.6, 45.9}	{31.8, 40.6}	{29.8, 38.3}	{29.0, 36.2}
1980–1989	[35.4, 35.5]	[38.0, 38.2]	[32.0, 33.5]	[32.9, 34.2]	[30.4, 30.5]
	{34.6, 36.3}	{36.8, 39.3}	{30.9, 34.6}	{31.7, 35.4}	{28.3, 32.6}
Change over time	[−5.6, 0.6]	[−6.9, −0.5]	[−6.9, −0.0]	[−3.9, 2.9]	[−3.4, −0.9]
	{−6.7, 1.7}	{−8.4, 1.0}	{−8.8, 1.9}	{−5.8, 4.8}	{−6.6, 2.2}
Fraction overlapping bounds	0.781	0.244	0.434	0.985	0.346

Notes: Table 1 shows estimates of full sample and subgroup bottom-half mobility (μ_0^{50}) for the 1960–1969 and 1980–1989 birth cohorts for father-son (panel A) and father-daughter (panel B) pairs. We show both bounds (in square brackets) and 90 percent confidence sets (in curly braces) on those bounds. The table also reports the bounds and 90 percent confidence sets on the change in bottom-half mobility between these two time periods. We obtain confidence sets by generating 1,000 bootstrap draws, estimating bounds on each bootstrap draw, and following the framework in Chernozhukov, Hong, and Tamer (2007) to form 90 percent confidence sets from bootstrapped bounds. Because these are confidence sets rather than confidence intervals, instead of p-values we show the fraction of bootstraps in which the 1960–1969 and 1980–1989 bounds are overlapping.

Source: IHDS (2012)

some other group only if there are no changes in the persistence of rank in the top half of the distribution. Equivalently, if upward mobility had improved in India as a whole, then it would be possible for poorer members of *all* social groups to see improving chances to move up in relative terms.

Statistical Tests.—Table 1 summarizes the changes over time for the full sample and all the population subgroups, along with bootstrap confidence sets to account for sampling variation, calculated following Chernozhukov, Hong, and Tamer (2007).[47] Table 2 shows confidence sets for mobility differences between groups for the youngest (1985–1989) birth cohort.[48]

E. *Robustness of Bottom-Half Mobility to Alternate Assumptions*

Nonuniform Within-Bin Subgroup Distributions.—Our mobility calculations draw on the uniformity of the rank distribution. Ranks are uniform in the national

[47] Online Appendix Table A4 shows similar estimates with ranks calculated from the granular years of education in the IHDS; the mobility estimates and subgroup differences are nearly identical.

[48] The confidence sets in Tables 1 and 2 are wider than mobility confidence intervals from prior studies because they reflect both statistical variation and uncertainty due to coarse measurement of education, the latter of which has not been addressed by prior studies. The cross-group differences in the youngest birth cohort are all highly significant, as is the trend difference between SCs and Muslims.

Table 2—Group Differences in Upward Mobility

	F/O minus SC	F/O minus Muslim	SC minus Muslim
Father/son (μ_0^{50})	[4.6, 5.0] {2.8, 6.8}	[11.6, 12.1] {10.0, 13.8}	[6.9, 7.3] {4.5, 9.6}
Fraction overlapping bounds	0.000	0.000	0.000
Father/daughter (μ_0^{50})	[4.2, 4.5] {1.9, 6.8}	[5.1, 5.5] {2.9, 7.7}	[0.8, 1.1] {−2.0, 3.9}
Fraction overlapping bounds	0.001	0.000	0.511
Father/son (μ_{50}^{100})	[4.8, 5.3] {3.3, 6.8}	[9.0, 9.4] {5.8, 12.6}	[3.9, 4.3] {0.8, 7.5}
Fraction overlapping bounds	0.000	0.000	0.005
Father/daughter (μ_{50}^{100})	[7.7, 8.0] {4.0, 11.7}	[7.8, 8.2] {5.3, 10.8}	[0.0, 0.4] {−3.9, 4.3}
Fraction overlapping bounds	0.000	0.000	0.318

Notes: Table 2 shows estimates of cross-group differences in bottom-half mobility (μ_0^{50}) and top-half mobility (μ_{50}^{100}) in the 1980–1989 birth cohorts. F/O stands for Forwards/Others and SC for Scheduled Castes. We show both bounds (in square brackets) and 90 percent confidence sets (in curly braces) on those bounds. We obtain confidence sets by generating 1,000 bootstrap draws, estimating bounds on each bootstrap draw, and following the framework in Chernozhukov, Hong, and Tamer (2007) to form 90 percent confidence sets from bootstrapped bounds. Because these are confidence sets rather than confidence intervals, instead of p-values we show the fraction of the bounds for the two social groups that are overlapping. For example, the value of 0.511 in the final column indicates that 51.1 percent of the bootstraps generate overlapping bounds for the two groups.

Source: IHDS (2012)

within-gender samples by construction. In population subsamples, uniformity is not guaranteed: for instance, the distribution of ranks of Muslim fathers, conditional on being in the bottom half or the bottom education bin, is not necessarily uniform. Is it possible that, conditional on being in the bottom education bin, Muslim fathers now have lower ranks than the SCs in that bin, such that our bottom-half mobility calculations are biased? Such a scenario is certainly possible and needs to be considered when working with bottom-half mobility for population subgroups.[49] We present three pieces of evidence that the Muslim-SC divergence is not biased by nonuniformity in subgroup rank distributions.

First, online Appendix C.1.1 asks whether Muslim parents with low education lost substantial ground in other dimensions against SC parents with low education over the sample period. Muslims in the parent generation have higher education than SCs in the entire sample period; they also have higher consumption, including in the subset of households where the household head has zero years of education. The Muslim-SC education gap closed slightly from a 4-rank point difference to a 3-rank point difference. The consumption gap for individuals in the bottom half did not close at all. The evidence does not suggest a substantial decline in

[49] Note that this concern is relevant in all studies of intergenerational educational mobility with coarse education data; it is not specific to bottom-half mobility or to our method.

the socioeconomic status of Muslim parents vis-à-vis SC parents, conditional on being in the bottom half, or the bottom education bin.

Second, we show in online Appendix C.1.2 that both the divergence between Muslims and SCs and the convergence between SCs and Forwards/Others are robust to defining parent education ranks *within* social groups rather than within the national distribution. These new ranks are uniform by construction, mitigating this particular bias concern, and give very similar results. We do not use this measure for our primary results because it no longer conditions on parents with comparable socioeconomic status; a parent in the least educated 50 percent of forward castes has more education than a parent in the least educated 50 percent of SCs. Nevertheless, this robustness check suggests that the trends observed in the main analysis are unlikely to be driven strictly by latent rank distribution changes within education bins.

Third, in online Appendix C.1.3, we simulate an uncensored parent rank distribution by fitting various parametric distributions to the cohort $\times$ subgroup parent education data. This approach addresses the coarse education bin problem but requires a parametric assumption about the latent education distribution. Intuitively, this approach infers the structure of the within-bin latent rank distribution based on the across-bin data distribution for each group. The estimated parameters confirm the result above that the relative ranks of bottom-half Muslim and SC parents have not changed much over the sample period; even under the worst-case assumptions for our estimation, at most 10 percent of the growing gap between Muslims and SCs can be explained by unobserved changes in latent parent ranks.[50]

Interval Censoring of Child Ranks.—We have assumed thus far that child ranks are directly observed at the midpoint of each child's rank bin. But child ranks are also interval censored. We present two pieces of evidence that unobserved variation in latent child ranks is unlikely to affect our main results.

First, when we use uncensored measures of child outcomes, such as primary or high school completion (in Figure 5, panels C and D, and online Appendix Figure A6), we continue to find both substantial divergence of SCs and Muslims from bottom-half families and a lack of relative progress for bottom-half individuals. These measures still hold *parent rank* fixed across generations, so they are valid for cross-group comparisons over time. The Muslim-SC divergence thus cannot be an artifact of child rank censoring.

Second, the censoring problem for child ranks is much smaller than the censoring problem for parent ranks, because children are on average more educated than their parents. This makes their education bins more evenly distributed. In the 1960–1969 birth cohort, the bottom child education bin—still the largest bin—contains 26.5 percent of the population; in 1985–1989, it contains only 9 percent. The bias from child rank censoring will thus be considerably smaller than that from parent

[50] To understand why the potential bias is small, note that while SCs and Muslims are disadvantaged on average, their education distributions still have significant overlap with higher-caste Hindus. The perverse hypothetical case of Muslims being bunched entirely at the bottom of the distribution (in 1985 only) is a theoretical concern, but very implausible in practice given the distributions of the Muslim population across education bins.

rank censoring, where the largest bins contain 60 percent of the data. These results suggest that using the midpoint of a child's rank bin captures most of the meaningful variation in child ranks in our context.

V. Potential Mechanisms for Subgroup Mobility Differences

In this section, we explore mechanisms that could explain the upward mobility differences and changes across social groups in India. We focus on explaining our most striking subgroup finding: the growing mobility gap between SCs and Muslims. Understanding the divergence between these groups is important in its own right for explaining mobility trends for a third of India's population and for understanding the drivers of mobility in India and other developing countries. These analyses are suggestive, but they point toward affirmative action as a key mechanism for the SC/Muslim divergence. We largely reject (i) differential fertility, (ii) geography, and (iii) differential occupational patterns as mechanisms.

A. *Affirmative Action for Scheduled Groups*

First, we consider the hypothesis that the basket of programs and policies targeted to SCs has driven the increase in upward mobility of SCs relative to Muslims since the 1950s. Affirmative action for SCs consists of (i) reservations (positions only available to SCs) in higher education admissions, political offices, and public sector employment; (ii) direct educational benefits, such as scholarships and dedicated schools for SCs; (iii) welfare schemes where SC status is a specific criterion; and (iv) regional programs that use SC shares as a targeting criterion.

To study how affirmative action affects bottom-half mobility, we exploit a change in 1977 in the social groups eligible for SC status, studied in Cassan (2019). At independence, lists of SCs were defined in each state, largely following definitions from the British regime. The SC category comprises many *jatis*—more granular caste groups. In 1956, states were reorganized along linguistic lines, but the SC lists, which were maintained at a regional level, were not substantially changed. This created variation where the same group might be on the SC list in some states but not in others. In 1977, a federal law harmonized these lists within states, which resulted in the arbitrary addition of many regional social groups to the SC lists; the policy change made 2.4 million Indian citizens newly eligible for SC-targeted benefits, on a base of 80 million. The policy change makes it possible to examine the impact of SC status, while controlling for a group's ethnicity, historical experience, and narrow geographic region, and for state $\times$ jati differences. The changes in status are plausibly exogenous to educational trends for local social groups, because the SC list changes were guided by clear federal rules that were applied across the country without discretion.

We use data from Cassan (2019) to test whether groups who were newly added to scheduled lists experienced upward mobility gains. Our sample definition includes individuals who were on a SC list after the policy change in 1977. The control group consists of people in groups that were already on SC lists in their region in 1956; we

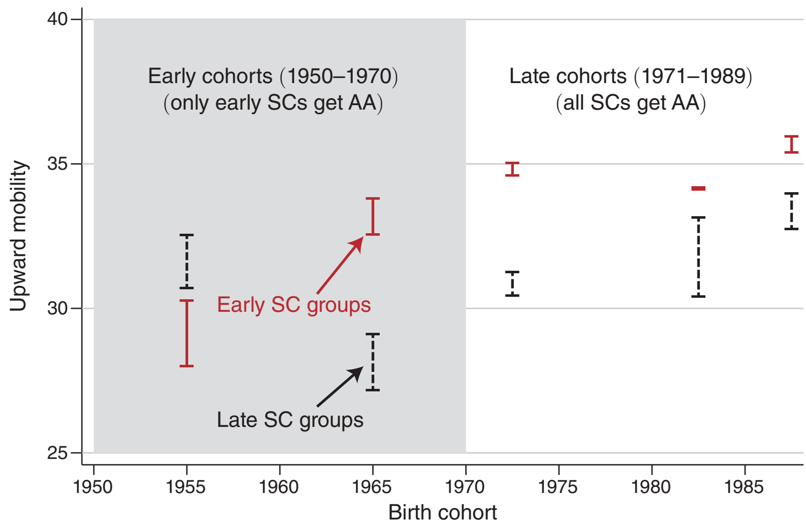

Figure 7. Jati Redesignation and Intergenerational Mobility

Notes: Figure 7 shows bounds on bottom-half mobility μ_0^{50} for two social groups in India. The red series shows upward mobility for groups that were designated as Scheduled Castes since before 1956. The black series shows upward mobility for groups that were not designated as Scheduled Castes until 1977; birth cohorts later than 1970 (outside the gray box) were six or younger when the policy change was implemented and thus would have had SC status at the primary school-starting age.

Source: IHDS (2012)

call these “early SCs.” The treatment group are those who only appeared on SC lists in 1977; we call these “late SCs.” As in Cassan (2019), our primary specification assumes that individuals needed to be six years old or younger in 1977 to benefit in terms of education from the change in status. We therefore treat individuals born later than 1970 as being in the late SC group.

We assign individuals to early and late SC groups using the IHDS’s jati-level group identifiers. Figure 7 shows the upward mobility trajectory of sons in the early and late SC groups and presents bounds on μ_0^{50} over time. In the 1950s and 1960s, the early SC group experiences rapid relative increases in mobility and diverges from the late SC group, which has not yet obtained protected status. This gap begins to close after the 1970 birth cohort, the first group where members of late SC groups are young enough to benefit from having protected status.

We formally estimate the impact of SC status with the following regression based on Cassan (2019):

$$(6.1) \qquad Y_{i,j,r,c} = \beta_0 + \beta_1 LateSC_{j,r} + \beta_2 post_c + \beta_3 (post_c \times LateSC_{j,r}) + \nu_{j,r} + \eta_{r,c} + \zeta_{j,c} + \epsilon_{i,j,r,c},$$

where $Y_{i,j,r,c}$ is the education rank of child i in jati j, region r, and birth cohort c. We include fixed effects for jati $\times$ region (ν), region $\times$ cohort (η), and jati $\times$ cohort (ζ). These fixed effects exploit the fact that the same jati group could be an early SC or a late SC depending on its region within a state. The coefficient of interest, β_3, therefore compares individuals in late SC groups born after 1970 to those born earlier in the same narrow social group, controlling for outcomes of individuals from early SC groups in the same region, and for outcomes of members of the same jati group in other regions. Regressions are clustered at both the jati and the region level.

Using bottom-half mobility as a dependent variable in this specification is challenging because, for each group, it can at best be bounded. Instead, we proxy bottom-half mobility by restricting the sample to a set of father ranks that is as close as possible to the bottom 50 percent but can also be point identified as consistently as possible across the different decades. To do this, we define Y as μ_0^{59} in the 1950s birth cohort, μ_0^{57} in the 1960s, and μ_0^{58} in the 1970s and 1980s.[51]

Table 3 shows the results, for men in panel A and women in panel B. Column 1 of panel A shows that male SC cohorts exposed to the basket of affirmative action policies experienced an 8-rank-point increase in upward mobility. Column 2 shows robustness to a specification where we limit the sample to sons of fathers with fewer than two years of education in all years instead of the sons of fathers in the approximate bottom 60 percent as noted above. The point estimate is similar. Column 3 (using the same sample as column 1) shows that both the 1970s and 1980s birth cohorts of late SCs benefited relative to the earlier cohorts. The same specifications in panel B show no evidence of mobility gains for women in late-designated SC groups.[52]

These results line up with Cassan (2019), who found that, for men, gaining SC status led to a 10 percentage point increase in literacy and a 7 percentage point increase in secondary school attainment, while there were no measurable gains for women. However, it was not necessarily the case that these gains would accrue to children of low education parents; indeed, affirmative action's critics in India often suggest that it is captured by a "creamy layer" of the already prosperous among targeted groups, a story that our findings reject.

The results are consistent with affirmative action causing a large increase in upward mobility for SC groups. The rank-point gain of late SCs is comparable in magnitude to the upward mobility gap in the 1980s between Muslims and SCs, and it emerged over only 20 years of affirmative action. This said, our empirical analysis does not directly test for the effect of affirmative action on SCs as a whole vis-à-vis Muslims, and general equilibrium effects could differ. However, if these treatment effects for late SCs are externally valid for the potential effect of affirmative action on the broader population of SCs and Muslims, then affirmative action could explain the entire contemporary mobility gap between these groups.

[51] We use the midpoint of the bounds as a point estimate; the bounds are less than one rank point in width in each case.

[52] There is not a clear cut definition of the first affected cohort; a four-year-old child could be too old to benefit from the new policy if parents have already decided that the child will not go to school. Alternately, an eight-year-old child in primary school might benefit if the policy keeps them in school for longer. Online Appendix Figure A7 shows that these results are robust to alternate definitions of the "post" variable; the figure shows estimates equivalent to column 1, but with "post" defined for a range of years from 1960 to 1985. An event-study specification with individual birth years would be desirable, but our sample is much too small to cut the data so finely (see notes to Table 3 for the subgroup sample sizes).

Table 3—Effect of Scheduled Caste Designation on Upward Mobility

	(1)	(2)	(3)
Panel A. Father-son pairs			
Post × *Late SC*	8.432	6.764	
	(1.794)	(1.555)	
1970–79 × *Late SC*			6.739
			(2.025)
1980–89 × *Late SC*			9.649
			(2.580)
Observations	4,502	3,746	4,502
R^2	0.32	0.34	0.32
Panel B. Father-daughter pairs			
Post × *Late SC*	−3.183	−1.604	
	(1.645)	(1.666)	
1970–79 × *Late SC*			−2.658
			(1.727)
1980–89 × *Late SC*			−1.685
			(2.342)
Observations	3,429	3,040	3,177
R^2	0.34	0.33	0.33

Notes: Table 3 shows estimates from Equation 6.1, which describes the impact of Scheduled Caste designation on upward mobility. The dependent variable is the child education rank. The sample consists of SC children of fathers in approximately the bottom 60 percent of the education distribution. *Late SC* is an indicator for jati groups that were added to Scheduled Caste lists in the caste redesignation of 1977. *Post* is an indicator for cohorts born after 1970. The sample in panel A is all men, while panel B is all women. In the column 1 estimation, the are 696 "late SC" men, of whom 438 were born after 1970. For women, these numbers are 516 and 391, respectively. All estimations control for region × cohort, jati × region, jati × cohort, and birth year, and are clustered at the jati and the region levels.

Source: IHDS (2012)

The nature of the natural experiment gives us little information as to which of the many dimensions of affirmative action were important; the direct investments in schooling, the general welfare schemes, or even changes in aspirations following the policy change are candidate results. These are left for future research, which may shed light on why only male members of SC groups appeared to benefit.

B. *Group Differences and Fertility*

Muslims on average have higher fertility than other groups. This section examines whether higher fertility could cause lower mobility for Muslims, perhaps through a household expenditure channel where children with many siblings receive fewer educational inputs.

We calculate individuals' number of siblings based on their mothers' responses to the IHDS women's survey, which has a question about the number of live births. This variable differs from total fertility by excluding children who have died. We only have information on mothers' fertility for children who live with their mothers; we therefore focus on sons in the youngest birth cohort (1985–1989, or ages 23–27), for whom the coresidence rate is highest.[53] The average number of siblings for Muslims is 4.1, compared with 3.0 for both SCs and STs, and 2.6 for Forwards/Others.

As in Section VA, we require a point estimate of upward mobility to use in a regression. We use μ_0^{51}, which can be point estimated as the education of children whose fathers completed two or fewer years of education.[54] We regress this mobility measure on a set of group indicators (Muslim, SC, ST), an urban indicator, and a set of state fixed effects.

The Muslim upward mobility disadvantage is 12 rank points, without adjusting for fertility (online Appendix Table A5, column 2). Controlling for the number of siblings brings the Muslim mobility gap down by 25 percent (column 3) but also reduces the SC mobility gap to an insignificant point estimate of 1.9 rank points.[55] After controlling for fertility, the Muslim-SC mobility gap has shrunk from approximately 9.3 rank points to 7.4.

High fertility can thus explain approximately 20 percent of the Muslim mobility disadvantage relative to SCs in the youngest birth cohorts. This is likely an upper bound, because household income is a direct cause of both children's education and parental fertility (Schultz 2003). Higher fertility can thus explain at most a small share of the present-day mobility disadvantage experienced by Muslims.

C. *Geography and Subgroup Differences*

We examine here whether Muslims live in low-mobility states and districts, or whether they have low mobility after conditioning on place. We recalculate mobility statistics using father and son education ranks *within states* and *within districts*. These estimates describe the outcomes of disadvantaged children relative to others within their own states and districts. If low upward mobility for Muslims is a function of living in districts where everyone has few opportunities, then their within-district mobility gap with Forwards/Others should be substantially smaller than the national mobility gap.

For the father-son mobility gap in 1985–1989, district of residence explains about 18 percent of the Muslim upward mobility gap, 44 percent of the SC upward mobility gap, and 60 percent of the ST mobility gap (online Appendix Figure A8A).[56] The results show that a large part of the mobility disadvantage for STs can be explained by geography—indeed, STs disproportionately live in remote places with low levels of public goods and educational attainment. In contrast, only a small part of Muslim disadvantage can be explained by district effects. These results do not rule out the possibility that finer geographic definitions (such as urban neighborhoods) could

[53] For daughters, coresidence begins to fall rapidly as soon as schooling is finished, leaving too little sample to estimate mobility among coresiders (see online Appendix Figures A1 and A2). Results are similar if we use younger children in the IHDS (e.g., ages 20–23), but we use the higher age bin as it matches the youngest birth cohort in the main analysis.

[54] Results are similar if we use children of fathers with strictly less than two years of education.

[55] An additional sibling is associated with 2.4 fewer rank points in the outcome distribution.

[56] IHDS districts are not representative so these results should be treated with caution; however, the ordering of the changes is the same when we use only within-state ranks—the middle set of bars in online Appendix Figure A8A.

explain a greater share of the mobility gap; unfortunately, higher resolution analysis is not possible with the data available at this time.

D. *Occupations, Returns to Education, and Subgroup Differences*

Can occupational choices and returns to education explain the low and falling upward mobility of Muslims? Online Appendix Figure B shows Mincerian returns to education for the different social groups.[57] Across a range of measures, there is no evidence suggesting that Muslims have lower returns to education than SCs or STs, though both Muslims and SCs have lower returns to education than Forward/Others. Mincerian returns may not reflect the causal effect of education on income and consumption, but they do not suggest that Muslims are choosing less education because their returns are lower.

Muslims are more likely to work as small-scale entrepreneurs than the other major social groups (online Appendix Figure A8C), but the growing Muslim-SC upward mobility gap is the same for entrepreneurial and non-entrepreneurial families (online Appendix Figure A8D). There is thus little evidence to support the idea that the mobility divergence of these groups is a mechanical function of differences in returns to education or occupational preferences.

VI. Conclusion

We have presented a set of tools that are well-suited to measuring intergenerational mobility in developing countries or other contexts where high-quality income data linked across generations are unavailable. Our partial identification approach takes seriously the loss of information given data that report education in coarse rank bins. We propose a measure, bottom-half mobility, which (i) isolates intergenerational mobility from growth and inequality; (ii) is analogous to the popular *absolute upward mobility* measure; (iii) is informative about intergenerational mobility even when education data are very coarse; (iv) has a simple interpretation for population subgroups; and (v) is easy to calculate.

In India, we find that in spite of enormous economic and political changes, upward mobility has barely changed from the 1950s to the 1980s birth cohorts. This lack of change overall reflects substantial gains for men from SCs/STs and substantial losses for Muslims. Our estimate of the causal effect of India's basket of affirmative action policies targeted to SCs suggests the effect of these policies may be large enough to explain the entire SC/Muslim divergence. However, more research is needed to elucidate the factors behind low Muslim mobility in India.

Intergenerational mobility is arguably an important policy objective even in the context of high economic growth and improvements in income for all social groups in modern India. The best opportunities remain scarce, and debates regarding who will be eligible for social programs and positions in universities that provide access

[57] We calculate Mincerian returns by regressing a measure of socioeconomic status (income, wages, or consumption) on individual years of education, age, and age squared for men aged 18–64. Results are robust to using younger ages, including women, and adding controls.

to those opportunities are extremely heated. Moreover, even if there is growth on average, the extent of intergenerational mobility across groups determines whether Muslims, SCs, and STs will occupy a position of permanent disadvantage in the long run. Research investigating the causes of geographic and social group differences in upward mobility has the potential to inform policies that expand the equality of opportunity in India and around the world.

REFERENCES

Ager, Philipp, Leah Boustan, and Katherine Eriksson. 2021. "The Intergenerational Effects of a Large Wealth Shock: White Southerners after the Civil War." *American Economic Review* 111 (11): 3767–94.

Alesina, Alberto, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou. 2023. "Religion and Educational Mobility in Africa." *Nature* 618: 134–43.

Alesina, Alberto, Sebastian Hohmann, Elias Papaioannou, and Stelios Michalopoulos. 2021. "Intergenerational Mobility in Africa." *Econometrica* 89 (1): 1–35.

Alesina, Alberto, Marlon Seror, David Yang, Yang You, and Weihong Zeng. 2020. "Persistence through Revolutions." NBER Working Paper 27053.

Asher, Sam, Paul Novosad, and Charlie Rafkin. 2024. "Replication Data for: Intergenerational Mobility in India: New Measures and Estimates Across Time and Social Groups." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E184504V1.

Azam, Mehtabul. 2016. "Household Income Mobility in India: 1993–2011." IZA Discussion Paper 10308.

Azam, Mehtabul, and Vipul Bhatt. 2015. "Like Father, Like Son? Intergenerational Educational Mobility in India." *Demography* 52 (6): 1929–59.

Bagde, Surendrakumar, Dennis Epple, and Lowell Taylor. 2016. "Does Affirmative Action Work? Caste, Gender, College Quality, and Academic Success in India." *American Economic Review* 106 (6): 1495–521.

Basant, Rakesh, and Abusaleh Shariff. 2010. *Handbook of Muslims in India: Empirical and Policy Perspectives.* Oxford: Oxford University Press.

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4 (3B): S1–S39.

Becker, Sascha O., and Ludger Woessmann. 2009. "Was Weber Wrong? A Human Capital Theory of Protestant Economic History." *Quarterly Journal of Economics* 124 (2): 531–96.

Berenschot, Ward. 2012. *Riot Politics: Hindu-Muslim Violence and the Indian State*. New York: Columbia University Press.

Bhalotra, Sonia and Bernarda Zamora. 2010. "Social Divisions in Education in India." In *Handbook of Muslims in India*, edited by Rakesh Basant and Abusaleh Shariff, 165–95. New Delhi: Oxford University Press.

Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, Vol. 4B, edited by Orley Ashenfelter and David Card, 1487–541. Amsterdam: North Holland.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "Why the Apple Doesn't Fall: Understanding Intergenerational Transmission of Human Capital." *American Economic Review* 95 (1): 437–49.

Blakeslee, David S. 2018. "The Rath Yatra Effect: Hindu Nationalist Propaganda and the Rise of the BJP." Unpublished.

Boserup, Simon Halphen, Wojciech Kopczuk, and Claus Thustrup Kreiner. 2014. "Stability and Persistence of Intergenerational Wealth Formation: Evidence from Danish Wealth Records of Three Generations." Unpublished.

Bratberg, Espen, Jonathan Davis, Bhashkar Mazumder, Martin Nybom, Daniel D. Schnitzlein, and Kjell Vaage. 2017. "A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US." *Scandinavian Journal of Economics* 119 (1): 72–101.

Bratsberg, Bernt, Knut Røed, Oddbjørn Raaum, Robin Naylor, Markus Ja¨ntti, Tor Eriksson, and Eva O¨sterbacka. 2007. "Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons." *Economic Journal* 117 (519): C72–C92.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, Vol. 3A, edited by Orley C. Ashenfelter and David Card, 1801–63. Amsterdam: Elsevier.

Card, David, Ciprian Domnisoru, and Lowell Taylor. 2022. "The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility." *Journal of Labor Economics* 40 (S1): S39–S95.

Cassan, Guilhem. 2019. "Affirmative Action, Education and Gender: Evidence from India." *Journal of Development* 136: 51–70.

Chancel, Lucas, and Thomas Piketty. 2019. "Indian Income Inequality, 1922–2015: From British Raj to Billionaire Raj?" *Review of Income and Wealth* 65 (S1): S33–S62.

Chernozhukov, Victor, Han Hong, and Elie Tamer. 2007. "Estimation and Confidence Regions for Parameter Sets in Econometric Models." *Econometrica* 75 (5): 1243–84.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *Quarterly Journal of Economics* 135 (2): 711–83.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014a. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics* 129 (4): 1553–623.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014b. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review* 104 (5): 141–47.

Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang. 2017. "The Fading American Dream: Trends in Absolute Income Mobility since 1940." *Science* 356 (6336): 398–406.

Connolly, Marie, Miles Corak, and Catherine Haeck. 2019. "Intergenerational Mobility Between and Within Canada and the United States." *Journal of Labor Economics* 37 (S2): S595–S641.

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

Derenoncourt, Ellora. 2022. "Can You Move to Opportunity? Evidence from the Great Migration." *American Economic Review* 112 (2): 369–408.

Emran, M. Shahe, and Forhad Shilpi. 2015. "Gender, Geography, and Generations: Intergenerational Educational Mobility in Post-Reform India." *World Development* 72: 362–80.

Eshaghnia, Sadegh, James J. Heckman, Rasmus Landersø, and Rafeh Qureshi. 2022. "Intergenerational Transmission of Family Influence." NBER Working Paper 30412.

Frisancho, Veronica, and Kala Krishna. 2016. "Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch." *Higher Education* 71 (5): 611–49.

Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies* 60 (1): 35–52.

Güell, Maia, José V. Rodríguez Mora, and Christopher I. Telmer. 2015. "The Informational Content of Surnames, the Evolution of Intergenerational Mobility, and Assortative Mating," *Review of Economic Studies* 82 (2): 693–735.

Hertz, Tom. 2008. "A Group-Specific Measure of Intergenerational Persistence." *Economics Letters* 100 (3): 415–17.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina. 2007. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends" *B.E. Journal of Economic Analysis & Policy* 7 (2): 1–48.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh Paul. 2012. "Castes and Labor Mobility." *American Economic Journal: Applied Economics* 4 (2): 274–307.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh B. Paul. 2013. "Breaking the Caste Barrier: Intergenerational Mobility in India." *Journal of Human Resources* 48 (2): 435–73.

Ito, Takahiro. 2009. "Caste Discrimination and Transaction Costs in the Labor Market: Evidence from Rural North India." *Journal of Development Economics* 88 (2): 292–300.

Jácome, Elisa, Ilyana Kuziemko, and Suresh Naidu. 2021. "Mobility for All: Representative Intergenerational Mobility Estimates over the 20th century." NBER Working Paper 29289.

Khamis, Melanie, Nishith Prakash, and Zahra Siddique. 2012. "Consumption and Social Identity: Evidence from India." *Journal of Economic Behavior & Organization* 83 (3): 353–71.

Khanna, Gaurav. 2020. "Does Affirmative Action Incentivize Schooling? Evidence from India," *Review of Economics and Statistics* 102 (2): 219–33.

Kuran, Timur. 2018. "Islam and Economic Performance: Historical and Contemporary Links." *Journal of Economic Literature* 56 (4): 1292–359.

Lamba, Rohit, and Arvind Subramanian. 2020. "Dynamism with Incommensurate Development: The Distinctive Indian Model." *Journal of Economic Perspectives* 34 (1): 3–30.

Li, Hao, Daniel Millimet, and Punarjit Roychowdhury. 2019. "Measuring Economic Mobility in India Using Noisy Data: A Partial Identification Approach." Unpublished.

Long, Jason, and Joseph Ferrie. 2013. "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *American Economic Review* 103 (4): 1109–37.

Loury, Glenn C. 1981. "Intergenerational Transfers and the Distribution of Earnings," *Econometrica* 49 (4): 843–67.

Maitra, Pushkar, and Anurag Sharma. 2009. "Parents and Children: Education Across Generations in India." Unpublished.

Manski, Charles F., and Elie Tamer. 2002. "Inference on Regressions with Interval Data on a Regressor or Outcome." *Econometrica* 70 (2): 519–46.

McCleary, Rachel M., and Robert J. Barro. 2006. "Religion and Economy." *Journal of Economic Perspectives* 20 (2): 49–72.

Mohammed, A. R. Shariq. 2019. "Does a Good Father Now Have to Be Rich? Intergenerational Income Mobility in Rural India." *Labour Economics* 60: 99–114.

Munshi, Kaivan, and Mark Rosenzweig. 2006. "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy." *American Economic Review* 96 (4): 1225–52.

Narayan, Ambar, Roy Van der Weide, Alexandru Cojocaru, Christoph Lakner, Silvia Redaelli, Daniel Gerszon Mahler, Rakesh Gupta N. Ramasubbaiah, and Stefan Thewissen. 2018. *Fair Progress? Economic Mobility across Generations around the World.* Washington, DC: World Bank Group.

Novosad, Paul, Charlie Rafkin, and Sam Asher. 2022. "Mortality Change among Less Educated Americans." *American Economic Journal: Applied Economics* 14 (4): 1–34.

Platas, Melina R. 2018. "Culture and the Persistence of Educational Inequality: Lessons from the Muslim–Christian Education Gap in Africa." Unpublished.

Rodrik, Dani, and Arvind Subramanian. 2005. "From "Hindu Growth" to Productivity Surge: The Mystery of the Indian Growth Transition." *IMF Staff Papers* 52 (2): 193–228.

Roemer, John E., and Alain Trannoy. 2016. "Equality of Opportunity: Theory and Measurement." *Journal of Economic Literature* 54 (4): 1288–332.

Rosenbaum, Dan T. 2000. "Ability, Educational Ranks, and Labor Market Trends: The Effects of Shifts in the Skill Composition of Educational Groups." Unpublished.

Sachar, Rajindar, Saiyid Hamid, T. K. Oommen, M. A. Basith, Rakesh Basant, Akhtar, Majeed, and Abusaleh Shariff. 2006. *Social, Economic and Educational Status of the Muslim Community of India*. New Delhi, IN: Government of India.

Schultz, T. Paul. 2003. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74 (1): 199–250.

Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*, Vol. 3A, edited by Orley Ashenfelter and David Card, 1761–800. Amsterdam: North Holland Press.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta. 2015. "Education and Human Capital Externalities: Evidence from Colonial Benin." *Quarterly Journal of Economics* 130 (2): 703–57.

Wilkinson, Steven I. 2006. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. Cambridge, UK: Cambridge University Press.