

# Intergenerational Mobility in India: Estimates from New Methods and Administrative Data\*

Sam Asher<sup>†</sup>

Paul Novosad<sup>‡</sup>

Charlie Rafkin<sup>§</sup>

March 2020

First version: February 2016

## Abstract

Estimating intergenerational mobility in developing countries is difficult because matched parent-child income records are unavailable and education is measured coarsely. Moreover, traditional measures like the rank-rank coefficient are not useful for studying subgroup mobility. We develop a partial identification method that resolves these issues and apply it to the study of upward mobility in India, using new administrative data. Intergenerational mobility for the population as a whole has remained constant since liberalization, but cross-group changes have been substantial. Rising mobility for Scheduled Caste boys is almost exactly offset by declining mobility among Muslims boys, a comparably sized group with few constitutional protections. These findings contest the conventional wisdom that marginalized groups in India have been catching up on average. Upward mobility is slightly lower for girls than for boys, but the gender gap is heterogeneous across space and social group. Muslim disadvantage cannot be explained by occupational patterns, fertility, differential returns to education, or location. However, affirmative action for Scheduled Castes (but not Muslims) has delivered substantial upward mobility gains for this group. We generate high-resolution geographic measures of intergenerational mobility across 5600 rural subdistricts and 2300 cities and towns. On average, children are most successful at exiting the bottom of the distribution in places that are southern, urban, or have higher average education levels.

---

\*We are thankful for useful discussions with Alberto Abadie, Emily Blanchard, Raj Chetty, Eric Edmonds, Shahe Emran, Francisco Ferreira, Nate Hilger, David Laibson, Ethan Ligon, Erzo Luttmer, Whitney Newey, Elias Papaioannou, Nina Pavcnik, Bruce Sacerdote, Na'ama Shenhav, Forhad Shilpi, Gary Solon, Bob Staiger, Doug Staiger, Chris Snyder and Elie Tamer, among others. Annaka Balch, Ali Champion, Toby Lunt, Ryu Matsuura, and Taewan Roh provided excellent research assistance. This project received financial support from the IZA GLM-LIC program and from the National Science Foundation Graduate Research Fellowship under Grant No. 1122374.

<sup>†</sup>Johns Hopkins University, sasher2@jhu.edu

<sup>‡</sup>Dartmouth College, paul.novosad@dartmouth.edu, corresponding author

<sup>§</sup>MIT, craffin@mit.edu

## 1 Introduction

The intergenerational transmission of economic status, a proxy for equality of opportunity, has implications for inequality, allocative efficiency and subjective well-being (Solon, 1999; Black and Devereux, 2011; Chetty et al., 2014a; Chetty et al., 2017). Studies of intergenerational mobility (IM) in developing countries have had less influence than in rich countries because of methodological challenges and an absence of data comparable to the administrative tax records that underlie much of the recent literature in richer countries. In this paper, we develop a set of measures of intergenerational educational mobility that perform well under the data constraints often faced in developing countries. Using sample survey and new administrative data, we apply these measures to the analysis of IM in India across time, across major population subgroups, and across space.

Because of data quality and availability, as well as the challenge of measuring individual income in households with joint production, studies of intergenerational mobility in developing countries (and in historical contexts) often use education as a proxy for social status.<sup>1</sup> A key challenge in the measurement of educational mobility is that education data is often coarsely measured; for instance, for the 1960–69 birth cohort in India, over 50% of fathers and 80% of mothers report a bottom-coded level of education. This makes it difficult to use recent rank-based measures of mobility (Chetty et al., 2014a), which require observing parents at specific percentiles in the socioeconomic status distribution. Studies of educational mobility have instead focused on estimators like the correlation coefficient between parents’ and children’s educational outcomes. These linear estimators have several limitations (discussed in Section 4.1), the chief of which is that they are not meaningful for subgroup analysis, because they measure individuals’ progress only against other members of their own group (Hertz, 2005).

In this paper, we show that with education data, the new generation of rank-based estimators of intergenerational mobility can at best be partially identified. Intuitively, conventional mobility estimators applied to educational mobility do not account for the loss of information associated with coarse measurement of ranks; they instead rely on untested assumptions about the unobserved rank data when

---

<sup>1</sup>Recent studies of intergenerational mobility focusing on education and occupation include Black et al. (2005), Güell et al. (2013), Wantchekon et al. (2015), Card et al. (2018), Derenoncourt (2018), and Alesina et al. (2019). More are summarized in Black and Devereux (2011).

generating rank-based estimates. Treating this as an interval data problem, we show that mobility can at best be partially identified, by considering best-/worst-case assumptions about the underlying distributions that can generate the observed data (Manski and Tamer, 2002; Novosad et al., 2020).

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 similar interpretation to other measures of upward mobility, but it can be bounded tightly even in contexts with extreme interval censoring.<sup>2</sup> In contrast, conventional measures have bounds that are too wide to be meaningful once they account for the underlying uncertainty associated with interval data. This paper thus relaxes the hidden assumptions underlying most mobility estimators in settings with education data and still obtains precise (if partially-identified) mobility estimates. To our knowledge, bottom half mobility is the first measure of intergenerational educational mobility that can be meaningfully compared across time and space, across countries, and across population subgroups.

We apply our measure to India, using data from the 2012 India Human Development Survey (IHDS), and the 2012 Socioeconomic and Caste Census (SECC). The former is a sample survey, and the latter is a socioeconomic census with high geographic resolution covering all individuals in the country. We document trends in educational mobility from the 1950–59 to the 1985–89 birth cohorts. We focus on measuring mobility from fathers to sons/daughters; mobility from mothers to children cannot be bounded tightly because the bottom-coding of mothers’ education is so severe.

We present three primary findings. First, upward mobility has remained constant for the past thirty years, in spite of dramatic gains in average levels of education and income. An Indian son born in the bottom half of the parent education distribution in 1985–89 (our youngest cohort) can expect to obtain the 37th percentile; a daughter obtains percentile 35.5.<sup>3</sup> A similar child in the U.S., which has

---

<sup>2</sup>Absolute upward mobility Chetty et al. (2014a) describes the expectation of a child outcome, conditional on that child being born to a parent at the 25th percentile. Bottom half mobility describes the expectation of a child outcome, conditional on that child being born to a parent between percentiles 0 and 50. If the conditional expectation function is linear in parent rank, the two measures are identical. If the CEF is concave (like many parent-child income CEFs), then bottom half mobility puts more weight on the outcomes of the least privileged children.

<sup>3</sup>Following prior work, we always rank children in the own-gender distribution.

low intergenerational mobility by OECD standards, on average attains education percentile 41.7.<sup>4,5</sup>

Second, we find significant changes in the cross-group distribution of upward mobility over time, particularly among boys. We divide the population into Scheduled Castes, Scheduled Tribes, Muslims, and Forwards/Others.<sup>6</sup> Consistent with prior work (Hnatkovska et al., 2012; Emran and Shilpi, 2015), we find that boys from India’s constitutionally protected marginalized groups, the Scheduled Castes and Tribes, have closed respectively 50% and 30% of the mobility gap to Forwards/Others. In contrast, upward mobility for Muslim boys has steadily declined from the 1960s to the present. The expected educational rank of a Muslim boy born in the bottom half of the parent distribution has fallen from between 31 and 34 to a dismal 29. Muslim boys now have considerably worse upward mobility today than both Scheduled Castes (38) and Scheduled Tribes (33), a striking finding given that STs tend to live in much more remote and low mobility areas than Muslims. The comparable figure for U.S. black men is 34. Higher caste groups have experienced constant and high upward mobility over time, a result that contradicts a popular notion that it is increasingly difficult for higher caste Hindus to get ahead.

Our measurements for father-daughter mobility are less precise, but the same subgroup patterns do not appear to hold. Girls from poor Muslim, SC, and ST households all have persistently lower mobility than Forwards/Others, and there is minimal convergence over the sample period.

Third, we describe substantial variation in upward mobility across 5600 rural subdistricts and 2000 cities and towns. Paralleling results from Chetty et al. (2014b), we find substantial heterogeneity even within small geographic regions. Upward mobility is highest in urban areas, and in places with high consumption, education, school supply, and manufacturing employment, all broad correlates of development. High mobility is inversely correlated with caste segregation and land inequality.

---

<sup>4</sup>In a society where children’s outcomes are independent of parents (i.e. total mobility), a child born in the bottom half of the distribution obtains the 50th 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 25th percentile.

<sup>5</sup>All of our mobility estimates are robust to different data construction methods, and we take care to show that survivorship bias, migration, or bias in estimates from coresident parent-child households are unlikely to substantially affect our results.

<sup>6</sup>We include non-Muslim OBCs in the “Others” category. Measuring OBC mobility is challenging because OBC definitions are less stable over time, are sometimes inconsistently classified between federal and state lists, and may be reported inconsistently by the same individual over time. These concerns apply to SC and ST groups, but at a considerably smaller scale. The very small number of Muslim SC/STs are categorized as Muslims; reclassifying them as SCs or STs, or excluding Sikhs, Jains and Christians from the “others” category do not affect our results.

Geography-subgroup interactions are important; for instance, girls have higher mobility than boys in urban areas, but lower in rural areas.

The final section of the paper examines several potential mechanisms for the divergence of Scheduled Castes from Muslims over the last twenty-five years. We find that this divergence cannot be explained by differential returns to education, occupational patterns, geography, or differential fertility. However, the basket of affirmative action policies targeted to India’s scheduled groups appears to have had a substantial impact on their mobility. Following Cassan (2019), we exploit a natural experiment that added many castes to the Scheduled Caste lists in 1977. We show that when a social group (jati) gets assigned to Scheduled Caste status, it experiences a 6–7 rank point increase in upward mobility over twenty years; this is more than two thirds of the mobility gap that has opened between Muslims and Scheduled Castes over the same period. This finding is consistent with the possibility that educational quotas, government job reservations, and other affirmative action policies (which benefited Scheduled Castes but not Muslims) are a driver of the increasing upward mobility gap between Scheduled Castes and Muslims.

Our paper’s contributions are both methodological and empirical. Methodologically, we develop methods for the measurement of intergenerational mobility that address the data limitations often faced in developing country and historical contexts. Bottom half mobility is the first educational mobility measure that is valid for comparing subgroups across different contexts. Prior researchers have used CEF-based mobility measures to examine subgroup outcomes, but the coarse measurement problem has forced them to use inconsistent measures over time. For example, Card et al. (2018) and Derenoncourt (2018) define upward mobility in the 1920s as the 9th grade completion rate of children whose parents have 5–8 years of school (or approximately parent percentiles 30–70) — they then compare this measure with  $p_{25}$  (i.e. children of parents at the 25th percentile) in the present. Alesina et al. (2019) define upward mobility in Africa as the likelihood that a child born to a parent who has not completed primary school manages to do so. While this measure captures the ability of poor children to exceed the education levels of their parents, it does not distinguish between average educational gains and changes in the ability of individuals to move up the socioeconomic distribution

in relative terms, an essential component of mobility.<sup>7</sup> In contrast, our measure precisely isolates individuals' ability to move up in the rank distribution across generations.

Empirically, we present several previously unknown facts about upward mobility in India. Our most striking finding is that Muslims are losing substantial ground in intergenerational mobility, and currently have lower mobility than either Scheduled Castes or Scheduled Tribes. Given the large amount of work studying access to opportunity among Scheduled Castes, our results highlight the importance of studying economic outcomes of Muslims in India, especially from poor families. Despite a population share comparable to Scheduled Castes, this group is overlooked by much of the economic literature on marginalized groups in India.<sup>8</sup> We also present causal evidence that the combination of affirmative action policies targeting Scheduled Castes has increased their intergenerational mobility.

Our findings imply that virtually all of the upward mobility gains in India over recent decades have accrued to Scheduled Castes and Tribes, groups with constitutional protections, reservations in politics and education, and who have been targeted by many development policies. There is no evidence that any of these gains have come at the expense of higher caste groups. For non-scheduled groups, there is little evidence that economic liberalization has made it easier to gain in relative status across generations, and for Muslim men, these opportunities have substantially deteriorated.

These patterns have to our knowledge not been identified because earlier studies have either (i) focused on 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). Studies on affirmative action in India have identified improvements in SC/ST access to higher education but have not examined impacts on Muslims (Frisancho Robles and Krishna, 2016; Bagde et al., 2016); our findings

---

<sup>7</sup>This measure also conditions on substantially different parts of the education distribution in different times and places; we show in Section 4 that this measure corresponds to, for example,  $E(y > 52 | x \in [0, 76])$  in Mozambique (where 76% of parents and 48% of children have not completed primary), but to  $E(y > 18 | x \in [0, 42])$  in South Africa.

<sup>8</sup>Notable exceptions include Khamis et al. (2012) and Bhalotra and Zamora (2010), who note poor education outcomes among Muslims. The Sachar Committee Report (2006) and Basant et al. (2010) summarize some recent research on Muslims in India, none of which addresses intergenerational mobility.

point to the importance of studying the effects of such policies on a wider set of marginalized groups.<sup>9</sup>

Our paper proceeds as follows. Sections 2 and 3 describe the context and data sources. Section 4 describes our methodological innovation in the context of other measures of intergenerational educational mobility. Section 5 presents results on national and cross-group mobility trends, the geographic distribution of intergenerational mobility, and our analysis of mechanisms. Section 6 concludes.

## **2 Context on Scheduled Castes, Scheduled Tribes, and Muslims**

While intergenerational mobility is of interest around the world, India's caste system and high levels of inequality make it a particularly important setting for such work. 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. Since independence in 1947, the government has systematically implemented policies intended to reduce the disadvantage of communities that are classified as Scheduled Castes or Scheduled Tribes. These groups are targeted by a range of government programs and benefit from reservations in educational and political institutions. If effective, these policies would substantially improve the intergenerational mobility of targeted groups.

India's Muslims constitute a similar population share as the Scheduled Castes and Scheduled Tribes (14% for Muslims vs. 17% for SCs and 14% 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. On the contrary, a large scale social movement (the Rashtriya Swayamsevak Sangh, or RSS) and several major political parties have successfully 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).

The last 30 years have seen tremendous growth in market opportunities in India as well as in educational attainment. While some have argued that economic growth is making old social and

---

<sup>9</sup>In an analogous finding, Bertrand et al. (2010) find that when Indian colleges intentionally select lower caste students, they end up admitting fewer women.

economic divisions less important to the economic opportunities of the young, caste remains an important predictor of economic opportunity (Munshi and Rosenzweig, 2006; Ito, 2009; Hnatkovska et al., 2013; Mohammed, 2016). Understanding how mobility has changed for these population groups is thus an essential component of understanding secular trends in intergenerational mobility in India. Further, whether economic progress can overcome traditional hierarchies of social class and religion is a central question for both India and the broader world.

### 3 Data

We describe the data in advance of the methods section, because data constraints in studying intergenerational mobility are a major motivator for our method.

We draw on two datasets that report matched parent-child educational attainment. The India Human Development Survey (IHDS) is a nationally representative survey of 41,554 households, with rounds in 2004-05 and 2011-12. Crucially, the IHDS solicits information on the education of parents of many respondents, even if those parents are not resident in the household. Many prior studies have estimated mobility by looking only at households where the parent and child live together. These estimates may be substantially biased as a minority of children over the age of 25 live with their parents, with the share shrinking significantly for women and at higher ages (Appendix Figure A1).

The IHDS records the father of the household head, and both the mother and father of women selected for the special women’s survey, as well as of their husbands. The women’s survey is a separate survey module that is given to one or two women aged 15–49 in each household. We created additional parent-child links using information from the household roster.<sup>10</sup> In many cases, a parent’s education is recorded in multiple ways, so we can cross-check the validity of the responses, which is high.<sup>11</sup>

The IHDS also identifies religion and Scheduled Tribe or Scheduled Caste status. We classify SC/ST Muslims, who make up less than 2% of SC/STs, as Muslims. About half of Muslims are Other

---

<sup>10</sup>For example, we can identify the education of the household head’s sister’s father using the household head’s response to the father education question.

<sup>11</sup>For example, the household head’s father’s education may be obtained from (i) the household roster (if he is coresident); (ii) from the household head’s response to the father education question; and (iii) from his wife’s responses to the husband’s father’s education question. The average correlation between pairs of responses is 0.9. Appendix Table A2 shows that the discrepancies between measures are not correlated with household characteristics.



Backward Castes (OBCs); we classify these as Muslims. We do not consider OBCs as a separate category in this paper, because OBC status is inconsistently reported across surveys, due to both misreporting and changes in the OBC schedules. Analysis of mobility of OBCs will therefore require detailed analysis of subcaste-level descriptors and classifications which are beyond the scope of the current work. We pool Christians, Sikhs, Jains and Buddhists, who collectively make up less than 5% of the population, with higher caste Hindus (*i.e.* forward castes and OBCs); we describe this group as “Forward/Other.” We find broadly similar results if we exclude these other religions from the sample.

Our primary dataset for time series is the 2011–2012 IHDS. We estimate a time series in mobility by drawing on the parent-child education links of older birth cohorts. To allay concerns that differential mortality across more or less educated fathers and sons might bias our estimates, we replicate our analysis on the *same* birth cohorts using the IHDS 2005. By estimating mobility on the same cohort at two separate time periods, we identify a small survivorship bias for the 1950–59 birth cohort (reflecting attrition of high mobility dynasties), but zero bias for the cohorts from the 1960s forward. Our results of interest largely describe trends from the 1960s forward, so survivorship bias among the oldest cohorts does not influence any of our conclusions. We pool the data into 10-year birth cohorts for 1950–69, and 5-year birth cohorts for 1970–1989 where we have more power. Because of the age restriction on the women’s survey, we do not observe links for mothers or daughters for the 1950–59 birth cohort.

The strengths of the IHDS for our study that it documents religion and records the education of non-coresident parents. However, the IHDS sample is too small to study geographic variation in much detail.

We therefore additionally draw on the 2011–12 Socioeconomic and Caste Census (SECC), an administrative socioeconomic database covering all individuals in the country that was collected to determine eligibility for various government programs. The data were posted to the internet by the government, with each village and urban neighborhood represented by hundreds of pages in PDF format. Each town/village was posted for only ninety days. Over a period of two years, we scraped over two million files, parsed the embedded data into text, and translated the text from twelve different Indian languages into English. The data include age, gender, education, an indicator for Scheduled

Tribe or Scheduled Caste status, and relationship with the household head. Assets and income are reported at the household rather than the individual level, and thus cannot be used to estimate mobility. Religion and subcaste were recorded but not included with the publicly posted data, and are therefore not available.<sup>12</sup> The SECC provides the education level of every parent and child residing in the same household. Non-coresident children cannot be linked to their parents' education in the SECC.

The scale of the SECC makes it possible to calculate upward mobility with high geographic precision, but the limitation is that we do not observe parent-child links for children who no longer live with their parents. We therefore focus on children aged 20 to 23, a set of ages where education is likely to be complete but coresidence is high enough that the bias from excluding non-coresident pairs is low. We selected these ages by examining the coresidence share at each child age, and examining the extent to which upward mobility estimates in the IHDS would be biased by excluding non-coresident parent-child pairs. Appendix Figure A2 plots this bias for father-son and father-daughter pairs, where vertical lines mark the sample selected for the SECC. For the pooled 20–23 age group, we can rule out a bias of more than two percentage points in the child rank for father-son pairs, but father-daughter mobility estimates based only on coresident households are already biased by 5 percentage points in this age range—more than the mobility difference between the United States and Finland. The bias is larger for girls because many have left home for marriage even by the age of 18; by age 23, only 53% are still living with their fathers, compared with 80% of sons.<sup>13</sup> At higher ages, the mobility bias rapidly grows for both men and women, justifying the use of this narrowly-aged sample. Earlier Indian mobility estimates which were based on coresident children as old as 40 should thus be treated with caution. The SECC sample thus consists of 31 million young men and their fathers. For the coresident father-son pairs that are observed in both datasets, IHDS and SECC produce similar point estimates for upward mobility.

Because of the strengths and limitations of each dataset, we use the SECC to study cross-sectional geographic variation in mobility, and we use the IHDS to study mobility differences across groups and across time.

---

<sup>12</sup>Additional details of the SECC and the scraping process are described in Asher and Novosad (2019).

<sup>13</sup>Appendix Figure A1 shows coresidence rates at all ages.

Education is reported in seven categories in the SECC.<sup>14</sup> IHDS records completed years of education. To make the two data sources consistent, we recode the SECC into years of education, based on prevailing schooling boundaries, and we downcode the IHDS so that it reflects the highest level of schooling completed, *i.e.*, if someone reports thirteen years of schooling in the IHDS, we recode this as twelve years, which is the level of senior secondary completion.<sup>15</sup> The loss in precision by downcoding the IHDS is minimal, because most students exit school at the end of a completed schooling level.

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. The youngest cohort in this study was born in 1989; cohorts born in the 1990s may not have completed their education at the time that they were surveyed and are therefore excluded.

## 4 Methods: Measuring Mobility in Developing Countries

When intergenerational mobility is low, the social status of individuals is highly dependent on the social status of their parents (Solon, 1999). In more mobile societies, individuals are less constrained by the circumstances of their birth. A growing literature, facilitated by new administrative datasets, has documented differences in intergenerational mobility across countries, across groups within countries, and across time.<sup>16</sup>

In this section, we present a brief background on the measurement of intergenerational mobility in various contexts. We then present the key challenge for using modern mobility measures in developing country contexts. We show that rank-based mobility measures can at best be partially identified, and that a novel measure of upward mobility, *bottom half mobility*, can provide very tight estimates while conventional measures are too coarse to be meaningful.<sup>17</sup>

---

<sup>14</sup>The categories are (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. These are standard categories used in many of India’s surveys, including the National Sample Surveys.

<sup>15</sup>We code the SECC category “literate without primary” as two years of education, as this is the number of years that corresponds most closely to this category in the IHDS data, where we observe both literacy and years of education. Results are not substantively affected by this choice.

<sup>16</sup>See Hertz et al. (2008) for cross-country comparisons, and Solon (1999), Corak (2013), Black and Devereux (2011), and Roemer (2016) for review papers.

<sup>17</sup>Code and documentation to generate all the measures used in this paper are available at

## 4.1 Background: Measurement of Intergenerational Mobility

The first generation of intergenerational mobility studies described matched parent-child outcome distributions with a single linear parameter, such as the correlation coefficient between children’s earnings and parents’ earnings (Solon, 1999; Black and Devereux, 2011). Such gradient measures are easy to calculate and interpret and they form the basis of studies in dozens of countries. The three main limitations of gradient measures are that (i) they do not distinguish between changes in opportunity at the top and the bottom of the distribution;<sup>18</sup> and (ii) they are not well-suited for between-group comparisons. The parent-child outcome gradient in a population subgroup effectively compares children’s outcomes against more advantaged members of their own group. A subgroup can therefore have a lower gradient (suggesting higher mobility) and yet worse outcomes at every point in the parent distribution. For a striking example, see Chetty et al. (2018), who show that the parent-child income rank gradient is virtually identical for whites and blacks, indicating that black children suffer the same rank disadvantage at every point in the parent rank distribution.

More recent studies have analyzed the entire joint parent-child income distribution, and in particular, the conditional expectation function of child rank given parent rank (Boserup et al., 2014; Chetty et al., 2014a; Bratberg et al., 2015; Hilger, 2016). Studying transitions in ranks is useful because it holds constant any changes in average income levels or inequality over time, isolating individuals’ relative movement in the social hierarchy. The most widely used measure in the new generation of studies is *absolute upward mobility*, defined as expectation of a child’s income rank, conditional on having a parent at the 25th income percentile, or  $p_{25} = (E(y|x=25))$ , where  $y$  is the child rank and  $x$  is the parent rank. As noted by Chetty et al. (2014a), this describes the expected rank of a child born to the median parent in the bottom half of the parent rank distribution.<sup>19</sup>

<https://github.com/paulnov/nra-bounds/>.

<sup>18</sup>This weakness is consequential only if the conditional expectation of the child outcome given the parent outcome is non-linear. But many studies find non-linear CEFs (Bratsberg et al., 2007; Boserup et al., 2014; Bratberg et al., 2015); the linearity of the income rank CEF described by Chetty et al. (2014a) is an exception. The Indian education rank CEF that we study below is clearly non-linear.

<sup>19</sup>Chetty et al. (2014a) use the term “absolute mobility” because this measure does not depend on the value of the CEF at any other point in the parent rank distribution, distinguishing it from the rank-rank correlation which they describe as a *relative* mobility measure. Other authors use the term “absolute mobility” to describe the set of mobility measures which use child levels as outcomes rather than child ranks. To avoid confusion, we use the

Unlike the gradient estimators, absolute upward mobility is valid for cross-group comparisons: with  $p_{25}$ , we can compare outcomes of children in different population subsets who are born to parents with similar incomes. It is also comparable across contexts or countries with different levels of income or income inequality. However, as we discuss below, methodological challenges have made it difficult to use this measure in developing countries.

## 4.2 Educational Mobility and Income Mobility

In the study of upward mobility in developing countries, education is often a better proxy of social status than income, for three reasons.

First, matched parent-child education data are more widely available than matched income data. Because education is fixed early in life and parents' education is generally known by the next generation, it is possible to record matched parent-child education data even when the parent is no longer alive. Second, income is subject to significant measurement error in developing countries. Transitory incomes are noisy estimates of lifetime income, especially in agriculture, and depend on an individual's place in the life cycle. Subsistence consumption is difficult to measure, and many individuals report zero income. Measurement error in income is consequential because it biases mobility estimates upward (Zimmerman, 1992). In contrast, education levels are measured more precisely than income, and rarely change in adulthood, mitigating life cycle bias.

Third, permanent income is difficult to assign to households with joint production processes, like many of the rural poor. For a large number of individuals, like subsistence farmers, income arises from a joint household production function, of which reported wages represent a small or zero share. When these households span multiple generations, it is virtually impossible to distinguish between parent income and child income. Education, in contrast, is directly ascribed to an individual.

For these reasons, studies of intergenerational mobility in developing countries, as well in historical periods in developed countries, have focused on the intergenerational persistence of social status as measured by educational attainment (Solon, 1999; Güell et al., 2013; Wantchekon et al., 2015; Card et al., 2018; Derenoncourt, 2018; Alesina et al., 2019). But studies of educational mobility have

---

term "absolute mobility" only when making reference to the specific measure used by Chetty et al. (2014a).

thus far not come up with a satisfactory analog to absolute upward mobility.

The key challenge is that the coarse measurement of educational completion makes it impossible to identify a parent at a precise education percentile. This challenge is demonstrated in Figure 1A, which shows the average child education rank in each parent education rank bin, for two Indian birth cohorts: 1960–69 (circles) and 1985–89 (x’s). The solid and dashed vertical lines respectively show the boundary for the bottom-coded education bin in the two cohorts. In the 1960–69 birth cohort, a full 58% of fathers report a bottom-coded education level; in the 1985–89 cohort, this figure is 36%. How does one identify the expected child rank given a parent at the 25th percentile in these birth cohorts?<sup>20</sup>

Figure 1B shows two conditional expectation functions that are both perfectly consistent with the 1960–69 moments. The data available cannot distinguish between these two functions, but they have very different implications for upward mobility: the flatter of the two functions implies a much higher expected rank for a child growing up at the bottom of the distribution. As conveyed by this example, the CEF of child rank given parent rank can be at best partially identified by the data; Section 4.3 below shows that we can nevertheless recover useful estimates of upward mobility from the data.

Coarse measures of parental status like these are extremely common in the mobility literature. In older generations in developing countries, bottom-coding rates in excess of 50% are widespread (Narayan and Van der Weide, 2018), as in older generations in richer countries. Internationally comparable censuses often report education in as few as four or five categories. Income mobility is also often based on censored estimates; in the well-known British Cohort Study, one income bin contains more than 30% of the data. Table 1 reports the number of parent education bins used in a set of recent studies of intergenerational mobility from several rich and poor countries. Many studies observe education in fewer than ten bins; the population share in the bottom bin is often above 20%, and in many developing countries it is above 50%.

Comparing transition matrices over time poses a similar problem to that of estimating  $p_{25}$ : when

---

<sup>20</sup>The graph makes self-evident why we cannot assume that the expected child rank given a parent at the 25th percentile is equivalent to the expected child rank given a parent in the bottom education bin. The size of the bottom bin is a function of the granularity of the education coding and the national distribution of education; when this bin is smaller, children in that bin are more negatively selected and will have mechanically lower education. In the case of Figure 1A, the path traced by the CEFs of the two birth cohorts look highly similar in rank space, but the expected child rank in the bottom bin is mechanically lower in the 1985–89 birth cohort because of the smaller bin size.

education bin boundaries are changing over time (due to changes in the education distribution), there may be few cells of the matrix that can be directly compared in different years. Using a quantile transition matrix would resolve the problem of comparability, but calculating the quantile transition matrix from coarsely binned rank data poses the same challenge as calculating  $p_{25}$ .

### 4.3 Estimating Bounds on a CEF with Censored Rank Data

In this section, we apply the method of Novosad et al. (2020) to bound the conditional expectation function of child rank given parent rank (henceforth the CEF), given coarsely measured education data. Many measures of mobility, including absolute upward mobility and the rank-rank gradient, are functions of the CEF and can be similarly bounded. We will show that one function of the CEF, which we call bottom half mobility, is particularly useful in studying upward mobility. The methodological approach is based on Novosad et al. (2020), where we are concerned with identifying  $E(y|x=i)$ , where  $y$  is adult mortality and  $x$  is the adult education rank. The methodological contribution of this paper is to use these partial identification tools to solve the problem of measuring mobility with coarse data.

Formally, our goal is to measure  $E(y|x=i)$ , where  $y$  is a child outcome (primarily the child education rank in this paper) and  $x$  is a parent rank. We only observe that  $x$  lies within some interval  $[x_k, x_{k+1}]$ , where  $k$  indexes bins. The mean value of  $E(y|x)$  in bin  $k$  is  $r_k$ , which is observed.

We require only two assumptions to derive bounds on  $E(y|x)$ . First, we assume that there is a latent continuous parent education rank distribution; this implies a meaningful but unobserved continuous ranking of parents' education *within* each observed education category. Formally:

$E(y|x=i)$  is a continuous function with support for all values of  $i \in [0,100]$ . (Assumption 1)

This assumption arises directly from a standard human capital model where observed differences in education levels reflect individual differences in costs and benefits of seeking education. The latent education rank  $x$  reflects the amount that 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 the next

level. Individuals who are at the margin of obtaining the next level of education (*i.e.* they would need only a small increase in marginal benefit in order to do so) have the highest educational ranks within their rank bin. Individuals who would not advance further even if the net benefit changed a great deal will have the lowest ranks in the bin. Alternatively,  $x$  represents what an individual’s education rank would be if education choices were fully continuous and observed as such. Assumption 1 implies that the calculation of  $E(y|x=i)$  can be treated as an interval censoring problem (Manski and Tamer, 2002).

Our second assumption is that the expectation of a child’s education level is weakly increasing in the latent parent education rank. In other words, having a more advantaged parent cannot make a child worse off. Formally:

$$E(y|x) \text{ must be weakly increasing in } x. \quad (\text{Assumption 2})$$

Empirically, average socioeconomic outcomes of children are strongly monotonic in parent socioeconomic outcomes across many socioeconomic measures and countries (Dardanoni et al., 2012). Average child education is also monotonic in parent education across nearly all of the subgroup-cohorts that we study in India (see Appendix Table A1).<sup>21</sup> Note that the conventional linear estimation of educational mobility is also implicitly imposing monotonicity.

We now apply Proposition 1 (derived in Novosad et al. (2020)) to obtain sharp bounds on the mobility CEF.

**Proposition 1** (Novosad et al., 2020). *Under assumptions 1 and 2 and without additional infor-*

---

<sup>21</sup>In practice, when education data are highly granular, non-monotonicity may emerge from monotonic distributions due to sampling error. While this occurs for a minority of subgroup-cohorts in our data, it only occurs at the very top of the distribution, where bins are very small, and thus does not affect our calculations of upward mobility, which are focused on expected child ranks in the bottom half of the parent distribution. To use our analytical method in cases where non-monotonicity may be caused by sampling error, we advise pooling small bins into larger bins until monotonicity is restored.



mation, the following bounds on  $E(y|x)$  are sharp:<sup>22</sup>

$$\begin{cases} r_{k-1} \leq E(y|x) \leq \frac{1}{x_{k+1}-x_k}((x_{k+1}-x_k)r_k - (x-x_k)r_{k-1}), & x < x_k^* \\ \frac{1}{x-x_k}((x_{k+1}-x_k)r_k - (x_{k+1}-x)r_{k+1}) \leq E(y|x) \leq r_{k+1}, & x \geq x_k^* \end{cases}$$

where

$$x_k^* = \frac{x_{k+1}r_{k+1} - (x_{k+1}-x_k)r_k - x_k r_{k-1}}{r_{k+1} - r_{k-1}}.$$

Figure 2A shows the bounds on the CEF for the 1960–69 and the 1985–89 birth cohorts. While the bounds are tight in parts of the CEF where the education bins are small, they are very wide in the bottom half of the distribution where the data is heavily interval-censored. Note that absolute upward mobility ( $p_{25}$ ) can be read directly from this graph by examining the bounds of the CEF at the 25th percentile. In this case, they are far too wide for both cohorts to be meaningful.

We can obtain tighter bounds by imposing a constraint on the curvature of the CEF, equivalent to assuming that a marginal change in parent education rank cannot result in a dramatic change in the slope of the CEF. Figure 2B shows bounds on the CEF under a conservative curvature constraint. The bounds are tighter, but the bounds on  $p_{25}$  remain too wide to be meaningful. While the constrained curvature assumption is defensible (see Novosad et al. (2020)), the bounds on our primary mobility measures below are only marginally improved by the curvature constraint, so we take a parsimonious approach and present all results below with unrestricted curvature.

We can take a similar approach to bound a function of the CEF. We focus on two functions in particular. The first is the slope of the best linear approximator to the CEF (or the rank-rank gradient), which is analogous to the linear estimator used in the bulk of the prior educational mobility literature. The second is a new function, which we call *bottom half mobility*, defined as  $\mu_0^{50} = E(y|x \in [0, 50])$ . This measure describes the expected outcome of a child born to parents in the bottom half of the parent distribution.  $\mu_0^{50}$  is thus analogous to  $p_{25}$ , which describes the expected outcome of a child born to the median parent in the bottom half of the parent distribution.  $\mu_0^{50}$  can

---

<sup>22</sup>Appendix B presents the proof formally following Novosad et al. (2020) and works through a graphical example.

be calculated analytically, while we calculate the rank-rank gradient using a numerical optimization; both of these approaches are described in detail in Appendix B.

#### 4.4 Comparing Bounds on Different Functions of the CEF

Panels A through C of Figure 3 respectively show bounds on the rank-rank gradient (denoted  $\beta$ ), absolute upward mobility ( $p_{25}$ ), and bottom half mobility ( $\mu_0^{50}$ ) for the 1960–69 and the 1985–89 birth cohorts. As benchmarks, we also show mobility estimates for USA and Denmark.<sup>23</sup>

The bounds on  $\beta$  and  $p_{25}$  are not informative either in levels or in changes; they are consistent with both major declines and major increases in mobility from 1960–85. In contrast,  $\mu_0^{50}$  can be bounded tightly in the 1960s and can be nearly point estimated in the 1980s. On the basis of  $\mu_0^{50}$ , we can clearly distinguish upward mobility in India from the U.S. and Denmark (India is about as far below the U.S. as the the U.S. is below Denmark), and we can reject substantial changes over the sample period.

The figure shows the key advantage of bottom half mobility, or  $\mu_0^{50}$ : it can be tightly bounded even with severely interval-censored rank data. As noted,  $\mu_0^{50}$  has a very similar interpretation to  $p_{25}$ . If the CEF is linear,  $p_{25}$  is equal to  $\mu_0^{50}$ ; if the CEF is concave at the bottom of the parent distribution, then Jensen’s Inequality implies that  $\mu_0^{50}$  will be lower than  $p_{25}$ , reflecting the greater persistence of bad outcomes at the bottom of the distribution. For non-linear CEFs,  $\mu_0^{50}$  thus effectively puts more weight on outcomes at the very bottom of the parent distribution than  $p_{25}$ .

For intuition behind the tight bounds on  $\mu_0^{50}$ , note that  $\mu_a^b$  is point-identified when  $a$  and  $b$  correspond to intervals in the data. If  $X\%$  of parents are in the bottom-coded rank bin, then  $\mu_0^X$  is the expected child rank, conditional on having parents in the bottom-coded bin. In general,  $\mu_a^b$  will be tightly bounded when  $a$  and  $b$  are close to bin boundaries in the data, by virtue of the continuity of the CEF and uniformity of the rank distribution. In sharp contrast, absolute mobility at percentile  $i$  ( $E(y|x=i)$ ) cannot be point identified for any value of  $i$ . For further clarification, Appendix Figure A3 presents an annotated set of graphs describing how these relatively tight bounds are obtained.

Note that the wide uncertainty around changes in mobility as measured by  $\beta$  and  $p_{25}$  are not

---

<sup>23</sup>The rank-rank gradients are benchmarked against educational mobility estimates from Hertz et al. (2008). For  $p_{25}$  and  $\mu_0^{50}$ , we use income mobility estimates from Chetty et al. (2014a).

weaknesses of our partial identification approach, but strengths. 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. By delivering equivalently precise point estimates regardless of the coarseness of the data, conventional methods use hidden assumptions (such as linearity of the CEF, in the case of the rank-rank gradient) to convey excess precision.

Note finally that child ranks are also interval-censored in our context. However, interval censoring in the child distribution is unlikely to cause substantial bias for three reasons. First, because of increasing education over time, child rank bins are more evenly distributed than parent rank bins, which decreases potential bias from censoring. Second, we can estimate the extent of this bias by imputing within-bin ranks using additional data on children’s wages (which are not available for parents); doing so, we find virtually the same mobility estimates. This suggests that using the midpoint of a child’s rank bin is capturing most of the meaningful variation in child ranks. See Appendix C for details. Third, we can calculate all of our cross-group and cross-sectional estimates using an uncensored measure of child outcomes, such as an indicator function for high school completion; as we show below, we find similar trends and cross-group results when we do so. In the analysis below, we therefore assign the midpoint rank to each child in a given bin and treat the data as uncensored. As an alternative, we present a method that formally accounts for interval censoring in the child data in Appendix C, following a similar numerical optimization structure to that used here. This approach delivers identical results, but requires additional assumptions on the joint distribution of parent and child ranks.

The rank-rank gradient, absolute mobility, and bottom half mobility all capture slightly different characteristics of the intergenerational persistence of 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. To our knowledge, bottom half mobility is the first measure of intergenerational educational mobility that can be compared meaningfully across population subgroups, across countries, and across time. We therefore use this measure in our analysis below.

## 4.5 Comparison with Other Approaches

In this section, we briefly contrast our approach to measuring intergenerational educational mobility with several other recent approaches in the literature.

Card et al. (2018) use education data to compare geographic patterns in upward mobility between the 1920s and the 1980s. They define upward mobility in the 1920s as the 9th grade completion rate of children whose parents have 5–8 years of school, whom they describe as “roughly in the middle of the parental education distribution” — they then compare this measure with  $p_{25}$  for a birth cohort in the 1980s. Translating this into our framework, where  $x$  is a parent rank and  $y$  is a child outcome, Card et al. (2018) are comparing  $E(y_1|x=50)$  in the 1920s to  $E(y_2|x=25)$  in the present. This approach has the disadvantage of comparing upward mobility from the middle class in the 1920s to upward mobility from a considerably lower class in the present. The 1920s measure was chosen because it corresponds to bin boundaries in the education data. Our approach makes it possible to measure  $\mu_0^{50} = E(y|x \in [0,50])$  in both periods (or  $\mu_a^b$  for any other value of  $a$  or  $b$ ), regardless of the bin boundaries available in the data.

Alesina et al. (2019), who study intergenerational mobility across Africa face the same problem. They define upward mobility as the probability that a child born to a parent who has not completed primary school manages to do so. Both the  $x$  and the  $y$  variables represent different ranks in each country and time. In rank terms, this measure is approximately capturing  $E(y > 52|x \in [0,76])$  in Mozambique (where 76% of parents and 48% of children have not completed primary school) and  $E(y > 18|x \in [0,42])$  in South Africa. Using this measure implies comparing outcomes at very different points of the socioeconomic distribution in different countries and across time. These measures also do not distinguish between aggregate increases in education and the ability of individuals to rise to a new rank; these phenomena are independently interesting and our approach makes it possible to distinguish between them.

Finally, when constructing transition matrices from data that are not subdivided by exact quantiles, researchers often take the ad hoc approach of randomly reassigning individuals across bins to create the desired quantile bins. This approach is taken by the World Bank’s recent flagship report on

intergenerational mobility (Narayan and Van der Weide, 2018).<sup>24</sup> While this approach may seem innocuous, it in fact 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. For example, Narayan and Van der Weide (2018) find virtually identical outcomes for children growing up in the bottom three quartiles of the parent distribution in Ethiopia — this result is a mechanical artifact of over 80% of parents reporting the bottom-coded education level.

## 5 Results: Intergenerational Mobility in India

### 5.1 National Estimates

Figure 4 shows our main measure of mobility (bottom half mobility, or  $E(y|x \in [0,50])$ , where  $y$  is the child education rank). Panel A shows the father-son relationship. Upward mobility has been largely static over time, moving from  $[36.6, 39.0]$  in 1960–69 to  $[37.5, 37.9]$  for the 1980–85 birth cohort.<sup>25</sup> The bounds on the 1950–59 estimates are wider, leaving open the possibility of some gains from the 1950s to the 1960s birth cohorts.<sup>26</sup> The comparable measure for the 1980s birth cohort in the United States is 41.7, which is low by OECD standards.<sup>27</sup>

Panel B of the same figure describes mobility from fathers to daughters. This sample does not go back to the 1950s since there are no respondents from that birth cohort in the women’s surveys. We cannot reject a broadly similar pattern to the father-son results, though the bounds in the 1960s are wider than for sons, leaving open the possibility of mobility losses over this period. In fact, mobility for daughters has fallen by about half a percentage point from the 1970–79 birth cohort to the 1980–84 and 1985–89 birth cohorts. In the youngest birth cohort measured, father-daughter

---

<sup>24</sup>A similar approach is taken in Bound et al. (2015) and by Coile and Duggan (2019) to study mortality change in education quantiles.

<sup>25</sup>The measures are very tightly bounded for the 1980s birth cohort, because there is a rank boundary close to 50 in the parent distribution. When the distance between upper and lower bounds is less than 0.3, we report the midpoint as a point estimate.

<sup>26</sup>Appendix Figure A5 shows that these results are unlikely to be affected by survivorship bias. We estimate upward mobility for the same birth cohorts using the IHDS 2004–05; 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 show the same lack of change over time.

<sup>27</sup>Source: our calculations using  $\mu_0^{50}$ , based on data from Chetty et al. (2018). There is not yet a wide set of internationally comparable estimates of rank-based educational mobility, in part because of the methodological challenges described and addressed in this paper.

mobility is 35.6, about two percentage points lower than father-son mobility. Daughters are thus less likely to escape low socioeconomic status than sons.

Obtaining informative mobility estimates for the mother-child relationships is more difficult, because the distribution of women’s education is much more left-skewed (and thus censored in rank terms). Among mothers of the 1960s birth cohort, 82% had less than two years of education. For the 1985–89 birth cohort, this number was 65%. Under such severe censoring, we cannot estimate  $\mu_0^{50}$  with any precision. Even in the most recent 1985–89 birth cohort, we estimate bottom half mobility to be [37.5,41.4] for sons and [33.8,39.1] for daughters.<sup>28</sup> We thus focus on estimates of mobility based on fathers.

We can also calculate  $\mu_0^{50}$  with child education levels as the  $y$  variable. For example,  $E(\text{child years} \geq 12 | x \in (0,50))$  describes the likelihood that a child attains high school or greater, conditional on having a parent in the bottom half. Panels C and D show this measure for father-son and father-daughter links respectively. The graphs also show  $E(\text{child years} \geq 12 | x \in (50,100))$ , which is the likelihood of child high school attainment conditional on having a parent in the top half of the education distribution.<sup>29</sup> These graphs show the secular increase in high school attainment over time for children from privileged and underprivileged backgrounds. Girls from bottom half families have experienced the least gains, while girls born in the top half of the distribution have almost closed the gap with well-off boys. For both boys and girls, gains in high school attainment have accrued almost entirely to children from the top half of the distribution, a reflection of the stagnant overall upward mobility seen in panels A and B. However, these estimates confound mobility with aggregate increase in education, which is why we focus on  $\mu_0^{50}$  in rank terms.

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 which is unchanged. This result thus contradicts the narrative of India becoming a land of greater churn in terms of relative social status.

---

<sup>28</sup>Appendix Figure A4 shows the admittedly uninformative graph of this measure over time.

<sup>29</sup>This measure is similar in meaning to Chetty et al. (2014a)’s absolute downward mobility, or  $p_{75} = E(y|x=75)$ .

## 5.2 Changing Mobility in Across Social Groups

We next examine how these levels and trends differ across groups. Figure 5 presents results analogous to those above but subdivided into Muslims, Scheduled Castes, Scheduled Tribes, and all others. Panel A shows bottom half mobility ( $\mu_0^{50}$ ) from the 1950s to the 1980s for father-son pairs, revealing substantial differences across groups. As noted by other researchers, upward mobility for Scheduled Tribes, and especially for Scheduled Castes, has improved substantially (Hnatkovska et al., 2012; Emran and Shilpi, 2015). SCs born in the bottom half of the parent distribution in the 1960s could expect to obtain between the 33rd and 35th percentile; the comparable group in the 1980s obtains the 38th percentile, closing approximately half of the mobility gap with upper castes. Upward mobility for members of Scheduled Tribes rises from [29,31] to 33 over the same period.

In contrast with SCs and STs, Muslim intergenerational mobility declines substantially, falling from [31,34] in the 1960s to 29 in the 1985–89 birth cohort. These changes not only constitute a major decline in mobility, but make Muslim men the least upwardly mobile group in present-day India. Mobility for Muslim boys is lower even than for ST boys, who are often thought of as having benefited very little from Indian industrialization. The fact that a Muslim boy born to a family in the bottom half of the distribution can expect to obtain the 29th percentile implies almost no reversion to the national mean among this group. Finally, the “Forward/Others” group, which predominantly consists of higher caste Hindus, shows little change, with mobility shifting from [42,44] to 42. The static trend in upward mobility for boys can therefore be decomposed into gains for SCs and STs and losses for Muslims.

Panel B shows downward mobility ( $\mu_{50}^{100}$ ) for father-son links over the same period; this measure reflects the persistence of high status among 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 was found for upward mobility.<sup>30</sup>

---

<sup>30</sup>Appendix Figure A6 shows analogous results to Figure 5, but with education levels (at least primary, and at least high school) as outcomes, rather than education ranks. The results are consistent with the rank-based estimates, confirming that the separation between Scheduled Caste and Muslim boys is not driven by unobserved changes in latent ranks for children from these groups. The upward trend in the levels graphs reflects the overall rise in educational attainment seen in Figures 4C and 4D.

Panels C and D of Figure 5 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 were observed among boys. Table 2 summarizes the changes over time for the full sample and all the population subgroups, along with bootstrap confidence sets for partially identified data calculated following Chernozhukov et al. (2007). Table 3 shows confidence sets for static mobility differences between groups for the youngest (1985–89) birth cohort.<sup>31</sup>

To summarize, we observe a sharp divergence between upward mobility for boys from Scheduled Caste and Muslim groups. Muslim boys from poor families have declining mobility and very little opportunity to improve their relative social status. This low mobility may also adversely affect female Muslims, since marriage is nearly universal in India and almost entirely within subgroup, and female labor force participation is very low. Understanding how marriage ties interact with the upward mobility of boys and girls would be valuable, but is beyond the scope of the present paper.

### 5.3 Upward Mobility Across Geographic Areas

We next describe the geographic variation in upward mobility. The IHDS is of limited use here because it is representative only at the state level and does not have a large enough sample to accurately measure mobility at anything other than the largest geographic groups. Instead we use the SECC, whose large sample makes it possible to generate mobility estimates in very small geographic areas. However, in the SECC, we cannot measure mobility for Muslims (because religion is not recorded) or for father-daughter pairs (because our SECC sample is limited to coresident 20–23 year old boys, as described in Section 3). Finally, we do not explore time series patterns across geography, because we do not observe where children grew up, only where they are in 2011–12.

We can therefore disaggregate geography by subgroup and gender only at high levels of aggregation. Appendix Figure A7 shows bottom half mobility, disaggregated by child gender, social group, and

---

<sup>31</sup>The confidence sets 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. Nevertheless, the cross-group differences in the present are all highly significant, and we can state with 90% confidence that Muslim mobility among boys is declining.



rural urban status. Mobility is systematically higher in urban areas, but subgroup disadvantage varies substantially by geography. The urban-rural gap is much higher for girls than for boys, such that urban girls in fact have about five rank points higher upward mobility than urban boys. Muslims and SCs on average have higher mobility in cities, but their relative position with respect to Forwards/Others is worse.

The remainder of this section uses the SECC as a data source, reporting pooled social group estimates for father-son pairs. We first map the distribution of intergenerational mobility across India. Figure 6A presents a heat map of upward mobility across 4000 subdistricts and 2000 major towns across all of India. The graph shows the midpoint of the bounds; in 90% of cases, the bound width is less than two rank points.<sup>32</sup>

The geographic variation is substantial. Upward mobility is consistently highest in southern India—Tamil Nadu and Kerala—and is also noticeably high in the hilly states of the North. Parts of the Hindi-speaking belt—especially the state of Bihar—and the Northeast are among the lowest mobility parts of India. Gujarat is noteworthy as a state with high economic growth but relatively low mobility.

In broad regions of high mobility, there are low mobility islands, such as the hilly region between Andhra Pradesh and Karnataka. Cities and towns for the most part stand out as islands of higher mobility. However, there is not a single subdistrict or town in Bihar with higher average mobility than the southern states.

There is substantial variation in mobility based on neighborhood of residence even within a single city. Figure 6B shows a ward-level mobility map of Delhi.<sup>33</sup> The highest mobility wards have upward mobility that is 38% higher than the lowest mobility wards. Children in the dense and industrial areas of Northeast Delhi have the least opportunity; the average child from a bottom half family in this area can expect to obtain the 32nd percentile nationally. Children from similarly-ranked families in Southwest Delhi, about 5km away, can expect to obtain the 44th percentile.

---

<sup>32</sup>There are 8000 towns in the 2011 Population Census, on which the SECC is based. While our coverage of rural areas is almost complete, the data posted online described only 2000 of the towns. The town sample is broadly representative of the nation in demographics and income.

<sup>33</sup>See <http://www.dartmouth.edu/novosad/mobility-delhi.html> for an interactive version of this map with higher resolution.

To explore some of the potential drivers of geographic variation in upward mobility, Figure 7 presents the association between bottom half mobility and several correlates identified by the earlier literature on India and other countries. Panel A presents bivariate correlations between upward mobility and location characteristics across all rural subdistricts in India. Panel B presents analogous results across the town sample. The indicators cover four broad areas: subgroup distribution (specifically, presence of SCs and STs, and the residential segregation of SCs and STs); inequality (consumption and land inequality); development (manufacturing jobs per capita, average consumption, average education, and remoteness); and local public goods (schools, paved roads, and electricity). All of the measures are standardized to mean zero and standard deviation one for meaningful comparison.<sup>34</sup>

At the rural level, the traditional markers of economic development — manufacturing jobs, monthly consumption, and average levels of education — are the strongest correlates of upward mobility. Local public goods are also positively correlated with upward mobility. Interestingly, availability of primary schools is the least important of these, but availability of high schools is highly correlated with mobility. Surprisingly, the share of SCs and STs is positively correlated with upward mobility. Segregation and land inequality are negatively correlated with upward mobility, a parallel result to that found in the United States (Chetty et al., 2014a).

Turning to urban places, we find that upward mobility is higher in towns that have (i) higher population; (ii) more SCs; (iii) more educated populations; and (iv) more high schools per capita. As in rural areas, SC/ST segregation is negatively associated with mobility. By contrast, consumption inequality is positively associated with mobility.

These local mobility estimates have two limitations. First, they are based on the educational outcomes of children born between 1989 and 1992, the majority of whom finished their education by 2010. They therefore reflect the circumstances that drove education choices in the period 2000–2010, which may be different in the present. Second, the estimates do not account for migration, as we do not observe respondents’ location of birth. Low mobility in Northeast Delhi could in part be the effect of immigration from poor parts of rural North India. The more local the mobility estimate, the

---

<sup>34</sup>All of these variables are defined in Appendix D.

greater is the potential bias from migration. Ideally, we would have local surveys that record both location of origin and parental education; to our knowledge, there are no such surveys with high geographic precision. The rural (subdistrict-level) estimates are less likely to be biased by migration, because permanent migration in rural areas in India is extremely low.<sup>35</sup>

The picture that emerges is one where place of residence appears to be highly important to upward mobility. Cities have higher mobility for all groups, but the effect of geography may be different for individuals of different genders and population subgroups. We leave additional exploration of these patterns for future research.

## 5.4 Potential Mechanisms for Subgroup Mobility Differences

In this section, we explore a series of mechanisms that could potentially explain the upward mobility differences and changes across social groups in India, with a focus on the growing mobility gap between Scheduled Castes and Muslims. We examine the extent to which these group differences can be explained by: (i) affirmative action for Scheduled Castes; (ii) differential fertility across social groups; (iii) patterns in location of residence; and (iv) different occupational patterns causing differential returns to education. These analyses are suggestive, but they point toward affirmative action as a key mechanism for the SC/Muslim divergence, and largely reject the other three mechanisms.

### 5.4.1 Mechanisms: Affirmative Action for Scheduled Groups

First, we consider the hypothesis that the basket of programs and policies targeted to Scheduled Castes and Scheduled Tribes has driven the increase in upward mobility of SCs and STs relative to Muslims since the 1950s. To estimate the causal effect of affirmative action on bottom half mobility, we exploit a change in the population groups eligible for Scheduled Caste status that was made in 1977, previously studied by Cassan (2019). Between 1948 and 1977, there were discrepancies across state lists regarding the caste groups eligible for the Scheduled Caste (SC) designation. In 1977, these lists were harmonized, arbitrarily moving many additional groups into the SC designation and thus making them eligible for SC-targeted benefits. As shown by Cassan (2019), this policy

---

<sup>35</sup>Foster and Rosenzweig (2007) find decadal rural-to-urban migration rates for 15–24 year old males of about 3% in 1961–2001; Munshi and Rosenzweig (2016) present confirmatory evidence in IHDS and the DHS.

change makes it possible to examine the impact of Scheduled Caste status, while controlling for a group’s ethnicity, historical experience, and geographic region.<sup>36</sup>

To test for the impact of affirmative action on upward mobility, we divide SCs into two groups: (i) those who were classified as SCs at the time when states were reorganized on linguistic lines in 1956 (whom we call early SCs); and (ii) a group which obtained protected status only in 1977 (whom we call late SCs). Following Cassan (2019), we assume that individuals needed to be 6 or younger in 1977 to benefit in terms of education from the change in status.<sup>37</sup> We therefore treat individuals born later than 1966 as being in the late SC group.

We assign individuals to early and late SC groups using jati-level group identifiers in the IHDS.<sup>38</sup> Figure 8 demonstrates the upward mobility trajectory of the early and late SC groups, showing bounds on  $\mu_0^{50}$  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. Beginning with the 1970 birth cohorts when the late SC group obtains protected status, it begins to close the mobility gap. The mobility gap peaks right before the late SC group gains SC status and steadily closes through the sample period.

We can formally estimate the impact of SC status in a regression based on the specification of Cassan (2019), that takes the format:

$$Y_{i,j,r,c} = \beta_0 + \beta_1 \text{LateSC}_{j,r} + \beta_2 \text{post}_c + \beta_3 (\text{post}_c \times \text{LateSC}_{j,r}) + \nu_{j,r} + \eta_{r,c} + \zeta_{j,c} + \epsilon_{i,j,r,c}, \quad (5.1)$$

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 ( $\nu$ ), region  $\times$  cohort ( $\eta$ ), and jati  $\times$  cohort ( $\zeta$ ).  $\beta_3$  therefore compares individuals in late SC groups born after 1970 to those in the same narrow social group and the same region, controlling for outcomes of individuals from early SC groups in the same region, and for outcomes of

<sup>36</sup>Looking primarily at educational outcomes, Cassan (2019) finds that newly scheduled groups experience literacy and schooling gains, but these accrue mostly to boys.

<sup>37</sup>We find similar results if we use a cutoff of 11 or younger, also used in Cassan (2019).

<sup>38</sup>A jati is a caste identifier that is more granular than the broad Scheduled Caste category, which includes many jatis.

members of the same jati group in other regions. Regressions are clustered at both the jati and the region level. Because of interval-censored parent education data, we approximate upward mobility by limiting the sample to children of parents in the bin that is as close to the bottom 50% as possible.

The result of this regression is shown in Table 4. Column 1 shows the specification above. The cohorts exposed to the basket of affirmative action policies experience a six rank point change in upward mobility. Column 2 shows robustness to a specification where we limit the sample to sons of fathers with less than two years of education instead of the set of fathers in the bottom 50%. Column 3 shows that both the 1970s and 1980s birth cohorts of late SCs benefit relative to the earlier cohorts.

The results suggest that affirmative action has had a large effect on upward mobility for Scheduled Caste groups. The six rank point gain of late SCs is more than 50% of the gap in the 1980s between Muslims and SCs, and it emerged over only 20 years of affirmative action. If these treatment effects are externally valid for disadvantaged Muslims, then affirmative action is potentially a key driver of the growing gap between Scheduled Castes and Muslims.

#### 5.4.2 Group Differences and Fertility

Muslims on average have higher fertility than either Scheduled Castes, Scheduled Tribes, or Forward Castes and other groups. In this section, we consider whether higher fertility could cause lower mobility for Muslims, perhaps through a household expenditure channel where children with many siblings received fewer educational inputs. We explore this question in a regression framework. We estimate the number of siblings of each individual based on their mothers' responses to the IHDS women's survey, which has a question about the number of 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 boys under the age of 30, for whom the coresidence rate is highest.<sup>39</sup> The average number of siblings for Muslims is 5.5, compared with 4.3 for SCs and STs, and 4.2 for Forwards/Others.

To obtain a point estimate of upward mobility to use in a regression, we limit the sample to children

---

<sup>39</sup>For girls, coresidence begins to fall rapidly as soon as schooling is finished, leaving too little sample to estimate mobility among coresiders. Restricting the sample to individuals aged 20–23 as we did for the SECC would cut our sample too much to obtain informative estimates.

born in the 1980s whose father completed two or fewer years of education. The average child rank in this subsample corresponds to  $\mu_0^{51}$ , which can be point estimated.<sup>40</sup> We regress this mobility measure on a set of group indicators (Muslim, SC, ST), an urban indicator, and a set of state fixed effects, showing the results in Table 5. Column 1 shows the Muslim mobility gap in the full sample and Column 2 shows the same gap in the set of boys whose coresident mother answered the women’s survey. An additional sibling is associated with 4 fewer ranks on the upward mobility scale. The Muslim upward mobility disadvantage is 13 rank points in the full sample and 12 in the restricted sample. Column 3 adds a control for the number of siblings, which brings the Muslim mobility gap down by 25%.

High fertility can thus explain at most 25% of the Muslim mobility disadvantage relative to SCs. This is likely to be an upper bound, because there is household income is a direct cause of both children’s education and parental fertility (Schultz, 2003). Higher fertility can thus explain only a small share of the present-day mobility disadvantage experienced by Muslims.

### 5.4.3 Geography and Subgroup Differences

To describe the extent to which Muslim disadvantage in upward mobility can be explained by geography, we examine here whether Muslims live in low mobility places, or whether they have low mobility after conditioning on place. Because the SECC does not record religion, we use the IHDS and explore cross-state and cross-district variation.

SCs, STs and Muslims are unevenly distributed across the country; the 25th-percentile district in SC population is only 8% SC. The equivalent numbers for STs and Muslims (0.4% and 2.7%), reflecting the great geographic concentration of these groups.

To examine the relationship between place and subgroup outcomes, we regenerate father and son education ranks within states and within districts. Mobility estimates generated in this way thus describe the ability of disadvantaged children to increase their relative rank within their own district. If low overall mobility for Muslims is a function of living in districts where everyone has low opportunity, then their within-district mobility gap with Forwards / Others should be substantially

---

<sup>40</sup>We do this to eliminate the challenge of estimating a regression where the outcome variable is only known to lie within a bound.

smaller than the national mobility gap. We focus on the static father-son mobility gap for the most recent birth cohort, which are individuals born in the 1980s.

These results are shown in Figure 9. The first set of bars shows the relative mobility gap to Forwards / Others for the three marginalized groups using national education ranks; these gaps correspond to the differences between groups in Figure 5A in the 1980s. For simplicity, we show the midpoint of the bounds; the width of the bounds is less than one percentage point in all cases. Upward mobility for the Forward / Others reference group is 41. The following two sets of bars show the same gaps for within-state and within-district ranks.

The extent to which group disadvantages in upward mobility can be explained by location differ substantially by group. District of residence explains about 9% of the Muslim upward mobility gap, 14% of the Scheduled Caste upward mobility gap, and a full 59% of the Scheduled Tribe mobility gap.<sup>41</sup> The result for Scheduled Tribes is consistent with the fact that STs disproportionately live in remote areas of the country with low levels of public goods and educational attainment. Given the uneven distribution of SCs and Muslims throughout India, the unimportance of district as an explanation for their lower mobility is worthy of note. Muslim disadvantage cannot be explained by the broad regions in which Muslims live. However, these results do not rule out the possibility that finer geographic definitions (such as urban neighborhoods) could explain a greater share of the mobility gap; unfortunately, higher resolution analysis is not possible with the data available at this time. To be clear, these results show that location is not a major mediator of SC and Muslim disadvantage; the prior section (Section 5.3) shows that location is an important predictor of mobility in the aggregate.

#### 5.4.4 Occupations, Returns to Education, and Subgroup Differences

We next examine whether occupational choices and returns to education can explain the low and falling upward mobility of Muslims. Muslims are more likely to work as small-scale entrepreneurs than the other major social groups (Figure 11A). If the returns to education are lower in entrepreneurship than in wage work, the low education outcomes of Muslim children born into poor families would

---

<sup>41</sup>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 Figure 9.

reflect a different expected career path but not necessarily a lack of opportunity.

To evaluate this hypothesis, we first examine the Mincerian returns to education for the different social groups. Figure 10 shows point estimates from Mincerian return regressions for each social group, calculated in three ways: (i) household log income on household head education (IHDS); (ii) individual log wages on individual education (NSS); and (iii) household consumption on household head education.<sup>42</sup> Across all three measures, there is no evidence that Muslims have lower returns to education than Scheduled Castes or Tribes. The point estimates for Muslims are higher than for SCs in all cases, though both Muslims and SCs have lower returns to education than Forward/Other castes. Mincerian returns are not certain to reflect the causal effect of education on income and consumption, but there is no evidence here that Muslims are choosing less education because their returns are lower.

Even if returns to education for Muslims are not lower overall, they could be low for Muslims who choose to run small businesses. But the data rejects this hypothesis as well. Figure 11B divides the IHDS sample into individuals who own their own business (right panel) and individuals who do not (left panel). We pool SCs and STs for this graph because very few SCs and STs own businesses, leading to small samples. The divergence in upward mobility between SC/STs and Muslims is sustained and of similar magnitude both among business- and non-business owning families.

In conclusion, the evidence does not support the idea that the decline in Muslim upward mobility is an artifact of Muslim occupational choice or returns to education.

## 6 Conclusion

In this paper, we present a set of tools that are well-suited to measuring intergenerational educational mobility in developing countries and other contexts where high quality income data are unavailable and education is coarsely binned. Our partial identification approach takes seriously the loss of information when a very large share of the population reports a bottom-coded education level. We have shown that bottom half mobility is a measure that is easy to calculate, analogous to the popular

---

<sup>42</sup>In each case, we regress the outcome variable on individual years of education, age, and age squared. We restrict the data to men aged 18–64 to avoid concerns about selection into labor markets. Results are similar if we restrict to young ages (reflecting education of the youngest birth cohorts), if we include women and men, and if we include additional controls.



absolute upward mobility measure, and informative regarding intergenerational mobility even when education data are very coarse. Bottom half mobility is also the first measure of intergenerational educational mobility that is meaningful for cross-group analysis across contexts; the absence of such a measure has prevented researchers from studying subgroup mobility in developing countries.

Despite enormous economic and political changes, bottom half mobility in India has barely changed from the 1950s to the 1980s birth cohorts. This lack of change overall can be decomposed into substantial gains for SC/STs and substantial losses for Muslims, a growing disparity which can be explained at least in part by the basket of affirmative action policies that have targeted SC/ST disadvantage since Indian independence. The falling mobility of Muslims has not previously been noted in part because there has previously been no methodology for creating comparable rank bins across cohorts.

Our work has only begun to describe the wide geographic and cross-group variation in intergenerational mobility in India. As in the U.S., location is a very strong predictor of intergenerational mobility, even if cross-group differences appear to be largely invariant to location for Muslims and SCs. Individuals growing up in different parts of India, even conditional on similar economic conditions in the household, can expect vastly different opportunities and outcomes throughout their lives. Future work describing the geographic variation in mobility in more detail, and moving toward causal estimates of the impact of place, will be an important basis for policies that create opportunities for those who are currently being left behind in India and in other developing countries around the world.

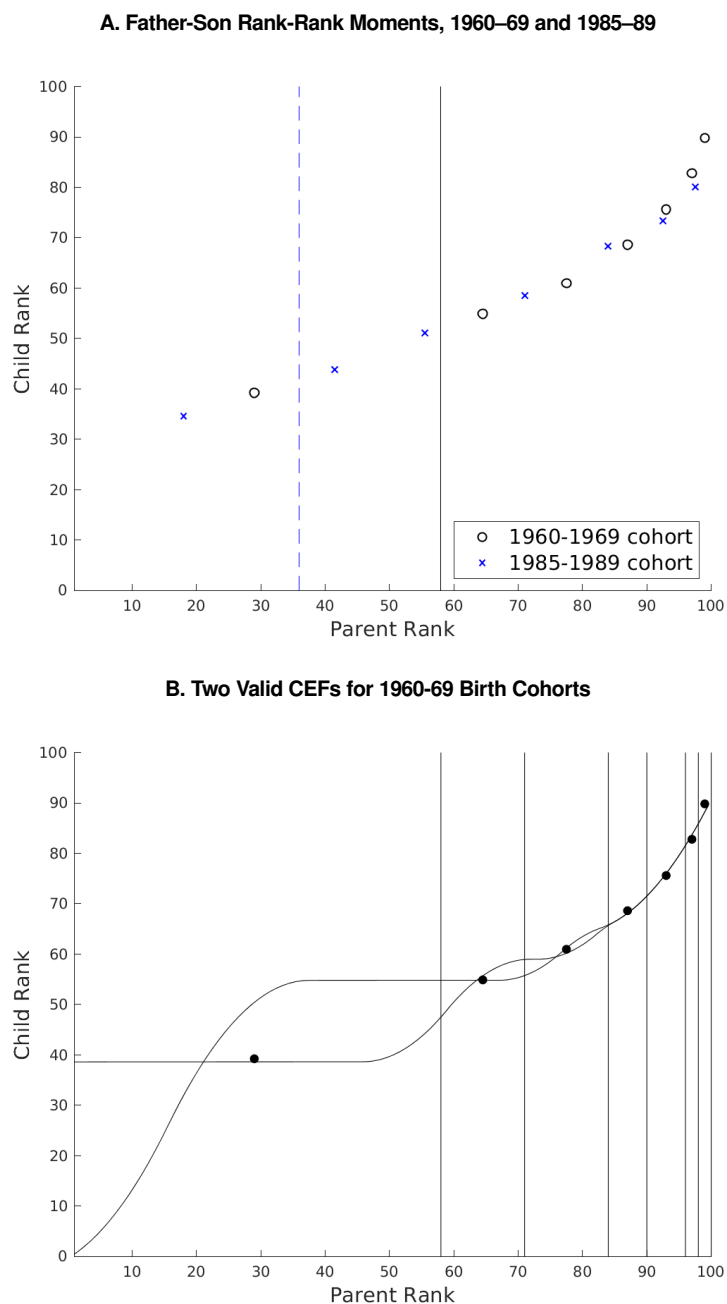
## References

- Alesina, Alberto, Sebastian Hohmann, Elias Papaioannou, and Stelios Michalopoulos, "Intergenerational Mobility in Africa," 2019. NBER Working Paper No.25534.
- Asher, Sam and Paul Novosad, "Rural Roads and Local Economic Development," *American Economic Review* (forthcoming), 2019.
- Azam, Mehtabul and Vipul Bhatt, "Like Father, Like Son? Intergenerational Educational Mobility in India," *Demography*, 2015, 52 (6).
- Bagde, Surendrakumar, Dennis Epple, and Lowell Taylor, "Does affirmative action work? Caste, gender, college quality, and academic success in India," *American Economic Review*, 2016, 106 (6).
- Basant, Rakesh, Abusaleh Shariff et al., *Handbook of Muslims in India: Empirical and policy perspectives*, Oxford University Press, 2010.
- Berenschot, Ward, "Riot politics: Hindu-Muslim violence and the Indian state," 2012.
- Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan, "Affirmative action in education: Evidence from engineering college admissions in India," *Journal of Public Economics*, 2010, 94 (1-2), 16–29.
- Bhalotra, Sonia and Bernarda Zamora, "Social divisions in education in India," in Rakesh Basant and Abusaleh Shariff, eds., *Handbook of Muslims in India*, New Delhi: Oxford University Press, 2010.
- Black, S, P Devereux, and Kjell G Salvanes, "Why the Apple Doesn't Fall: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, 2005, 95 (1).
- Black, Sandra E. and Paul J. Devereux, "Recent Developments in Intergenerational Mobility," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Amsterdam: North Holland Press, 2011.
- Blakeslee, David S, "The Rath Yatra Effect: Hindu Nationalist Propaganda and the Rise of the BJP," 2018. Working Paper.
- Boserup, Simon Halphen, Wojciech Kopczuk, and Claus Thustrup Kreiner, "Stability and persistence of intergenerational wealth formation: Evidence from Danish wealth records of three generations," 2014.
- Bound, John, Arline T. Geronimus, Javier M. Rodriguez, and Timothy A. Waidmann, "Measuring Recent Apparent Declines In Longevity: The Role of Increasing Educational Attainment," *Health Affairs*, 2015, 34 (12), 2167–2173.
- Bratberg, Espen, Jonathan Davis, Bhashkar Mazumder, Martin Nybom, Daniel Schnitzlein, and Kjell Vaage, "A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden and the U.S.," *The Scandinavian Journal of Economics*, 2015, 119 (1).
- Bratsberg, Bernt, Knut Røed, Oddbjørn Raaum, Robin Naylor, Markus Jäntti, Tor Eriksson, and Eva Österbacka, "Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons," *Economic Journal*, 2007, 117 (519).
- Card, David, Ciprian Domnisoru, and Lowell Taylor, "The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility," 2018. NBER Working Paper No.25000.
- Cassan, Guilhem, "Affirmative Action, Education and Gender: Evidence from India," *Journal of Development Economics*, 2019, 136.
- Chernozhukov, Victor, Han Hong, and Elie Tamer, "Estimation and confidence regions for parameter sets in econometric models," *Econometrica*, 2007, 75 (5).
- Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan, "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility," 2017. NBER Working Paper No. 23618.

- , **Nathaniel Hendren**, **Maggie R Jones**, and **Sonya R Porter**, “Race and Economic Opportunity in the United States: An Intergenerational Perspective,” 2018.
- , — , **Patrick Kline**, and **Emmanuel Saez**, “Where is the land of opportunity? The geography of intergenerational mobility in the United States,” *Quarterly Journal of Economics*, 2014, 129 (4), 1553–1623.
- , — , — , — , and **Nicholas Turner**, “Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility,” *American Economic Review: Papers & Proceedings*, 2014, 104 (5), 141–147.
- Coile, Courtney and Mark Duggan**, “When Labor’s Lost: Health, Family Life, Incarceration, and Education in a Time of Declining Economic Opportunity for Low-Skilled Men,” 2019. NBER Working Paper No. 25569.
- Corak, Miles**, “Income Inequality, Equality of Opportunity, and Intergenerational Mobility,” *Journal of Economic Perspectives*, 2013, 27 (3).
- Dardanoni, Valentino, Mario Fiorini, and Antonio Forcina**, “Stochastic Monotonicity in Intergenerational Mobility Tables,” *Journal of Applied Econometrics*, 2012, 27.
- Derenoncourt, Ellora**, “Can you move to opportunity? Evidence from the Great Migration,” 2018.
- Dunn, Christopher E.**, “The intergenerational transmission of lifetime earnings: Evidence from Brazil,” *B.E. Journal of Economic Analysis and Policy*, 2007, 7 (2).
- Emran, M.S. and F. Shilpi**, “Intergenerational Occupational Mobility in Rural Economy: Evidence from Nepal and Vietnam,” *Journal of Human Resources*, 2011, 46 (2).
- and **Forhad Shilpi**, “Gender, Geography, and Generations: Intergenerational Educational Mobility in Post-Reform India,” *World Development*, 2015, 72.
- Foster, Andrew D. and Mark R. Rosenzweig**, “Economic Development and the Decline of Agricultural Employment,” in “Handbook of Development Economics,” Vol. 4 2007, chapter 47, pp. 3051–3083.
- Frisancho Robles, Veronica C. and Kala Krishna**, “Affirmative Action in Higher Education in India: Targeting, Catch Up and Mismatch,” *Higher Education*, 2016, 71 (5).
- Güell, Maia, José V Rodríguez Mora, and Christopher I. Telmer**, “The informational content of surnames, the evolution of intergenerational mobility, and assortative mating,” *Review of Economic Studies*, 2013, 82 (2).
- Guest, Avery M, Nancy S Landale, and James C Mccann**, “Intergenerational Occupational Mobility in the Late 19th Century United States,” *Social Forces*, 1989, 68 (2).
- Hertz, Tom**, “Rags, riches and race: The intergenerational economic mobility of black and white families in the United States,” in Samuel Bowles, Herbert Gintis, and Melissa Osborne Groves, eds., *Unequal Chances: Family Background and Economic Success*, Princeton University Press, 2005.
- , **Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina**, “The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends,” *The B.E. Journal of Economic Analysis & Policy*, 2008, 7 (2).
- Hilger, Nathaniel G.**, “The Great Escape: Intergenerational Mobility Since 1940,” 2016. NBER Working Paper No. 21217.
- Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh B. Paul**, “Castes and Labor Mobility,” *American Economic Journal: Applied Economics*, 2012, 4 (2).
- , — , and — , “Breaking the caste barrier: intergenerational mobility in India,” *The Journal of Human Resources*, 2013, 48 (2).
- Hungerford, Thomas and Gary Solon**, “Sheepskin Effects in the Returns to Education,” *The Review of Economics and Statistics*, 1987, 69 (1).
- Ito, Takahiro**, “Caste discrimination and transaction costs in the labor market: Evidence from rural

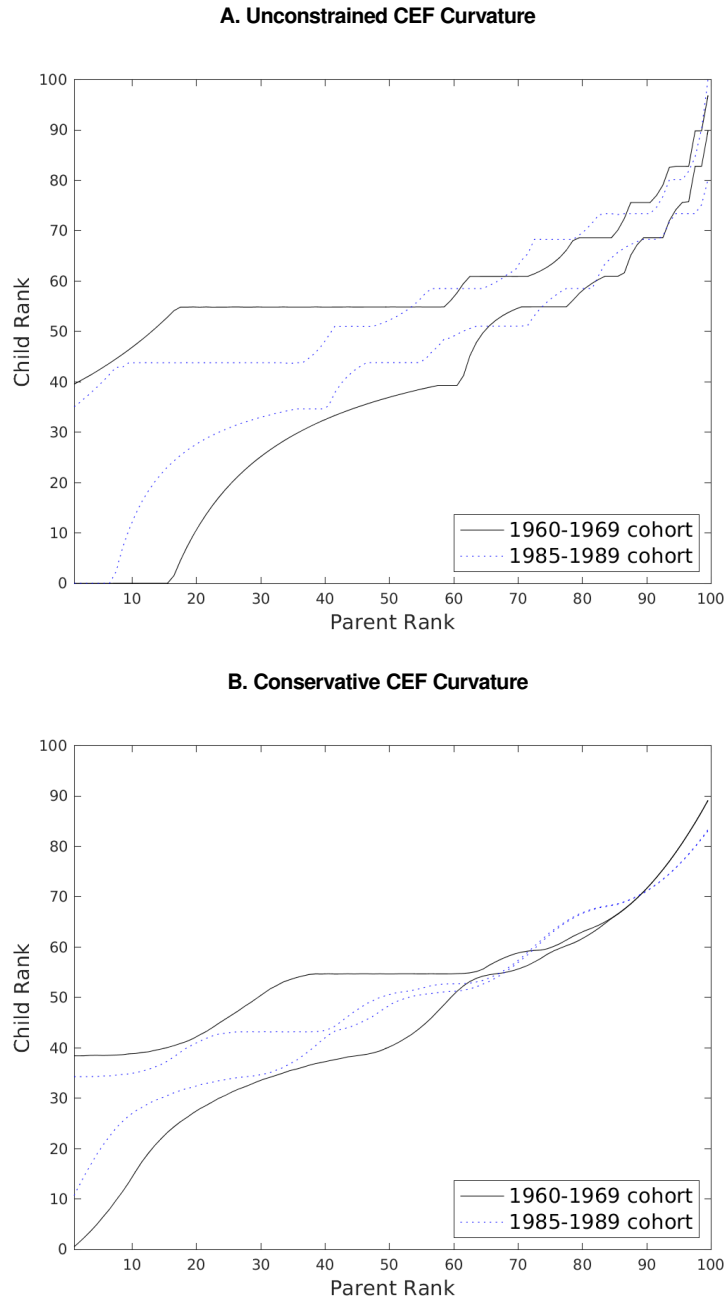
- North India,” *Journal of Development Economics*, 2009, 88 (2), 292–300.
- Khamis, Melanie, Nishith Prakash, and Zahra Siddique**, “Consumption and social identity: Evidence from India,” *Journal of Economic Behavior & Organization*, 2012, 83 (3), 353–371.
- Knight, John, Terry Sicular, and Ximing Yue**, “Educational Inequality in China: The Intergenerational Dimension,” 2011. Center for Human Capital and Productivity Working Paper.
- Lindahl, Mikael, Mårten Palme, Sofia Sandgren Massih, and Anna Sjögren**, “Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations,” *Journal of Human Resources*, 2012, 50 (1).
- Long, Jason and Joseph Ferrie**, “Intergenerational Occupational Mobility in Great Britain and the United States Since 1850,” *American Economic Review*, 2013, 103 (4).
- Maitra, Pushkar and Anurag Sharma**, “Parents and Children: Education Across Generations in India,” 2009. Working Paper.
- Manski, Charles F. and Elie Tamer**, “Inference on Regressions with Interval Data on a Regressor or Outcome,” *Econometrica*, 2002, 70 (2), 519–546.
- Mohammed, A R Shariq**, “Does A Good Father Now Have To Be Rich? Intergenerational Income Mobility in Rural India,” 2016. Working paper.
- Munshi, Kaivan and M Rosenzweig**, “Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap,” *American Economic Review*, 2016, 106 (1), 46–98.
- and **Mark Rosenzweig**, “Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy,” *American Economic Review*, 2006, 96 (4), 1225–1252.
- Narayan, Ambar and Roy Van der Weide**, *Fair Progress? Economic Mobility Across Generations Around the World*, World Bank Group, 2018.
- Novosad, Paul, Charlie Rafkin, and Sam Asher**, “Mortality Change Among Less Educated Americans,” 2020. Working Paper.
- Piraino, Patrizio**, “Intergenerational Earnings Mobility and Equality of Opportunity in South Africa,” *World Development*, 2015, 67 (March), 396–405.
- Ponomareva, Maria and Elie Tamer**, “Econometrics Misspecification in moment inequality models: back to moment equalities?,” *The Econometrics Journal*, 2011, 14.
- Roemer, John.**, “Equality of Opportunity: Theory and Measurement,” *Journal of Economic Literature*, 2016, 54 (4).
- Sachar Committee Report**, “Social, Economic and Educational Status of the Muslim Community of India,” Technical Report, Government of India 2006.
- Schultz, T Paul**, “School subsidies for the poor: evaluating the Mexican Progresa poverty program,” *Journal of Development Economics*, 2003, 74, 199–250.
- Solon, Gary**, “Intergenerational Mobility in the Labor Market,” in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Amsterdam: North Holland Press, 1999, pp. 1761–1800.
- Wantchekon, Leonard, Marko Klačnja, and Natalija Novta**, “Education and Human Capital Externalities: Evidence from Benin,” *The Quarterly Journal of Economics*, 2015, 130 (2), 703–757.
- Wilkinson, Steven I**, *Votes and violence: Electoral competition and ethnic riots in India*, Cambridge University Press, 2006.
- Zimmerman, David J**, “Regression toward mediocrity in economic stature,” *The American Economic Review*, 1992.

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



Panel A of the figure shows the average child education rank in each parent education rank bin for the 1960–69 and 1985–89 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 1960–69 birth cohort and the dashed line to the 1985–89 birth cohort. Points are displayed at the midpoint of each parent rank bin. Panel B shows the 1960–69 moments again, along with two simulated conditional expectation functions which are equally good fits to the moments. Source: IHDS 2012.

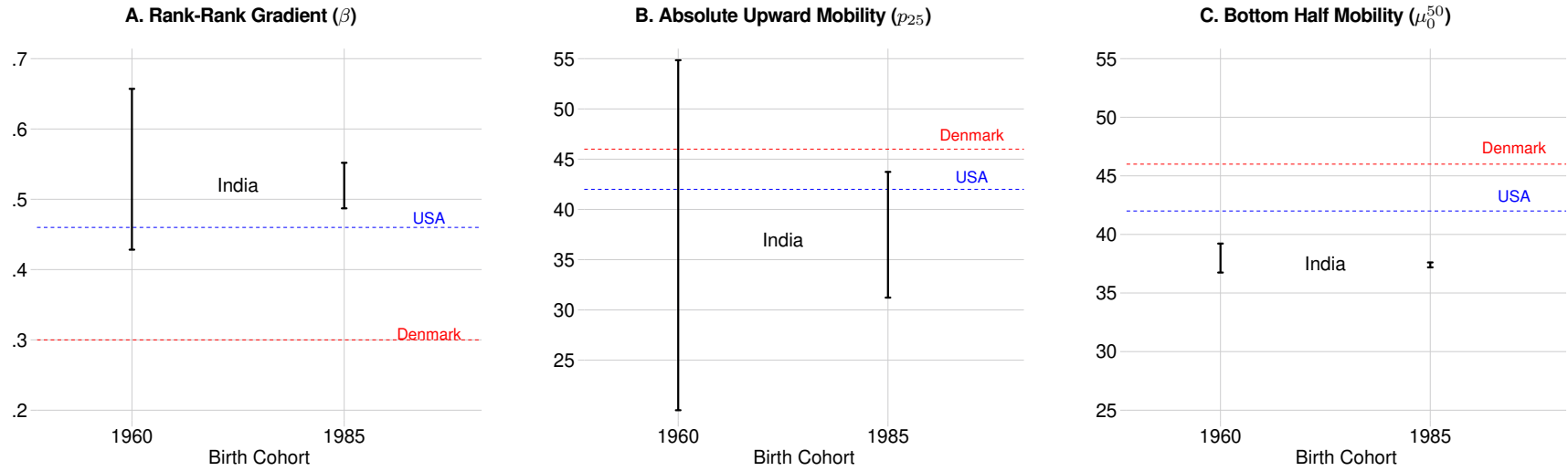
**Figure 2**  
Changes in Bounded Father-Son Rank CEF  
from 1960–69 to 1985–89 Birth Cohorts



The figure presents the change over time in the rank-rank relationship between Indian fathers and sons born in the 1960–69 and the 1985–89 birth cohorts. The graph shows bounds on the expected child education rank, given a parent at a given rank in the parent education distribution. Panel A shows bounds on father-son CEFs with unconstrained curvature. Panel B shows bounds using a conservative curvature constraint equal to double the maximum curvature of parent-child *income* rank-rank CEFs from Sweden, Denmark, Norway, and the United States (Chetty et al., 2014b; Boserup et al., 2014; Bratberg et al., 2015). Source: IHDS 2012.

**Figure 3**

Bounds on Mobility Measures in India:  
1960–69 and 1985–89 Birth Cohorts



The figure shows bounds on three mobility statistics for the 1960–69 and 1985–89 birth cohorts, estimated on father-son pairs in India. For reference, we display estimates of similar statistics from USA and Denmark. Data on rank-rank education gradients are from Hertz et al. (2008). For  $p_{25}$  and  $\mu_0^{50}$ , the USA 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 25th percentile.  $\mu_0^{50}$  is bottom half mobility, which is the expected rank of a son born below to a father below the 50th percentile. Source: IHDS 2012.

**Figure 4**  
Bottom Half Mobility, Fathers to Sons and Daughters

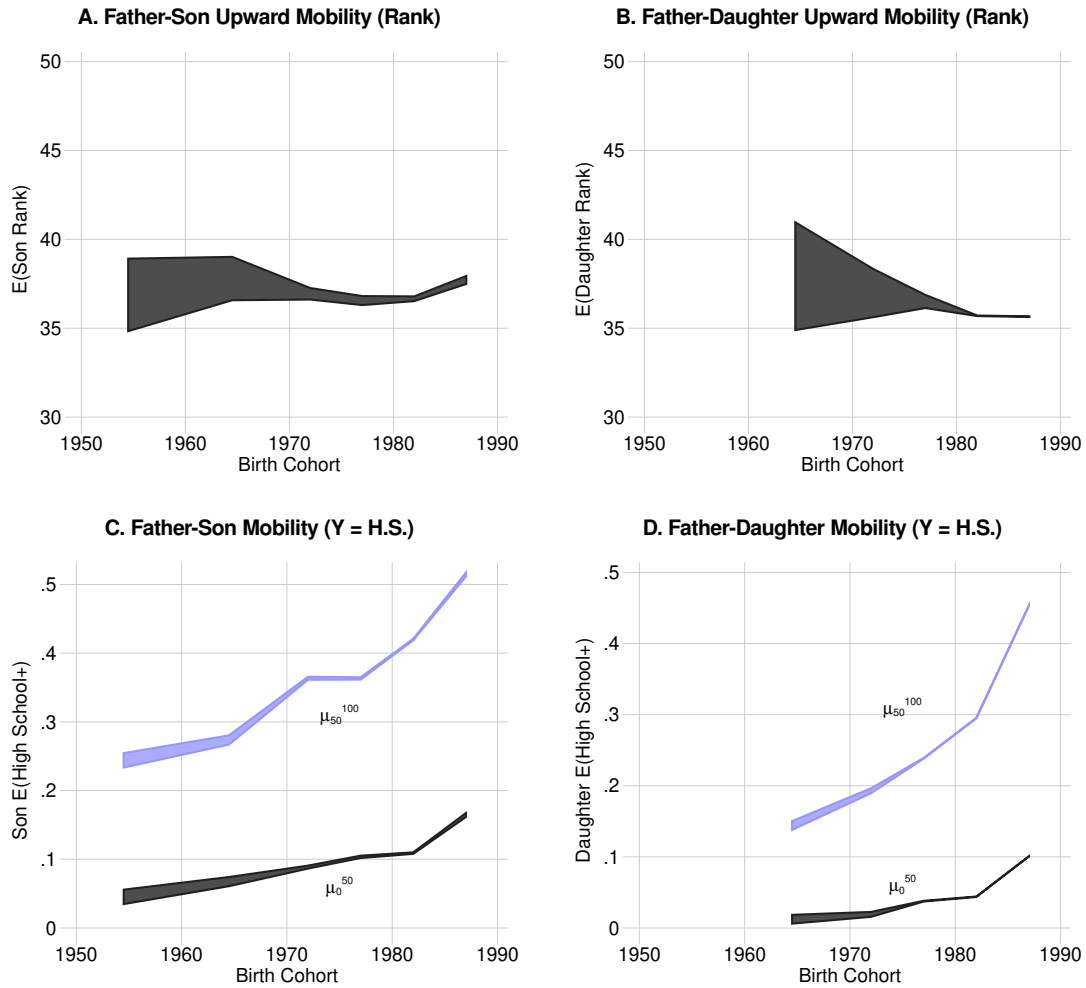


Figure 4 presents bounds on aggregate trends in intergenerational mobility, using cohorts born from 1950 through 1989. Panels A and B show bottom half mobility ( $\mu_0^{50} = E(y|x \in [0,50])$ ), where  $x$  is parent rank and  $y$  is child rank. This is the average rank attained by children born 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  $E(HS|x \in [50,100])$  (blue). 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 (blue) is the share of children completing high school, conditional on parents in the top half of the parent distribution. Source: IHDS 2012.



**Figure 5**  
Trends in Mobility by Subgroup, 1950–1989 Birth Cohorts

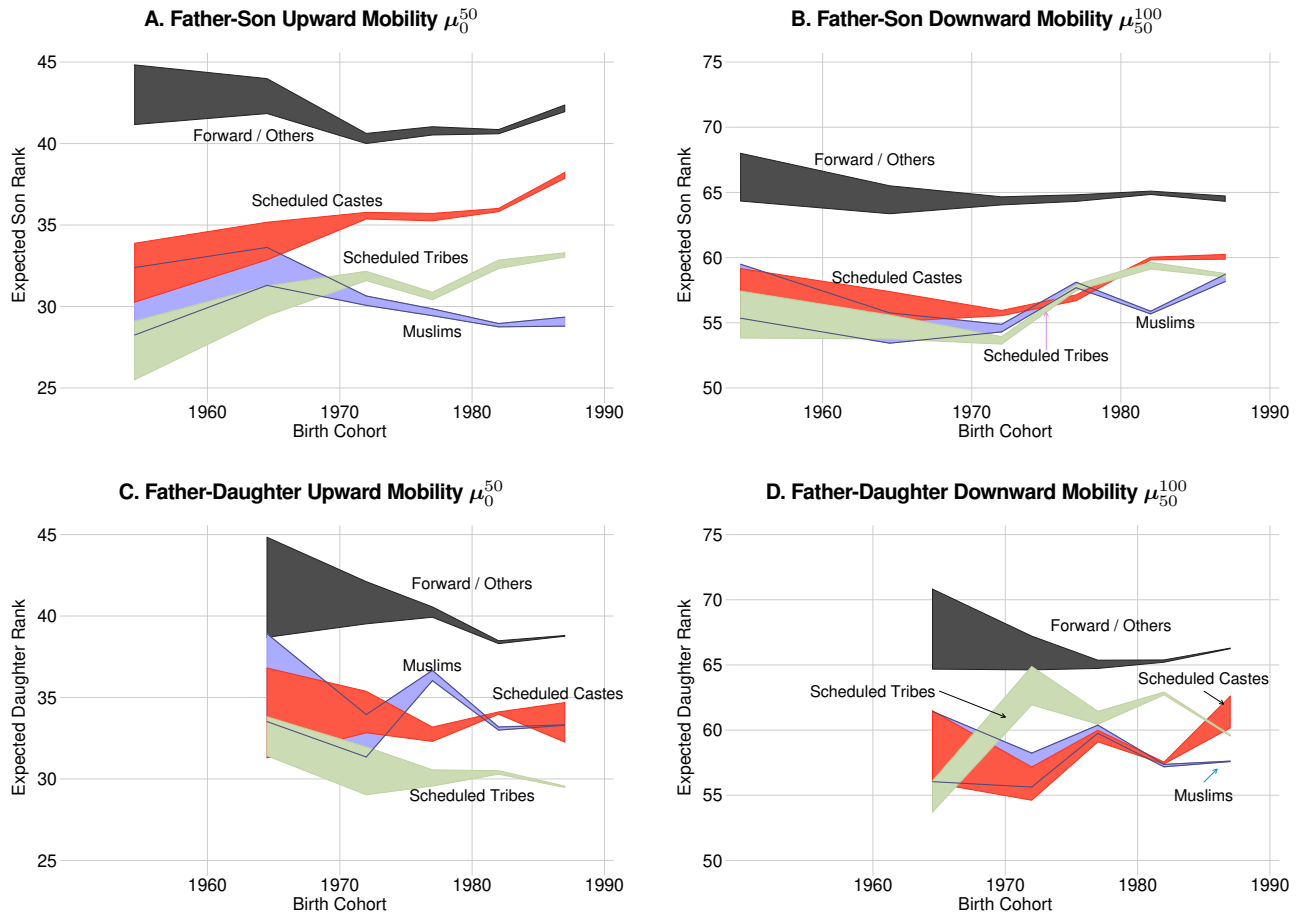


Figure 5 presents bounds on trends in intergenerational mobility, stratified by four prominent social groups in India: Scheduled Castes, Scheduled Tribes, Muslims, and Forward Castes/Others. The mobility measure in Panels A and B is bottom half mobility ( $\mu_0^{50}$ ), or the average rank among children born to fathers in the bottom half of the father education distribution. The measure in Panels C and D is top half mobility ( $\mu_{50}^{100}$ ), 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 1950–59 birth cohort. Source: IHDS 2012.

**Figure 6**  
Upward Mobility by Geographic Location: National and Neighborhood Estimates

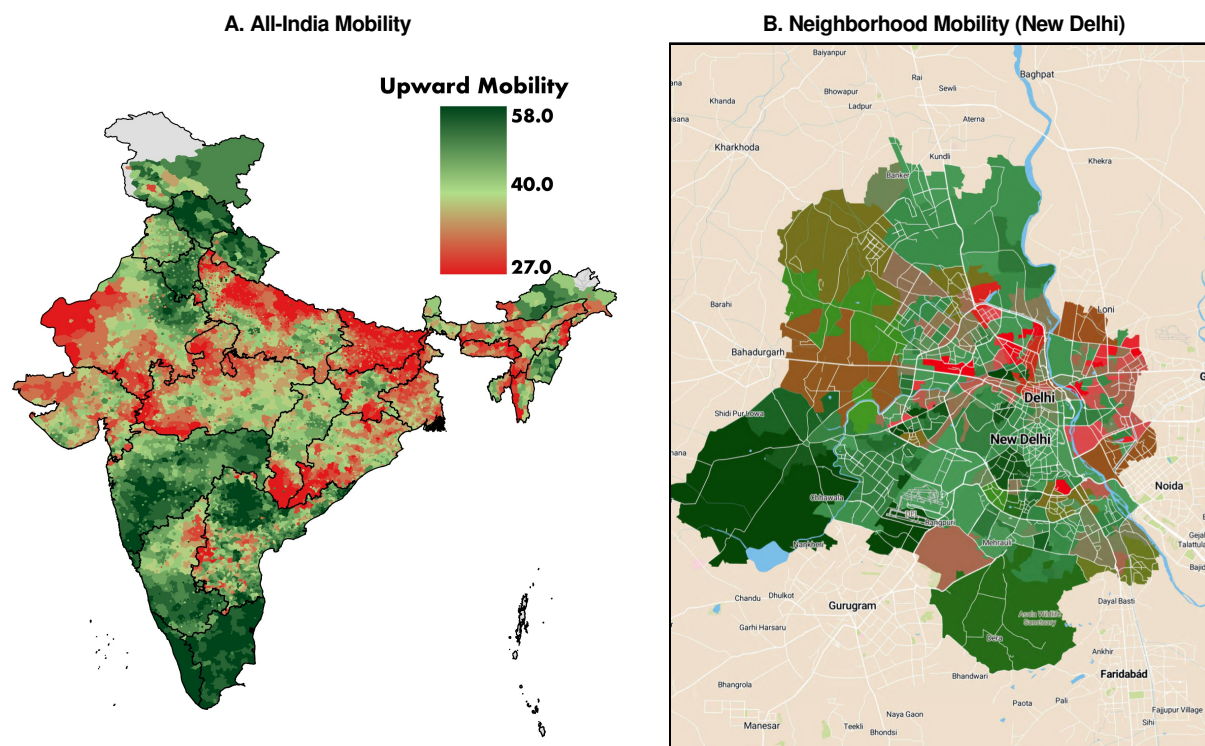


Figure 6 Panel A presents a map of the geographic distribution of upward mobility across Indian subdistricts and towns. Panel B shows a map of the geographic distribution of upward mobility across the wards of Delhi. Upward mobility ( $\mu_0^{50}$ ) is the average education rank attained by sons born to fathers who are in the bottom half of the father education distribution. Green areas have the highest mobility and red areas the lowest. The heat map legend applies to both panels of the figure. Source: SECC 2012.

**Figure 7**  
Correlates of Upward Mobility, 1985 Birth Cohort

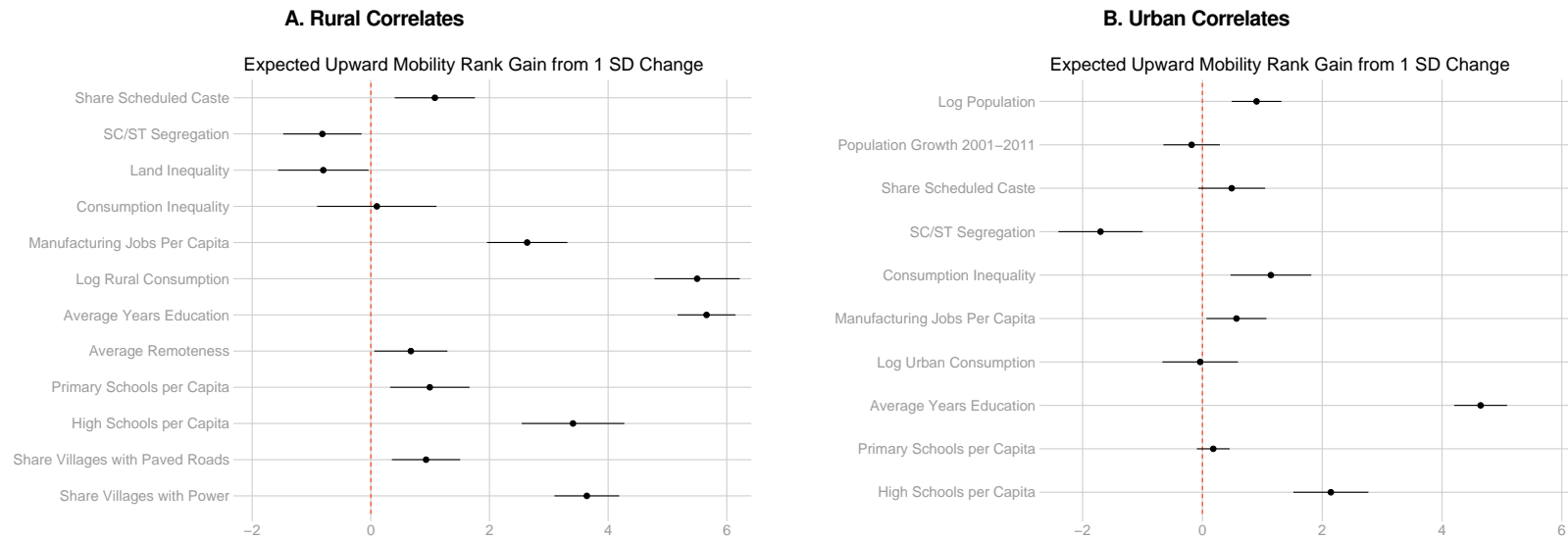


Figure 7 presents coefficients that illustrate the relationship between district-level characteristics and upward mobility. Bottom half mobility here ( $\mu_0^{50}$ ) is the average rank attained by sons born to fathers who are in the bottom half of the education distribution. Panel A presents correlations for rural areas, while Panel B presents correlations for urban areas. Source: SECC 2012.

**Figure 8**  
Jati Redesignation and Intergenerational Mobility

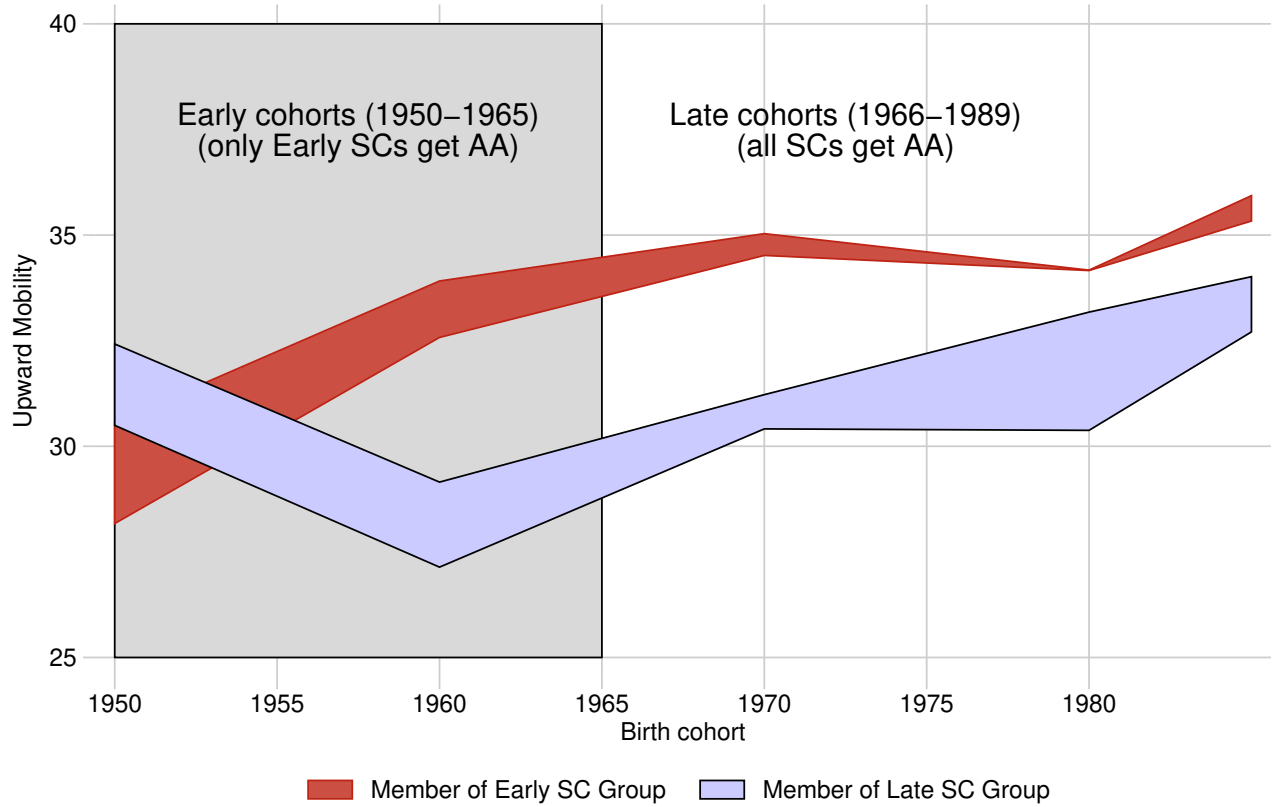


Figure 8 shows bounds on bottom half mobility  $\mu_0^{50}$  for two social groups in India. The dark red series shows upward mobility for groups that were designated as Scheduled Castes beginning in the 1950s. The light blue series shows upward mobility for groups that were not designated as Scheduled Castes until 1977; birth cohorts later than 1966 (where the vertical line is drawn) are those who were young enough to benefit. Source: IHDS 2012.

**Figure 9**  
Within-State and Within-District Mobility Gaps

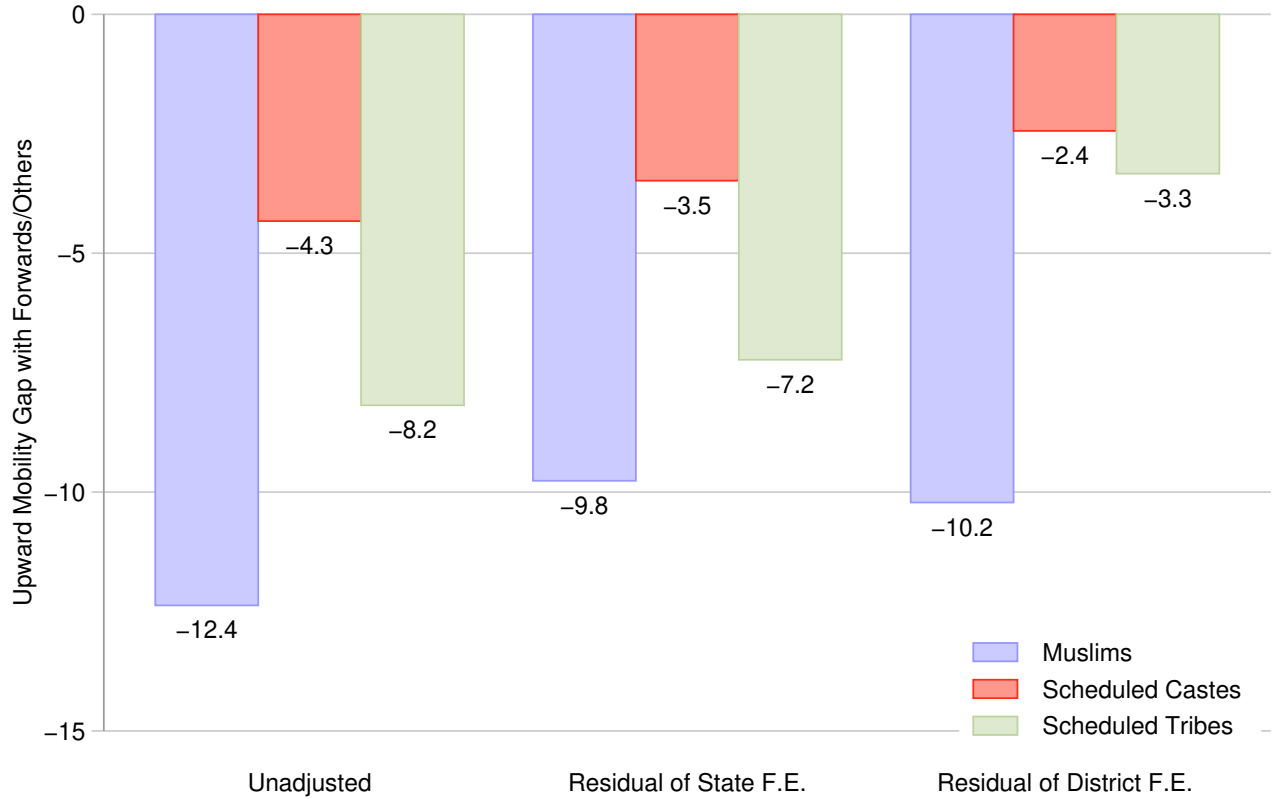


Figure 9 presents the upward mobility disadvantage relative to Forwards/Others faced by Muslims, Scheduled Castes and Scheduled Tribes. The first set of bars shows the aggregate mobility disadvantage of each group in rank terms. The second set of three bars shows the gaps calculated using within-state father and son education ranks. The third set shows gaps calculated using within-district father and son education ranks. Upward mobility in all cases is defined as  $\mu_0^{50} = E(y|x \in [0,50])$ , where  $x$  is parent rank and  $y$  is child rank. Upward mobility is partially identified; for simplicity, we show the midpoint of the bounds, which in all cases span less than a single rank. Source: IHDS 2012.

**Figure 10**  
Mincerian Returns and Intergenerational Mobility

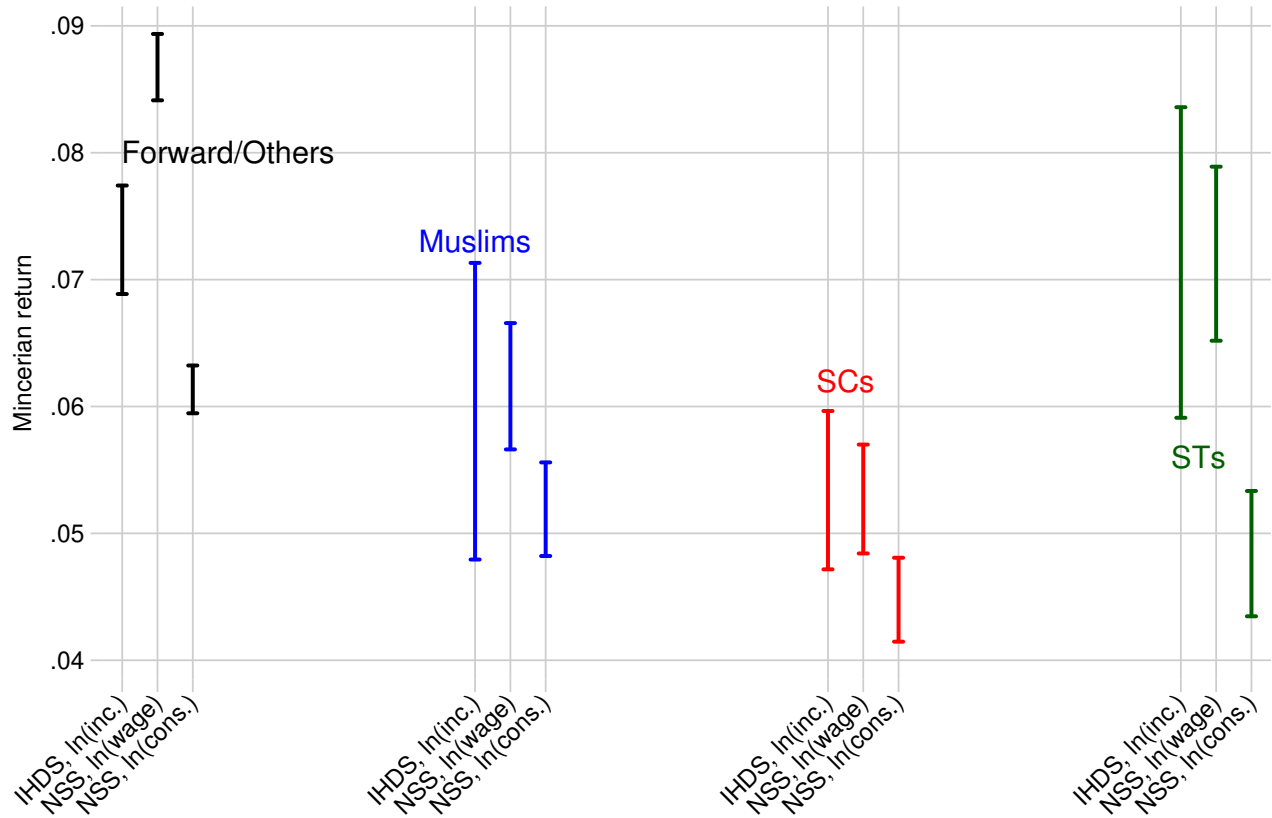
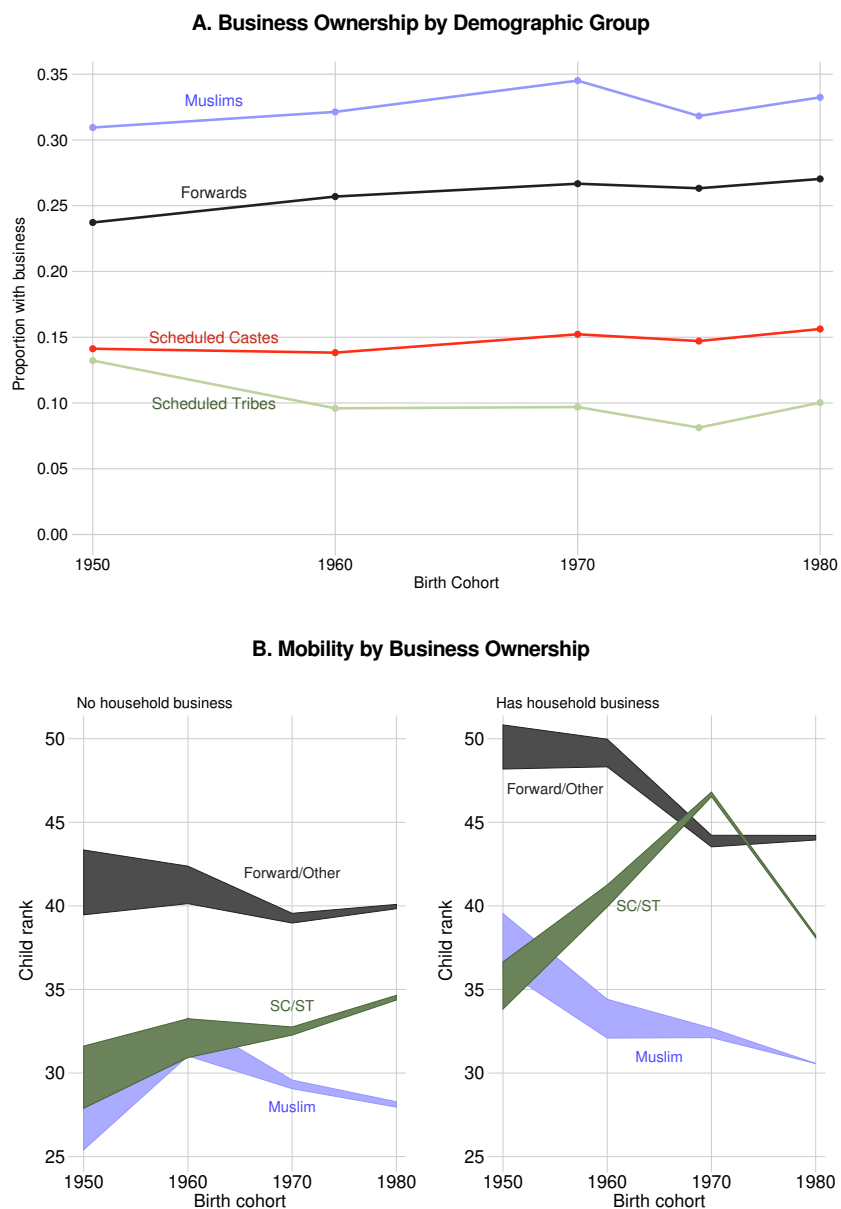


Figure 10 shows Mincerian returns to education for each major social group. Each set of points shows estimates of the Mincerian return to household log income (IHDS 2012), individual log wages (NSS 2012), and household log mean per capita income (NSS 2012).

**Figure 11**  
Business Ownership and Intergenerational Mobility



Panel A of Figure 11 shows the share of individuals who report that they work in their own business, by social group and time. Source: NSS (2012). Panel B shows bottom half mobility ( $\mu_0^{50}$ ) for the major social groups, separated by individual business ownership. Scheduled Castes and Tribes are pooled to increase power, since few members of either group own businesses.

**Table 1**  
Bin Sizes in Studies of Intergenerational Mobility

Study	Country	Birth Cohort of Son	Number of Parent Outcome Bins	Population Share in Largest Bin
Alesina et al. (2019)	Many countries in Africa <sup>43</sup>	1960–2005	5	83%
Card et al. (2018), Derenoncourt (2018)	USA	1920s	5	41% <sup>44</sup>
Dunn (2007)	Brazil	1972–1981	> 18	20% <sup>45</sup>
Emran and Shilpi (2011)	Nepal, Vietnam	1992–1995	2	83%
Güell et al. (2013)	Spain	~ 2001	9	27% <sup>46</sup>
Guest et al. (1989)	USA	~ 1880	7	53.2%
Hnatkovska et al. (2013)	India	1918–1988	5	Not reported
Knight et al. (2011)	China	1930–1984	5	29% <sup>47</sup>
Lindahl et al. (2012)	Sweden	1865–2005	8	34.5%
Long and Ferrie (2013)	Britain	~ 1850	4	57.6%
	Britain	~ 1949–55	4	54.2%
	USA	~ 1850–51	4	50.9%
	USA	~ 1949–55	4	48.3%
Piraino (2015)	South Africa	1964–1994	6	36%

Table 1 presents a review of papers analyzing educational and occupational mobility. The sample is not representative: we focus on papers where interval censoring may be a concern. The column indicating number of parent outcome bins refers to the number of categories for the parent outcome used in the main specification. The outcome is education in all studies with the exception of Long and Ferrie (2013) and Guest et al. (1989), where the outcome is occupation.

<sup>43</sup>Many countries are studied; the table shows illustrative statistics for Ethiopia, one of the largest countries in the sample.

<sup>44</sup>Source: Census Bureau (1940).

<sup>45</sup>Includes all people born after about 1990.

<sup>46</sup>Includes all people born after about 1960.

<sup>47</sup>This is the proportion of sons in 1976 who had not completed one year of education — an estimate of the proportion of fathers in 2002 with no education, which is not reported.

<sup>48</sup>Estimate is from the full population rather than just fathers.

<sup>49</sup>This reported estimate does not incorporate sampling weights; estimates with weights are not reported.



**Table 2**  
Changes in Upward Mobility Over Time

Panel A: Father/son pairs

	All groups	Forward/Others	Muslims	SCs	STs
1960–1969	[36.6, 39.0] {35.7, 39.8}	[41.8, 44.0] {40.6, 45.2}	[31.3, 33.6] {29.4, 35.5}	[32.9, 35.2] {31.5, 36.6}	[29.4, 31.3] {27.1, 33.6}
1980–1989	[37.1, 37.2] {36.4, 37.9}	[41.3, 41.3] {40.2, 42.4}	[28.9, 29.0] {27.5, 30.3}	[36.9, 37.0] {35.4, 38.6}	[33.1, 33.1] {31.1, 35.1}
Change over time	[-1.9, 0.6] {-2.9, 1.6}	[-2.7, -0.5] {-4.4, 1.1}	[-4.7, -2.3] {-7.1, 0.1}	[1.8, 4.1] {-0.4, 6.3}	[1.8, 3.7] {-1.3, 6.8}
Fraction overlapping bounds	0.818	0.322	0.050	0.102	0.118

Panel B: Father/daughter pairs

	All groups	Forward/Others	Muslims	SCs	STs
1960–1969	[34.9, 41.0] {34.1, 41.8}	[38.7, 44.8] {37.6, 46.0}	[33.5, 38.9] {31.7, 40.7}	[31.3, 36.8] {29.8, 38.3}	[31.4, 33.8] {29.0, 36.2}
1980–1989	[35.4, 35.5] {34.6, 36.2}	[38.0, 38.2] {36.8, 39.3}	[32.0, 33.5] {31.0, 34.6}	[32.9, 34.2] {31.6, 35.5}	[30.4, 30.5] {28.4, 32.4}
Change over time	[-5.6, 0.6] {-6.7, 1.7}	[-6.9, -0.5] {-8.5, 1.1}	[-6.9, -0.0] {-8.9, 2.0}	[-3.9, 2.9] {-5.9, 4.9}	[-3.4, -0.9] {-6.5, 2.1}
Fraction overlapping bounds	0.802	0.244	0.446	0.982	0.344

Table 2 shows estimates of full sample and subgroup bottom half mobility ( $\mu_0^{50}$ ) for the 1960–69 and 1980–89 birth cohorts. We show both bounds and 90% confidence sets on those bounds. The table also reports the confidence interval on the change in bottom half mobility between these two time periods. Source: IHDS (2012).

**Table 3**  
Group Differences in Upward Mobility

	F/O minus SC	F/O minus Muslim	SC minus Muslim
Father/son ( $\mu_0^{50}$ )	[4.6, 5.0]	[11.6, 12.1]	[6.9, 7.3]
	{2.8, 6.9}	{9.8, 13.9}	{4.5, 9.7}
Fraction overlapping bounds	0.000	0.000	0.000
Father/daughter ( $\mu_0^{50}$ )	[4.2, 4.5]	[5.1, 5.5]	[0.8, 1.1]
	{1.8, 6.9}	{3.0, 7.6}	{-2.1, 4.0}
Fraction overlapping bounds	0.000	0.000	0.509
Father/son ( $\mu_{50}^{100}$ )	[4.6, 5.0]	[11.6, 12.1]	[6.9, 7.3]
	{3.2, 6.5}	{5.7, 18.1}	{0.7, 13.4}
Fraction overlapping bounds	0.000	0.000	0.012
Father/son ( $\mu_{50}^{100}$ )	[4.6, 5.0]	[11.6, 12.1]	[6.9, 7.3]
	{3.2, 6.5}	{5.7, 18.1}	{0.7, 13.4}
Fraction overlapping bounds	0.000	0.000	0.329

Table 3 shows confidence intervals of subgroup differences in bottom half mobility in the 1980–89 birth cohorts. We show both bounds and 90% confidence sets on those bounds. Source: IHDS (2012).

**Table 4**  
Effect of Caste Redesignation on Upward Mobility

	(1)	(2)	(3)
Post * Late SC	6.163** (2.502)	6.764*** (1.555)	
1970-79 * Late SC			5.795* (2.422)
1980-89 * Late SC			6.476 (4.246)
N	4024	3746	4024
r <sup>2</sup>	0.33	0.34	0.33

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table 4 shows estimates from Equation 5.1, which describes the impact of scheduled caste redesignation on upward mobility. The dependent variable is the child education rank. The sample consists of SC sons of fathers with less than two years of education in the 1960s and 1970s, and SC sons of fathers with 2 or fewer years of education in the 1980s. Late SC is an indicator for jati groups that were added to Scheduled Caste lists in the caste redesignation of 1977. All estimations control for region\*cohort, jati\*region, jati\*cohort, birth year, and religion fixed effects, and are clustered at the jati and the region levels. Source: IHDS (2012).

**Table 5**  
Relationship Between Fertility and Subgroup Upward Mobility

	(1)	(2)	(3)
Muslim	-13.476*** (0.976)	-12.338*** (1.697)	-9.287*** (1.721)
Scheduled Caste	-4.163*** (0.749)	-2.608** (1.281)	-1.901 (1.268)
Scheduled Tribe	-9.075*** (1.076)	-8.040*** (1.851)	-8.291*** (1.829)
Urban	3.881*** (0.782)	3.812*** (1.276)	3.514*** (1.261)
Number of Siblings			-2.359*** (0.304)
N	6345	2347	2347
r <sup>2</sup>	0.11	0.15	0.18

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table 5 shows estimates from regressions of child education rank on social group indicators and an individual's number of siblings, a proxy for mother's fertility. The sample is limited to individuals born in 1985–89 to fathers with two or fewer years of education, or fathers in the bottom 50%. The outcome variable is thus a measure of bottom half mobility. Column 1 shows the estimation without the fertility measure for the full sample. Column 2 limits the data to the set of individuals for whom mother's fertility can be measured, and Column 3 adds the fertility variable. All regressions control for state fixed effects. Source: IHDS.

## A Appendix A: Additional Tables and Figures

**Table A1**  
Transition Matrices for Father and Son Education in India

**A: Sons Born 1950-59**

Father ed attained	Son highest education attained						
	< 2 yrs. (31%)	2-4 yrs. (11%)	Primary (17%)	Middle (13%)	Sec. (13%)	Sr. sec. (6%)	Any higher (8%)
<2 yrs. (60%)	0.47	0.12	0.17	0.11	0.09	0.03	0.03
2-4 yrs. (12%)	0.10	0.18	0.22	0.19	0.16	0.09	0.06
Primary (13%)	0.07	0.08	0.31	0.16	0.19	0.08	0.10
Middle (6%)	0.06	0.05	0.09	0.30	0.17	0.14	0.18
Secondary (5%)	0.03	0.02	0.04	0.12	0.37	0.11	0.30
Sr. secondary (2%)	0.02	0.00	0.03	0.11	0.11	0.35	0.38
Any higher ed (2%)	0.01	0.01	0.01	0.03	0.08	0.13	0.72

**B: Sons Born 1960-69**

Father ed attained	Son highest education attained						
	< 2 yrs. (27%)	2-4 yrs. (10%)	Primary (16%)	Middle (16%)	Sec. (14%)	Sr. sec. (7%)	Any higher (10%)
<2 yrs. (57%)	0.41	0.12	0.16	0.14	0.09	0.04	0.04
2-4 yrs. (13%)	0.12	0.17	0.18	0.22	0.15	0.08	0.08
Primary (14%)	0.09	0.05	0.26	0.18	0.20	0.09	0.13
Middle (6%)	0.06	0.04	0.09	0.29	0.21	0.13	0.19
Secondary (6%)	0.03	0.02	0.08	0.12	0.35	0.16	0.25
Sr. secondary (2%)	0.02	0.02	0.03	0.07	0.19	0.25	0.41
Any higher ed (2%)	0.01	0.01	0.02	0.03	0.09	0.11	0.73

**C: Sons Born 1970-79**

Father ed attained	Son highest education attained						
	< 2 yrs. (20%)	2-4 yrs. (8%)	Primary (17%)	Middle (18%)	Sec. (16%)	Sr. sec. (10%)	Any higher (12%)
<2 yrs. (50%)	0.33	0.10	0.19	0.17	0.12	0.05	0.04
2-4 yrs. (11%)	0.11	0.16	0.20	0.22	0.15	0.08	0.08
Primary (15%)	0.08	0.06	0.24	0.23	0.18	0.11	0.11
Middle (8%)	0.05	0.03	0.09	0.29	0.21	0.17	0.16
Secondary (9%)	0.03	0.02	0.06	0.12	0.31	0.19	0.27
Sr. secondary (3%)	0.01	0.01	0.02	0.08	0.17	0.29	0.42
Any higher ed (4%)	0.00	0.00	0.02	0.05	0.10	0.17	0.66

**D: Sons Born 1980-89**

Father ed attained	Son highest education attained						
	< 2 yrs. (12%)	2-4 yrs. (7%)	Primary (16%)	Middle (20%)	Sec. (16%)	Sr. sec. (12%)	Any higher (17%)
<2 yrs. (38%)	0.26	0.10	0.21	0.20	0.12	0.06	0.05
2-4 yrs. (11%)	0.08	0.17	0.19	0.24	0.15	0.09	0.08
Primary (17%)	0.05	0.04	0.22	0.23	0.20	0.13	0.13
Middle (12%)	0.03	0.02	0.10	0.28	0.20	0.17	0.20
Secondary (11%)	0.02	0.01	0.05	0.13	0.23	0.24	0.32
Sr. secondary (5%)	0.02	0.01	0.04	0.09	0.15	0.24	0.46
Any higher ed (5%)	0.01	0.01	0.02	0.05	0.10	0.16	0.65

Table A1 shows transition matrices by decadal birth cohort for Indian fathers and sons in the study.

**Table A2**  
Internal Consistency of Reports of Parents' Education

	Father-Son		Father-Daughter		Mother-Daughter	
	(1)	(2)	(3)	(4)	(5)	(6)
Age		0.005 (0.006)		-0.010 (0.012)		-0.002 (0.005)
Child years of education		0.012 (0.012)		0.041** (0.018)		0.003 (0.009)
Log household income		-0.002 (0.026)		-0.030 (0.051)		-0.016 (0.029)
Constant	0.057 (0.053)	-0.201 (0.364)	0.000 (0.091)	0.363 (0.651)	-0.016 (0.044)	0.210 (0.339)
N	1424	1421	545	545	892	891
r2	0.00	0.00	0.00	0.01	0.00	0.00

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table A2 shows measures of internal consistency when there are multiple reports of an individual's father in the IHDS. Each column is a regression of the level difference between two different measures of a parent's education. Columns 1, 3, and 5 show the average differences, and Columns 2, 4, and 6 regress that difference on individual characteristics to measure the extent to which they predict the discrepancy. Source: IHDS.

**Figure A1**  
Coresidence Rates by Age and Gender

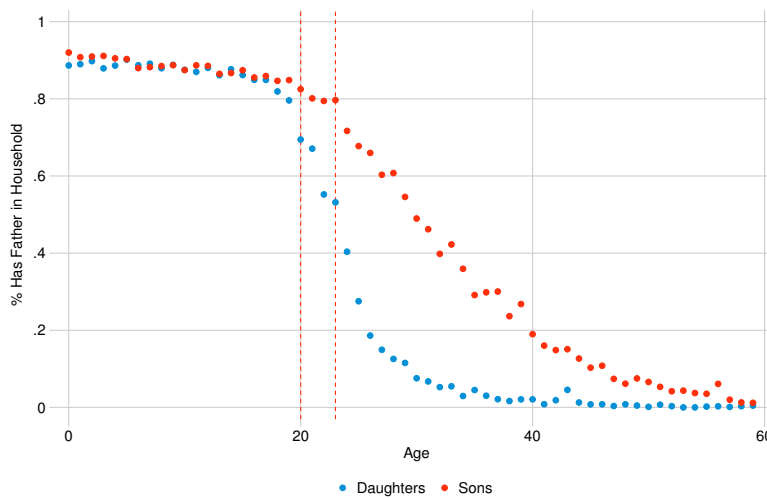


Figure A1 shows the share of individuals who live in the same household as their father as a function of gender and age.

**Figure A2**  
Bias in Mobility Estimates Based on Coresident Samples



Figure A2 shows the bias in a measure of upward mobility when non-coresident parent-child links are excluded. The mobility measure is bottom half mobility ( $\mu_0^{50}$ ), which is the expected child rank conditional on being born to a parent in the bottom half of the education distribution. We calculate bias as the coresident-only measure minus the full sample measure. Source: IHDS.



Figure A3  
Annotated Calculation of  $\mu_0^{50}$

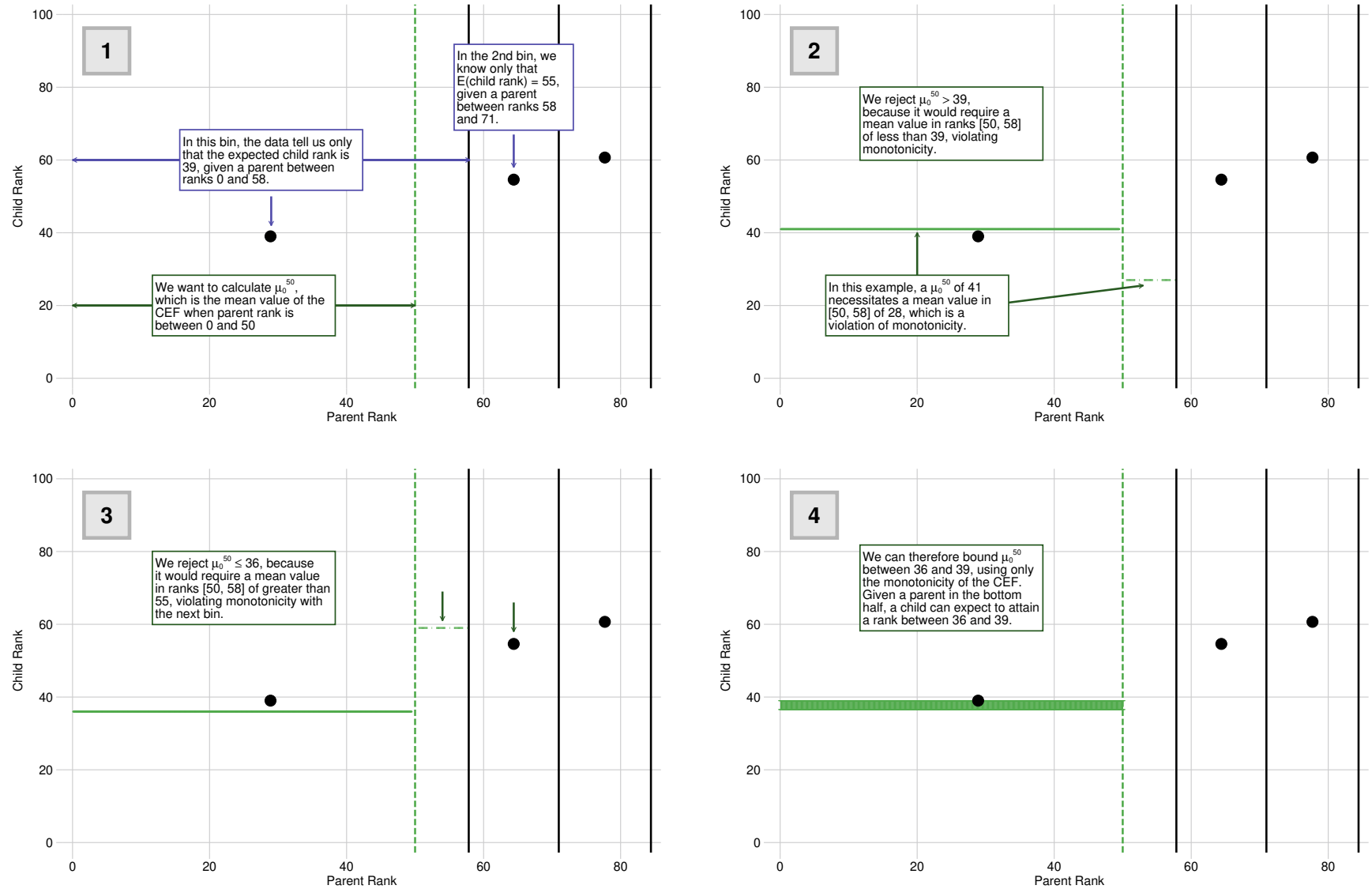


Figure A3 walks through the process of calculating bounds on  $\mu_0^{50} = E(y|x \in (0,50))$  using data from the 1960–69 birth cohort in India. Source: IHDS.



**Figure A4**  
Upward Interval Mobility ( $\mu_0^{50}$ ) for Mother-Son and Mother-Daughter Pairs

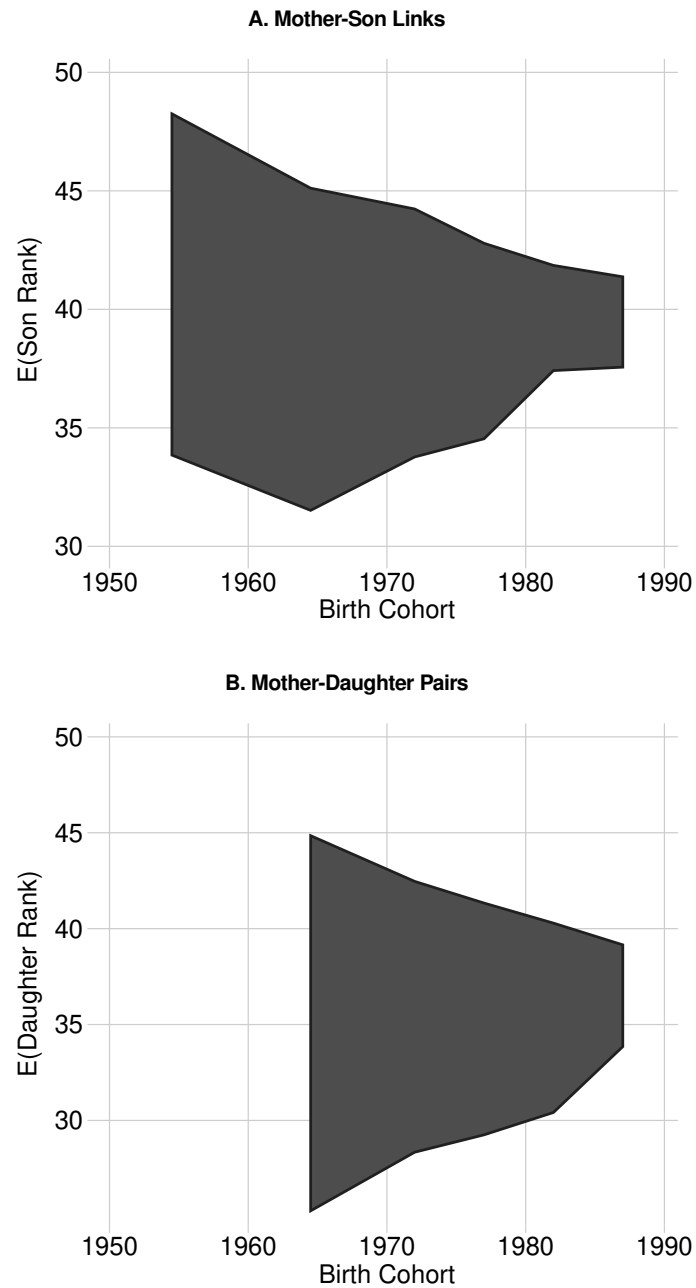


Figure A4 shows bounds on aggregate trends in intergenerational mobility, using cohorts born from 1950–59 through 1985–89, focusing on mother-son and mother-daughter links. The measure used is bottom half mobility ( $\mu_0^{50}$ ), which is the average rank attained by children born to parents who are in the bottom half of the education distribution. The bounds are very wide because of the large share of mothers who report bottom-coded education levels. Source: IHDS

**Figure A5**  
Robustness of Upward Mobility to Survivorship Bias

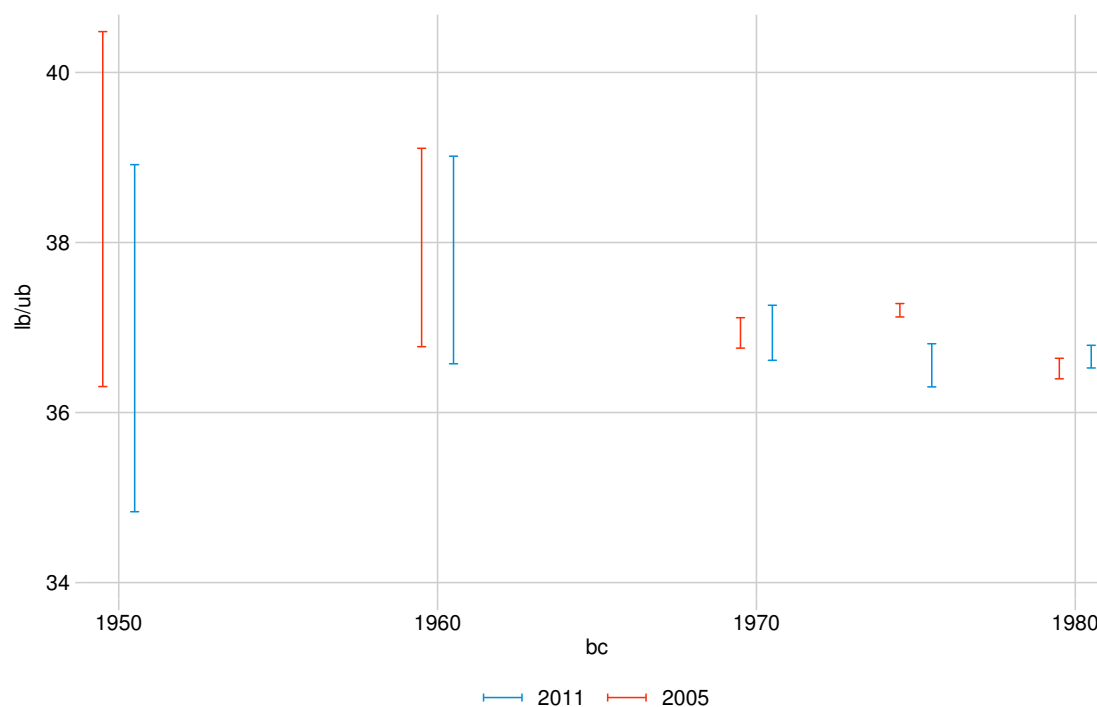


Figure A5 shows a test of survivorship bias in estimates of bottom half mobility in the 2012 IHDS. The figure shows estimates of bottom half mobility calculated for the same birth cohorts in the 2005 and 2012 IHDS. If there was substantial survivorship bias in the mobility measures, we would expect the estimates to differ across the two survivors because of the death of some of the respondents.

**Figure A6**  
Trends in Mobility by Subgroup, 1950–1989 Birth Cohorts  
Education Level Outcomes

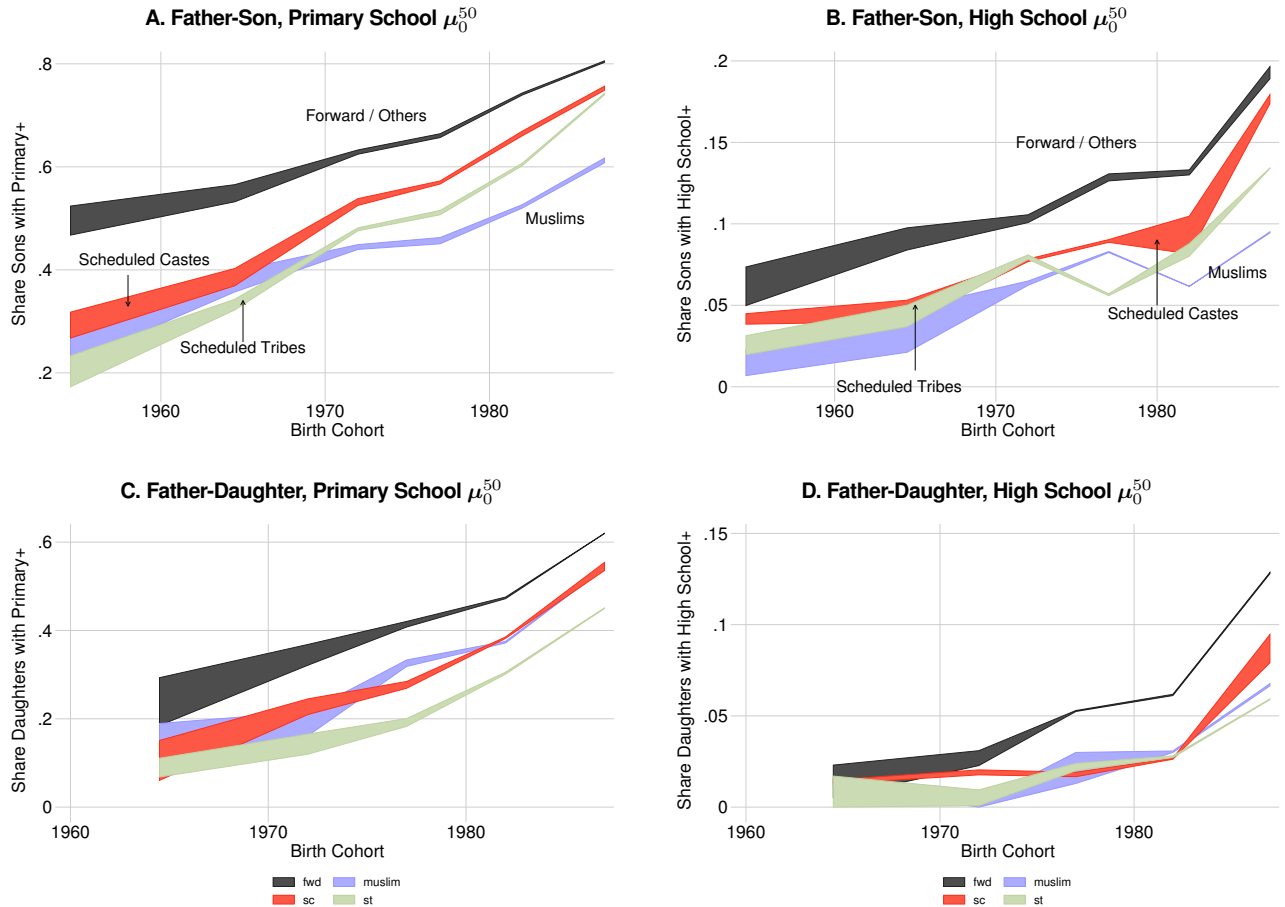


Figure A6 presents bounds on trends in intergenerational mobility, stratified by four prominent social groups in India: Scheduled Castes, Scheduled Tribes, Muslims, and Forward Castes/Others. The figure is analogous to Figure 5, but show the probability that a child attains a given education level (primary completion in Panels A and C, and secondary completion in Panels B and D), conditional on having a father in the bottom half of the father education distribution. Linked father-daughter education data are not available for the 1950–59 birth cohort. Source: IHDS.

**Figure A7**  
Urban/Rural Differences in Bottom Half Mobility for Population Subgroups

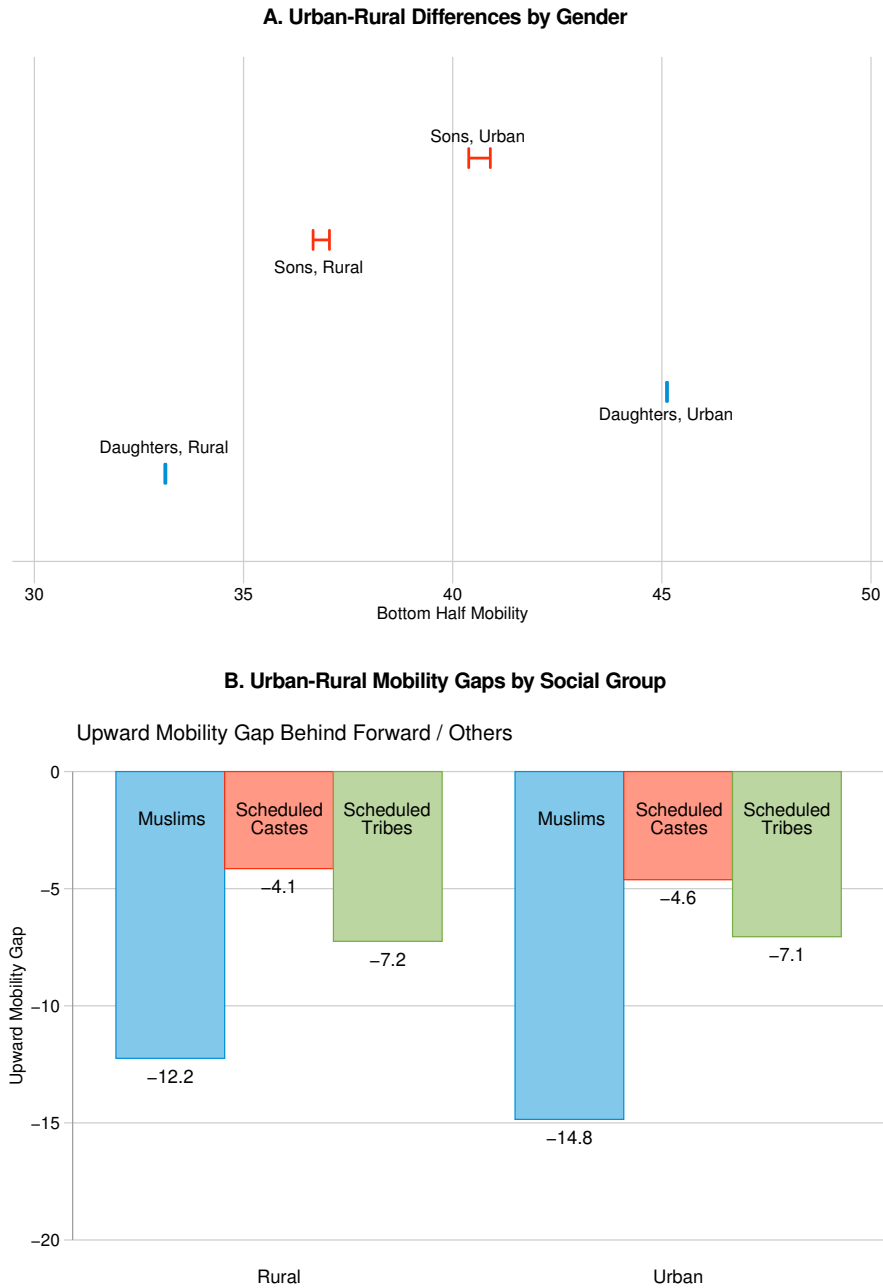


Figure A7A shows estimates of bottom half mobility ( $\mu_0^{50}$ ) for the 1985–89 birth cohort, disaggregated by gender and by urban/rural residence at the time of survey. Panel B of the figure shows the gap in upward mobility between each population subgroup and the Forward/Other group, disaggregated by urban and rural residence. Source: IHDS 2012.

## B Appendix B: Proofs and Numerical Optimizations

This section contains proofs of the propositions that define bounds on various measures of upward mobility, as well as the specification for the numerical optimization procedure for calculating these measures under CEFs with curvature constraints.

All of these proofs are based on Novosad et al. (2020), but are reproduced here for convenience. Novosad et al. (2020) is concerned with estimating bounds on  $E(y|x=i)$  and various functions of that CEF, where  $x$  is an interval-censored adult education rank and  $y$  is that same adult's mortality rate. This paper is concerned with the same mathematical problem, where  $x$  is an interval-censored parent education rank and  $y$  is a measure of child socioeconomic status. Note that the monotonicity condition here is reversed from that in Novosad et al. (2020). Here, we assume child status is ? in parent education rank; Novosad et al. (2020) assumes adult mortality is *decreasing* in adult education rank. The propositions and problem setups are otherwise the same.

### B.1 Proof of Proposition 1.

Let the function  $Y(x)=E(y|x)$  be defined on a known interval; without loss of generality, define this interval as  $x \in [0,100]$ . Assume  $Y(x)$  is integrable. We want to bound  $E(y|x)$  when  $x$  is known to lie in the interval  $[x_k, x_{k+1}]$ ; there are  $K$  such intervals. Define the expected value of  $y$  in bin  $k$  as

$$r_k = \int_{x_k}^{x_{k+1}} Y(x) f_k(x) dx.$$

Note that

$$r_k = E(y|x \in [x_k, x_{k+1}])$$

via the law of iterated expectations. Define  $r_0=0$  and  $r_{K+1}=100$ .

Restate the following assumptions from Manski and Tamer (2002):

$$\begin{aligned} P(x \in [x_k, x_{k+1}]) &= 1. && \text{(Assumption I)} \\ E(y|x) &\text{ must be weakly increasing in } x. && \text{(Assumption M)} \\ E(y|x \text{ is interval censored}) &= E(y|x). && \text{(Assumption MI)} \end{aligned}$$

From Manski and Tamer (2002), we have:

$$r_{k-1} \leq E(y|x) \leq r_{k+1} \quad \text{(Manski-Tamer bounds)}$$

Suppose also that

$$x \sim U(0,100). \quad \text{(Assumption U)}$$

In that case,

$$r_k = \frac{1}{x_{k+1} - x_k} \int_{x_k}^{x_{k+1}} Y(x) dx,$$

substituting the probability distribution function for the uniform distribution within bin  $k$ . Then we derive the following proposition.

**Proposition 1.** *Let  $x$  be in bin  $k$ . Under assumptions M, I, MI (Manski and Tamer, 2002) and*

$U$ , and without additional information, the following bounds on  $E(y|x)$  are sharp:

$$\begin{cases} r_{k-1} \leq E(y|x) \leq \frac{1}{x_{k+1}-x}((x_{k+1}-x_k)r_k - (x-x_k)r_{k-1}), & x < x_k^* \\ \frac{1}{x-x_k}((x_{k+1}-x_k)r_k - (x_{k+1}-x)r_{k+1}) \leq E(y|x) \leq r_{k+1}, & x \geq x_k^* \end{cases}$$

where

$$x_k^* = \frac{x_{k+1}r_{k+1} - (x_{k+1}-x_k)r_k - x_k r_{k-1}}{r_{k+1} - r_{k-1}}.$$

The intuition behind the proof is as follows. First, find the function  $z$  which meets the bin mean and is defined as  $r_{k-1}$  up to some point  $j$ . Because  $z$  is a valid CEF, the lower bound on  $E(y|x)$  is no larger than  $z$  up to  $j$ ; we then show that  $j$  is precisely  $x_k^*$  from the statement. For points  $x > x_k^*$ , we show that the CEF which minimizes the value at point  $x$  must be a horizontal line up to  $x$  and a horizontal line at  $r_{k+1}$  for points larger than  $x$ . But there is only one such CEF, given that the CEF must also meet the bin mean, and we can solve analytically for the minimum value the CEF can attain at point  $x$ . We focus on lower bounds for brevity, but the proof for upper bounds follows a symmetric structure.

*Part 1: Find  $x_k^*$ .* First define  $\mathcal{V}_k$  as the set of weakly increasing CEFs which meet the bin mean. Put otherwise, let  $\mathcal{V}_k$  be the set of  $v: [x_k, x_{k+1}] \rightarrow \mathbb{R}$  satisfying

$$r_k = \frac{1}{x_{k+1}-x_k} \int_{x_k}^{x_{k+1}} v(x) dx.$$

Now choose  $z \in \mathcal{V}_k$  such that

$$z(x) = \begin{cases} r_{k-1}, & x_k \leq x < j \\ r_{k+1}, & j \leq x \leq x_{k+1}. \end{cases}$$

Note that  $z$  and  $j$  both exist and are unique (it suffices to show that just  $j$  exists and is unique, as then  $z$  must be also). We can solve for  $j$  by noting that  $z$  lies in  $\mathcal{V}_k$ , so it must meet the bin mean. Hence, by evaluating the integrals,  $j$  must satisfy:

$$\begin{aligned} r_k &= \frac{1}{x_{k+1}-x_k} \int_{x_k}^{x_{k+1}} z(x) dx \\ &= \frac{1}{x_{k+1}-x_k} \left( \int_{x_k}^j r_{k-1} dx + \int_j^{x_{k+1}} r_{k+1} dx \right) \\ &= \frac{1}{x_{k+1}-x_k} ((j-x_k)r_{k-1} + (x_{k+1}-j)r_{k+1}). \end{aligned}$$

Note that these expressions invoke assumption U, as the integration of  $z(x)$  does not require any adjustment for the density on the  $x$  axis. For a more general proof with an arbitrary distribution of  $x$ , see section ??.

With some algebraic manipulations, we obtain that  $j = x_k^*$ .

*Part 2: Prove the bounds.* In the next step, we show that  $x_k^*$  is the smallest point at which no  $v \in \mathcal{V}_k$  can be  $r_{k-1}$ , which means that there must be some larger lower bound on  $E(y|x)$  for  $x \geq x_k^*$ . In other words, we prove that

$$x_k^* = \sup \left\{ x \mid \text{there exists } v \in \mathcal{V}_k \text{ such that } v(x) = r_{k-1} \right\}.$$

We must show that  $x_k^*$  is an upper bound and that it is the least upper bound.

First,  $x_k^*$  is an upper bound. Suppose that there exists  $j' > x_k^*$  such that for some  $w \in \mathcal{V}_k$ ,  $w(j') = r_{k-1}$ . Observe that by monotonicity and the bounds from Manski and Tamer (2002),  $w(x) = r_{k+1}$  for  $x \leq j'$ ; in other words, if  $w(j')$  is the mean of the mean of the prior bin, it can be no lower or higher than the mean of the prior bin up to point  $j'$ . But since  $j' > j$ , this means that

$$\int_{x_k}^{j'} w(x) dx < \int_{x_k}^{j'} z(x) dx,$$

since  $z(x) > w(x)$  for all  $h \in (j, j')$ . But recall that both  $z$  and  $w$  lie in  $\mathcal{V}_k$  and must therefore meet the bin mean; i.e.,

$$\int_{x_k}^{x_{k+1}} w(x) dx = \int_{x_k}^{x_{k+1}} z(x) dx.$$

But then

$$\int_{j'}^{x_{k+1}} w(x) dx > \int_{j'}^{x_{k+1}} z(x) dx.$$

That is impossible by the bounds from Manski and Tamer (2002), since  $w(x)$  cannot exceed  $r_{k+1}$ , which is precisely the value of  $z(x)$  for  $x \geq j$ .

Second,  $j$  is the least upper bound. Fix  $j' < j$ . From the definition of  $z$ , we have shown that for some  $h \in (j', j)$ ,  $z(h) = r_{k-1}$  (and  $z \in \mathcal{V}_k$ ). So any point  $j'$  less than  $j$  would not be a lower bound on the set — there is a point  $h$  larger than  $j'$  such that  $z(h) = r_{k-1}$ .

Hence, for all  $x < x_k^*$ , there exists a function  $v \in \mathcal{V}_k$  such that  $v(x) = r_{k-1}$ ; the lower bound on  $E(y|x)$  for  $x < x_k^*$  is no greater than  $r_{k-1}$ . By choosing  $z'$  with

$$z'(x) = \begin{cases} r_{k-1}, & x_k \leq x \leq j \\ r_{k+1}, & j < x \leq x_{k+1}, \end{cases}$$

it is also clear that at  $x_k^*$ , the lower bound is no larger than  $r_{k-1}$  (and this holds in the proposition itself, substituting in  $x_k^*$  into the lower bound in the second equation).

Now, fix  $x' \in (x_k^*, x_{k+1}]$ . Since  $x_k^*$  is the supremum, there is no function  $v \in \mathcal{V}_k$  such that  $v(x') = r_{k-1}$ . Thus for  $x' > x_k^*$ , we seek a sharp lower bound larger than  $r_{k-1}$ . Write this lower bound as

$$Y_{x'}^{min} = \min \left\{ v(x') \text{ for all } v \in \mathcal{V}_k \right\},$$

where  $Y_{x'}^{min}$  is the smallest value attained by any function  $v \in \mathcal{V}_k$  at the point  $x'$ .

We find this  $Y_{x'}^{min}$  by choosing the function which maximizes every point after  $x'$ , by attaining the value of the subsequent bin. The function which minimizes  $v(x')$  must be a horizontal line up to this point.

Pick  $\tilde{z} \in \mathcal{V}_k$  such that

$$\tilde{z}(x) = \begin{cases} \underline{Y}, & x_k \leq x' \\ r_{k+1}, & x' < x_{k+1} \leq x_{k+1} \end{cases}.$$

By integrating  $\tilde{z}(x)$ , we claim that  $\underline{Y}$  satisfies the following:

$$\frac{1}{x_{k+1}-x_k}((x'-x_k)\underline{Y}+(x_{k+1}-x')r_{k+1})=r_k.$$

As a result,  $\underline{Y}$  from this expression exists and is unique, because we can solve the equation. Note that this integration step also requires that the distribution of  $x$  be uniform, and we generalize this argument in ??.

By similar reasoning as above, there is no  $Y' < \underline{Y}$  such that there exists  $w \in \mathcal{V}_k$  with  $w(x') = Y'$ . Otherwise there must be some point  $x > x'$  such that  $w(x') > r_{k+1}$  in order that  $w$  matches the bin means and lies in  $\mathcal{V}_k$ ; the expression for  $\underline{Y}$  above maximizes every point after  $x'$ , leaving no additional room to further depress  $\underline{Y}$ .

Formally, suppose there exists  $w \in \mathcal{V}_k$  such that  $w(x') = Y' < \underline{Y}$ . Then  $w(x') < \tilde{z}(x')$  for all  $x < x'$ , since  $w$  is monotonic. As a result,

$$\int_{x_k}^{x'} \tilde{z}(x)dx > \int_{x_k}^{x'} w(x)dx.$$

But recall that

$$\int_{x_k}^{x_{k+1}} w(x)dx = \int_{x_k}^{x_{k+1}} \tilde{z}(x)dx,$$

so

$$\int_{x'}^{x_{k+1}} w(x)dx > \int_{x'}^{x_{k+1}} \tilde{z}(x)dx.$$

This is impossible, since  $\tilde{z}(x) = r_{k+1}$  for all  $x > x'$ , and by Manski and Tamer (2002),  $w(x) \leq r_{k+1}$  for all  $w \in \mathcal{V}_k$ . Hence there is no such  $w \in \mathcal{V}_k$ , and therefore  $\underline{Y}$  is smallest possible value at  $x'$ , i.e.  $\underline{Y} = Y_{x'}^{min}$ .

By algebraic manipulations, the expression for  $\underline{Y} = Y_x^{min}$  reduces to

$$Y_x^{min} = \frac{(x_{k+1}-x_k)r_k - (x_{k+1}-x)r_{k+1}}{x-x_k}, \quad x \geq x_k^*.$$

The proof for the upper bounds uses the same structure as the proof of the lower bounds.

Finally, the body of this proof gives sharpness of the bounds. For we have introduced a CEF  $v \in \mathcal{V}_k$  that obtains the value of the upper and lower bound for any point  $x \in [x_k, x_{k+1}]$ . For any value  $y$  within the bounds, one can generate a CEF  $v \in \mathcal{V}_k$  such that  $v(x) = y$ .  $\square$

## B.2 Proof of Bounds on $\mu_a^b$ (Proposition 2)

Define

$$\mu_a^b = \frac{1}{b-a} \int_a^b E(y|x)di.$$

Let  $Y_x^{min}$  and  $Y_x^{max}$  be the lower and upper bounds respectively on  $E(y|x)$  given by Proposition 1. We seek to bound  $\mu_a^b$  when  $x$  is observed only in discrete intervals.



**Proposition 2.** Let  $b \in [x_k, x_{k+1}]$  and  $a \in [x_h, x_{h+1}]$  with  $a < b$ . Let assumptions  $M$ ,  $I$ ,  $MI$  (?) and  $U$  hold. Then, if there is no additional information available, the following bounds are sharp:

$$\begin{cases} Y_b^{min} \leq \mu_a^b \leq Y_a^{max}, & h=k \\ \frac{r_h(x_k-a)+Y_b^{min}(b-x_k)}{b-a} \leq \mu_a^b \leq \frac{Y_a^{max}(x_k-a)+r_k(b-x_k)}{b-a}, & h+1=k \\ \frac{r_h(x_{h+1}-a)+\sum_{\lambda=h+1}^{k-1} r_\lambda(x_{\lambda+1}-x_\lambda)+Y_b^{min}(b-x_k)}{b-a} \leq \mu_a^b \leq \frac{Y_a^{max}(x_{h+1}-a)+\sum_{\lambda=h+1}^{k-1} r_\lambda(x_{\lambda+1}-x_\lambda)+r_k(b-x_k)}{b-a}, & h+1 < k. \end{cases}$$

The order of the proof is as follows. If  $a$  and  $b$  lie in the same bin, then  $\mu_a^b$  is maximized only if the CEF is minimized prior to  $a$ . As in the proof of proposition 1, that occurs when the CEF is a horizontal line at  $Y_x^{min}$  up to  $a$ , and a horizontal line  $Y_x^{max}$  at and after  $a$ . If  $a$  and  $b$  lie in separate bins, the value of the integral in bins that are contained between  $a$  and  $b$  is determined by the observed bin means. The portions of the integral that are not determined are maximized by a similar logic, since they both lie within bins. We prove the bounds for maximizing  $\mu_a^b$ , but the proof is symmetric for minimizing  $\mu_a^b$ .

*Part 1: Prove the bounds if  $a$  and  $b$  lie in the same bin.* We seek to maximize  $\mu_a^b$  when  $a, b \in [x_k, x_{k+1}]$ . This requires finding a candidate CEF  $v \in \mathcal{V}_k$  which maximizes  $\int_a^b v(x)dx$ . Observe that the function  $v(x)$  defined as

$$v(x) = \begin{cases} Y_a^{min}, & x_k \leq x < a \\ Y_a^{max}, & a \leq x \leq x_{k+1} \end{cases}$$

has the property that  $v \in \mathcal{V}_k$ . For if  $a \geq x_k^*$ ,  $v = \tilde{z}$  from the second part of the proof of proposition 1. If  $a < x_k^*$ , the CEF in  $\mathcal{V}_k$  which yields  $Y_a^{max}$  is precisely  $v$  (by a similar argument which delivers the upper bounds in proposition 1).

This CEF maximizes  $\mu_a^b$ , because there is no  $w \in \mathcal{V}_k$  such that

$$\frac{1}{b-a} \int_a^b w(x)dx > \frac{1}{b-a} \int_a^b v(x)dx.$$

Note that for any  $w \in \mathcal{V}_k$ ,  $\frac{1}{x_{k+1}-x_k} \int_{x_k}^{x_{k+1}} w(x)dx = \frac{1}{x_{k+1}-x_k} \int_{x_k}^{x_{k+1}} v(x)dx = r_k$ . Hence in order that  $\int_a^b w(x)dx > \int_a^b v(x)dx$ , there are two options. The first option is that

$$\int_{x_k}^a w(x)dx < \int_{x_k}^a v(x)dx.$$

That is impossible, since there is no room to depress  $w$  given the value of  $v$  after  $a$ . If  $a < x_k^*$ , then it is clear that there is no  $w$  giving a larger  $\mu_a^b$ , since  $r_{k-1} \leq w(x)$  for  $x_{k-1} \leq x \leq a$ , so  $w$  is bounded below by  $v$ . If  $a \geq x_k^*$ , then  $v(x) = r_{k+1}$  for all  $a \leq x \leq x_{k+1}$ . That would leave no room to depress  $w$  further; if  $\int_{x_k}^a w(x)dx < \int_{x_k}^a v(x)dx$ , then  $\int_a^{x_{k+1}} w(x)dx > \int_a^{x_{k+1}} v(x)dx$ , which cannot be the case if  $v = r_{k+1}$ , by the bounds given in ?.

The second option is that

$$\int_b^{x_k} w(x)dx < \int_b^{x_k} v(x)dx.$$

This is impossible due to monotonicity. For if  $\int_a^b w(x)dx > \int_a^b v(x)dx$ , then there must be some point  $x' \in [a, b]$  such that  $w(x') > v(x')$ . By monotonicity,  $w(x) > v(x)$  for all  $x \in [x', x_{k+1}]$  since

$v(x) = Y_a^{max}$  in that interval. As a result,

$$\int_b^{x_k} w(x)dx > \int_b^{x_k} v(x)dx,$$

since  $b \in (x', x_{k+1})$ . (If  $b = x_{k+1}$ , then only the first option would allow  $w$  to maximize the desired  $\mu_a^b$ .)

Therefore, there is no such  $w$ , and  $v$  indeed maximizes the desired integral. Integrating  $v$  from  $a$  to  $b$ , we obtain that the upper bound on  $\mu_a^b$  is  $\frac{1}{b-a} \int_a^b Y_a^{max} dx = Y_a^{max}$ . Note that there may be many functions which maximize the integral; we only needed to show that  $v$  is one of them.

To prove the lower bound, use an analogous argument.

*Part 2: Prove the bounds if  $a$  and  $b$  do not lie in the same bin.* We now generalize the set up and permit  $a, b \in [0, 100]$ . Let  $\mathcal{V}$  be the set of weakly increasing functions such that  $\frac{1}{x_{k+1} - x_k} \int_{x_k}^{x_{k+1}} v(x)dx = r_k$  for all  $k \leq K$ . In other words,  $\mathcal{V}$  is the set of functions which match the means of every bin. Now observe that for all  $v \in \mathcal{V}$ ,

$$\begin{aligned} \mu_a^b &= \frac{1}{b-a} \int_a^b v(x)dx \\ &= \frac{1}{b-a} \left( \int_a^{x_{h+1}} v(x)dx + \int_{x_{h+1}}^{x_k} v(x)dx + \int_{x_k}^b v(x)dx \right), \end{aligned}$$

by a simple expansion of the integral.

But for all  $v \in \mathcal{V}$ ,

$$\int_{x_{h+1}}^{x_k} v(x)dx = \sum_{\lambda=h+1}^{k-1} r_\lambda (x_{\lambda+1} - x_\lambda)$$

if  $h+1 < k$  and

$$\int_{x_{h+1}}^{x_k} v(x)dx = 0$$

if  $h+1 = k$ . For in bins completely contained inside  $[a, b]$ , there is no room for any function in  $\mathcal{V}$  to vary; they all must meet the bin means.

We proceed to prove the upper bound. We split this into two portions: we wish to maximize  $\int_a^{x_{h+1}} v(x)dx$  and we also wish to maximize  $\int_{x_k}^b v(x)dx$ . The values of these objects are not codependent. But observe that the CEFs  $v \in \mathcal{V}_k$  which yield upper bounds on these integrals are the very same functions which yield upper bounds on  $\mu_a^{x_{h+1}}$  and  $\mu_{x_k}^b$ , since  $\mu_s^t = \frac{1}{t-s} \int_s^t v(x)dx$  for any  $s$  and  $t$ . Also notice that  $a$  and  $x_{h+1}$  both lie in bin  $h$ , while  $b$  and  $x_k$  both lie in bin  $k$ , so we can make use of the first portion of this proof.

In part 1, we showed that the function  $v \in \mathcal{V}$ ,  $v: [x_h, x_{h+1}] \rightarrow \mathbb{R}$ , which maximizes  $\mu_a^{x_{h+1}}$  is

$$v(x) = \begin{cases} Y_a^{min}, & x_h \leq x < a \\ Y_a^{max}, & a \leq x \leq x_{h+1}. \end{cases}$$

As a result

$$\max_{v \in \mathcal{V}} \left\{ \int_a^{x_{h+1}} v(x)dx \right\} = \int_a^{x_{h+1}} Y_a^{max} dx = Y_a^{max} (x_{h+1} - a).$$

Similarly, observe that  $x_k$  and  $b$  lie in the same bin, so the function  $v: [x_k, x_{k+1}] \rightarrow \mathbb{R}$ , with  $v \in \mathcal{V}$  which maximizes  $\int_{x_k}^b v(x)dx$  must be of the form

$$v(x) = \begin{cases} Y_{x_k}^{min}, & x_k \leq x < a \\ Y_{x_k}^{max}, & b \leq x \leq x_{k+1}. \end{cases}$$

With identical logic,

$$\max_{v \in \mathcal{V}} \left\{ \int_{x_k}^b v(x)dx \right\} = \int_{x_k}^b Y_{x_k}^{max} dx = Y_{x_k}^{max}(b - x_k).$$

And by proposition 1,  $x_k \leq x_k^*$  so  $Y_{x_k}^{max} = r_k$ . (Note that if  $x_k = x_k^*$ , substituting  $x_k^*$  into the second expression of proposition 1 still yields that  $Y_{x_k}^{max} = r_k$ .)

Now we put all these portions together. First let  $h+1 = k$ . Then  $\int_{x_{h+1}}^{x_k} v(x)dx = 0$ , so we maximize  $\mu_a^b$  by

$$\frac{1}{b-a} (Y_a^{max}(x_{h+1} - a) + r_k(b - x_k)).$$

Similarly, if  $h+1 < k$  and there are entire bins completely contained in  $[a, b]$ , then we maximize  $\mu_a^b$  by

$$\frac{1}{b-a} \left( Y_a^{max}(x_{h+1} - a) + \sum_{\lambda=h+1}^{k-1} r_\lambda(x_{\lambda+1} - x_\lambda) + r_k(b - x_k) \right).$$

The lower bound is proved analogously. Sharpness is immediate, since we have shown that the CEF which delivers the endpoints of the bounds lies in  $\mathcal{V}$ . As a result, there is a function delivering any intermediate value for the bounds.  $\square$

### B.3 Numerical Bounds Interval-Censored CEF and Functions of the CEF

This paper describes two circumstances where bounds on  $E(Y|X=i)$  (the CEF) or some function of that CEF may be difficult to calculate analytically. In the first case, we may wish to impose a curvature constraint on the CEF; for example, we may wish to rule out candidate CEFs that have large discrete jumps in value when they cross arbitrary percentiles that are not associated with completing a given schooling level. In the second case, we may wish to calculate functions of the CEF with very complex analytical forms, such as the slope of the best linear approximator to the CEF, an analog to the rank-rank gradient. In these cases, a numerical optimization may be used to derive bounds on the statistic of interest.

We solve this problem numerically by calculating the set of solutions to a pair of minimization and maximization problems that take the following structure. We write the conditional expectation function in the form  $Y(x) = s(x, \gamma)$ , where  $\gamma$  is a finite-dimensional vector that lies in parameter space  $G$  and serves to parameterize the CEF through the function  $s$ . For example, we could estimate the parameters of a linear approximation to the CEF by defining  $s(x, \gamma) = \gamma_0 + \gamma_1 * x$ . We can approximate an arbitrary nonparametric CEF by defining  $\gamma$  as a vector of discrete values that give the value of the CEF in each of  $N$  partitions; we take this approach in our numerical optimizations, setting  $N$  to 100.<sup>50</sup>

<sup>50</sup>For example,  $s(x, \gamma_{50})$  would represent  $E(y|x \in [49, 50])$ .

Any statistic  $m$  that is a single-valued function of the CEF, such as the average value of the CEF in an interval  $(\mu_a^b)$ , or the slope of the best fit line to the CEF, can be defined as  $m(\gamma) = M(s(x, \gamma))$ .

Let  $f(x)$  again represent the probability distribution of  $x$ . Define  $\Gamma$  as the set of parameterizations of the CEF that obey monotonicity and minimize mean squared error with respect to the observed interval data:

$$\Gamma = \underset{g \in G}{\operatorname{argmin}} \sum_{k=1}^K \left\{ \int_{x_k}^{x_{k+1}} f(x) dx \left( \left( \frac{1}{\int_{x_k}^{x_{k+1}} f(x) dx} \int_{x_k}^{x_{k+1}} s(x, g) f(x) dx \right) - \bar{r}_k \right)^2 \right\} \quad (\text{B.1})$$

such that

$s(x, g)$  is weakly increasing in  $x$ . (Monotonicity)

Decomposing this expression,  $\frac{1}{\int_{x_k}^{x_{k+1}} f(x) dx} \int_{x_k}^{x_{k+1}} s(x, g) f(x) dx$  is the mean value of  $s(x, g)$  in bin  $k$ , and  $\int_{x_k}^{x_{k+1}} f(x) dx$  is the width of bin  $k$ . The minimand is thus a bin-weighted MSE.<sup>51</sup> Recall that for the rank distribution,  $x_1 = 0$  and  $x_{K+1} = 100$ .

The bounds on  $m(\gamma)$  are therefore:

$$\begin{aligned} m^{\min} &= \inf \{m(\gamma) \mid \gamma \in \Gamma\} \\ m^{\max} &= \sup \{m(\gamma) \mid \gamma \in \Gamma\}. \end{aligned} \quad (\text{B.2})$$

For example, bounds on the best linear approximation to the CEF can be defined by the following process. First, consider the set of all CEFs that satisfy monotonicity and minimize mean-squared error with respect to the observed bin means.<sup>52</sup> Next, compute the slope of the best linear approximation to each CEF. The largest and smallest slope constitute  $m^{\min}$  and  $m^{\max}$ . Note that this definition of the best linear approximator to the CEF corresponds to the *least squares set* defined by Ponomareva and Tamer (2011).

This problem setup is valid both for calculating complex functions of the CEF, like the rank-rank gradient, and for imposing additional constraints, like a curvature constraint.

### B.3.1 Specifying a Curvature Constraint on the CEF

The set of CEFs that describe the upper and lower bounds in Proposition 1 are step functions with substantial discontinuities. If such functions are implausible descriptions of the data, then the researcher may wish to impose an additional constraint on the curvature of the CEF, which will generate tighter bounds. For example, examination of the parent-child *income* rank relationships (which can be estimated at each of 100 income ranks) from developed countries suggest no such discontinuities (Chetty et al., 2014a). Alternately, in a context where continuity has a strong theoretical underpinning but monotonicity does not, a curvature constraint can substitute for a monotonicity constraint and in many cases deliver useful bounds.

We consider a curvature restriction with the following structure:

$$s(x, \gamma) \text{ is twice-differentiable and } |s''(x, \gamma)| \leq \bar{C}. \quad (\text{Curvature Constraint})$$

<sup>51</sup>While we choose to use a weighted mean squared error penalty, in principle  $\Gamma$  could use other penalties.

<sup>52</sup>In many cases, and in all of our applications, there will exist many such CEFs that exactly match the observed data and the minimum mean-squared error will be zero.

This is analogous to imposing that the first derivative is Lipschitz.<sup>53</sup> Depending on the value of  $\overline{C}$ , this constraint may or may not bind.

The most restrictive curvature constraint,  $\overline{C}=0$ , is analogous to the assumption that the CEF is linear. Note that the best linear approximation to  $E(rank_{child}|rank_{parent})$  is a canonical mobility estimator. A moderate curvature constraint is therefore a *less* restrictive assumption than the approach taken in many studies. We discuss the choice of curvature restriction below.

### B.3.2 Numerical Calculation of CEF Bounds

This section describes a method to numerically solve the constrained optimization problem suggested by Equations B.1 and B.2. We take a nonparametric approach for generality: explicitly parameterizing an unknown CEF with limited data is unsatisfying and could yield inaccurate results if the interval censoring conceals a non-linear within-bin CEF. In the context of intergenerational mobility, many CEFs of interest do not appear to obey a familiar parametric form (see the figures in ? and Bratberg et al. (2015)).

To make the problem numerically tractable, we solve the discrete problem of identifying the feasible mean value taken by  $E(y|x)$  in each of  $N$  discrete partitions of  $x$ . We thus assume  $E(y|x)=s(x,\gamma)$ , where  $\gamma$  is a vector that defines the mean value of the CEF in each of the  $N$  partitions. We use  $N=100$  in our analysis, corresponding to integer rank bins, but other values may be useful depending on the application. In other words, we will numerically calculate upper and lower bounds on  $E(y|x \in [0,1])$ ,  $E(y|x \in [1,2])$ , ...,  $E(y|x \in [99,100])$ . Given continuity in the latent function, the discretized CEF will be a very close approximation of the continuous CEF; in our analysis, increasing the value of  $N$  increases computation time but does not change any of our results.

We solve the problem through a two-step process. Define a  $N$ -valued vector  $\hat{\gamma}$  as a candidate CEF. First, we calculate the minimum MSE from the constrained optimization problem given by Equation B.1. We then run a second pair of constrained optimization problems that respectively minimize and maximize the value of  $m(\hat{\gamma})$ , with the additional constraint that the MSE is equal to the value obtained in the first step, denoted  $\underline{\text{MSE}}$ . Equation B.3 shows the second stage setup to calculate the lower bound on  $m(\hat{\gamma})$ . Note that this particular setup is specific to the uniform rank distribution, but setups with other distributions would be similar.

$$m^{min} = \min_{\hat{\gamma} \in [0,100]^N} m(\hat{\gamma}) \quad (\text{B.3})$$

such that

$$s(x, \hat{\gamma}) \text{ is weakly increasing in } x \quad (\text{Monotonicity})$$

$$|s''(x, \hat{\gamma})| \leq \overline{C}, \quad (\text{Curvature})$$

$$\sum_{k=1}^K \left[ \frac{\|X_k\|}{100} \left( \left( \frac{1}{\|X_k\|} \sum_{x \in X_k} s(x, \hat{\gamma}) \right) - \bar{r}_k \right)^2 \right] = \underline{\text{MSE}} \quad (\text{MSE Minimization})$$

---

<sup>53</sup>Let  $X, Y$  be metric spaces with metrics  $d_X, d_Y$  respectively. The function  $f: X \rightarrow Y$  is *Lipschitz continuous* if there exists  $K \geq 0$  such that for all  $x_1, x_2 \in X$ ,

$$d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2).$$

$X_k$  is the set of discrete values of  $x$  between  $x_k$  and  $x_{k+1}$  and  $\|X_k\|$  is the width of bin  $k$ . The complementary maximization problem obtains the upper bound on  $m(\hat{\gamma})$ .

Note that setting  $m(\gamma) = \gamma_x$  (the  $x^{\text{th}}$  element of  $\gamma$ ) obtains bounds on the value of the CEF at point  $x$ . Calculating this for all ranks  $x$  from 1 to 100 generates analogous bounds to those derived in Proposition 1, but satisfying the additional curvature constraint. Similarly  $m(\gamma) = \frac{1}{b-a} \sum_{x=a}^b \gamma_x$  obtains bounds on  $\mu_a^b$ .

The numerical method can easily permit the curvature constraint to vary over the CEF. For example, one might believe that there are discontinuities in the CEF at bin boundaries, due to sheepskin effects (Hungerford and Solon, 1987). In other settings, researchers might impose that the CEF has a large (but finite) curvature in one portion of its domain and be more constrained elsewhere. In this paper, we focus on a uniform curvature constraint for concision.

Matlab code to run these numerical optimizations for more complex functions (as well as with the curvature constraints described below) are posted on the corresponding author’s web site.<sup>54</sup>

---

<sup>54</sup>See <https://github.com/paulnov/nra-bounds/>.

## C Appendix C: CEF Bounds When $x$ and $y$ are Interval-Censored

In the main part of the paper, we focus on bounding a function  $Y(x) = E(y|x)$  when  $y$  is observed without error, but  $x$  is observed with interval censoring. In this section, we modify the setup to consider simultaneous interval censoring in the conditioning variable  $x$  and in observed outcomes  $y$ . This arises, for example, in the study of educational mobility, where latent education ranks of both parents and children are observed with interval censoring.

We first present a setup that takes a similar approach to the bounding method presented in Section 4. We can define bounds on the CEF  $E(y|x)$  when both  $y$  and  $x$  are interval-censored as a solution to a constrained optimization problem. The number of parameters is an order of magnitude higher than the problem in Section 4, and proved too computationally intensive to solve in the Indian test case (where interval censoring is severe). We therefore present a sequential approach that yields theoretical bounds on the double-censored CEF for the case of intergenerational mobility.

Specifically, we define the theoretical best- and worst-case latent distributions of  $y$  variables for a given intergenerational mobility statistic. The best- and worst-case assumptions each generate a bound on the feasible value of  $y$  for each  $x$  bin. We then use the method in Section 4 to calculate bounds on the mobility statistic under each case. The union of these bounds is a conservative bound on the mobility statistic given censoring in both the  $y$  and  $x$  variables.

Finally, we can shed light on the distribution of the true value of  $y$  in each  $x$  bin if other data is available. In the context of intergenerational mobility, and in our specific empirical context, it is frequently the case that more information is available about children than about their parents. We use data on child wages to predict whether the true latent child rank distribution ( $y$ ) is better represented by the best- or worst-case mobility scenario. The joint wage distribution suggests that the true latent distribution of  $y$  in each bin is very close to the best case distribution, which we used in Section ??, because there is little effect of parent education on child wages after conditioning on child education.

## C.1 Solution Definition for CEF Bounds with Double Censoring

We are interested in bounding a function  $E(y|x)$ , where  $y$  is known only to lie in one of  $H$  bins defined by intervals of the form  $[y_h, y_{h+1}]$ , and  $x$  is known only to lie in one of  $K$  bins defined by intervals of the form  $[x_k, x_{k+1}]$ . For simplicity, we focus on the case where both  $y$  and  $x$  are uniformly distributed on the interval  $[0, 100]$ .<sup>55</sup>

Where Section 4 focused on bounding the *cumulative expectation function* (CEF) of  $y$  given  $x$ , we focus here on bounding a separate *conditional distribution function* (CDF) for  $y$ , given each value of  $x$ . Each value of  $x$  implies a different CDF for  $y$ , as follows:

$$F(r, x) = P(y \leq r | x = X) \quad (\text{C.1})$$

This CDF is related to the CEF  $E(y|x)$  as follows:

$$E(y|x) = \int_0^{100} r f(x, r) dr \quad (\text{C.2})$$

where  $f(x, r)$  is the probability density function corresponding to the CDF in Equation C.1, when the conditioning variable takes the value  $x$ . Note that  $r$  in this case represents a child rank. This expression simply denotes that  $E(y|x)$  is the average value from 0 to 100 on the  $y$ -axis, holding  $x$  fixed.

We do not observe the sample analog of  $F(x, r)$  directly. Rather, we observe the sample analog of the following expression for each of  $H * K$  bin combinations:

$$P(y \leq y_{h+1} \mid x \in [x_k, x_{k+1}]) = \frac{1}{x_{k+1} - x_k} \int_{q=x_k}^{x_{k+1}} F(q, y_{h+1}) dq \quad (\text{C.3})$$

We denote this sample analog  $\hat{P}(y \leq y_{h+1} \mid x \in [x_k, x_{k+1}])$  as  $\hat{R}(k, h)$ . Equation C.3 states that the probability that  $y$  is less than  $y_{h+1}$  is the average value of the CDF in that bin. Since  $x$  is uniform, we can write its probability distribution function within the bin as  $\frac{1}{x_{k+1} - x_k}$ .

We parameterize each CDF as  $F(x, r) = S(x, r, \gamma_x)$ , where  $r$  is the outcome variable,  $x$  is the

---

<sup>55</sup>Taking a different known distribution into account would require imposing different weights on the mean-squared error function and budget constraint below, but would otherwise not be substantively different.



conditioning variable, and  $\gamma_x$  is a parameter vector in some parameter space  $G_x$ . Similarly let  $f(x,r)=s(x,r,\gamma_x)$ . In our numerical calculation, we define  $G_x$  as  $[0,1]^{100}$ , a vector which gives the value of the cumulative distribution function at each of 100 conditioning variable percentiles on the  $y$ -axis, for a given value of  $x$ . Put otherwise, holding  $x$  fixed, we seek the 100-valued column vector  $\gamma_x$  which contains the value of the CDF at each of the 100 possible  $y$  values:  $y=1,y=2,...,y=100$ . As a result,  $\gamma_x$  must lie within  $[0,1]^{100}$ . Note that there are as many vectors  $\gamma_x$  as there are possible values for the conditioning variable  $x$ . If we discretize also  $x$  as  $1,2,...,100$ , then we define the matrix of 100 CDFs, indexed by  $x$ , as  $\boldsymbol{\gamma}^{100}=[\gamma_1 \ \gamma_2 \ ... \gamma_{100}]$ . To be explicit,  $\boldsymbol{\gamma}^{100}$  is a  $100 \times 100$  matrix constructed by setting its  $x^{\text{th}}$  column as  $\gamma_x$ . We write that  $\boldsymbol{\gamma}^{100} \in G^{100}$ .

We also introduce a new monotonicity condition for this context. In this set up, monotonicity implies that the outcome distribution for any value of  $x$  first-order stochastically dominates the outcome distribution at any lower value of  $x$ . Put otherwise,

$$s(x,r,g_x) \text{ is weakly decreasing in } x \quad (\text{Monotonicity})$$

In the mobility context, this statement implies that the child rank distribution of a higher-ranked parent stochastically dominates the child rank distribution of a lower-ranked parent.<sup>56,57</sup>

The following minimization problem defines the set of feasible values of  $\gamma_x$  for each value  $x$ :

---

<sup>56</sup>? find that a similar conditional monotonicity holds in almost all mobility tables in 35 countries.

<sup>57</sup>A stronger monotonicity assumption would require that the hazard function is decreasing in  $x$ . This is equivalent to stating that the CDF must be weakly decreasing in  $x$  conditional upon  $x$  being above some value. In the mobility case, for example, the stronger assumption would imply that conditional on being in high school, a child of a better off parent must have a higher latent rank than the child of a worse off parent.

$$\Gamma = \underset{\mathbf{g} \in \mathbf{G}^{100}}{\operatorname{argmin}} \left\{ \sum_{k=1}^K \sum_{h=1}^H \left( \int_{q=x_k}^{x_{k+1}} S(q, h, g_q) dq - \hat{R}(k, h) \right)^2 \right\} \quad (\text{C.4})$$

such that

$$s(x, r, g_x) \text{ is weakly decreasing in } x \quad (\text{Monotonicity})$$

$$\frac{1}{100} \sum_{x=1}^{100} S(x, r, g_x) = r \quad (\text{Budget Constraint})$$

$$S(x, 0, g_x) = 0 \quad (\text{End Points})$$

$$S(x, 100, g_x) = 1.$$

In the above minimization problem,  $\mathbf{g}$  is a candidate vector satisfying the conditions; each  $g_x$  describes the candidate CDF holding  $x$  fixed. A valid set of cumulative distribution functions is one that minimizes error with respect to all of the observed data points and obeys the monotonicity condition. The budget constraint requires that the weighted sums of CDFs across all conditioning groups must add up to the population CDF. For example,  $J\%$  of children must on average attain less than or equal to the  $J^{th}$  percentile. The constraints on the end points of the CDF are redundant given the other constraints, but are included to highlight how the end points constrain the set of possible outcomes. For simplicity, we have not included a curvature constraint, but such a constraint would be a sensible further restriction on the feasible parameter space in many contexts.

Once a set of candidate CDFs have been identified, they have a one-to-one correspondence with the CEF given an interval censored conditioning variable, (described by Equation C.2), and thus with any function of the CEF. These statistics can be numerically bounded as in Section 4.

This problem is computationally more challenging than the problem of censoring only in the conditioning variable dealt with in Section 4. In the case of the rank distribution, if we discretize both outcome and conditioning variables into 100 separate percentile bins, then the problem has 10,000 parameters and 10,000 constraints, and an additional 9800 curvature constraint inequalities if desired. This problem proved computationally too difficult to resolve. Restricting the set of discrete

bins (e.g. to deciles) is unsatisfying because it requires significant rounding of the raw data which could substantively affect results. We proceed instead by taking advantage of characteristics that are specific to the problem of intergenerational mobility.

## C.2 Best and Worst Case Mobility Distributions

Our goal is to bound the parent-child rank CEF given interval censored data on both parent and child ranks. In this section, we take a sequential approach to the double-censoring problem. We use additional information about the structure of the mobility problem to obtain worst- and best-case parent CDFs for intergenerational mobility. From these cumulative distribution functions, we can obtain worst- and best-case CEFs using Equation C.2. First, we calculate bounds on the average value of the child rank in each child rank \* parent rank cell. We then apply the methods from Section 4 on the best and worst case bounds; the union of resulting bounds describes the bounds on the mobility statistic of interest. We focus on the rank-rank gradient and on  $\mu_0^{50}$ .

Given data where child rank is known only to lie in one of  $h$  bins, there are two hypothetical scenarios that describe the best and worst cases of intergenerational mobility. Mobility will be lowest if child outcomes are sorted perfectly according to parent outcomes *within* each child bin, and highest if there is no additional sorting within bins.<sup>58</sup>

Consider a simple 2x2 case. In the 1960s birth cohort in India, 27% of boys attained less than two years of education, the lowest recorded category. 55% of these had fathers with less than two years of education, and 45% had fathers with two or more years of education. We do not observe how the children of each parent group are distributed *within* the bottom 27%. For this case, mobility will be lowest if children of the least educated parents occupy the bottom ranks of this bottom bin, or ranks 0 through 15, and children of more educated parents occupy ranks 16 through 27. Mobility will be highest if parental education has no relationship with rank, conditional on the child rank bin. We do not consider the case of perfectly reversed sorting, where the children of the least educated parents occupy the highest ranks within each child rank bin, as it would violate the stochastic dominance condition (and is implausible).

---

<sup>58</sup>Specifically, these scenarios respectively minimize and maximize both the rank-rank gradient and  $\mu_0^x$  for any value of  $x$ . To minimize and maximize  $p_x$ , a different within-bin arrangement is required for every  $x$ . We leave this out for the sake of brevity, and because bounds on  $p_x$  are minimally informative even with uncensored  $y$ .

Appendix Figure C1 shows two set of CDFs that correspond to these two scenarios for the 1960–69 birth cohort. In Panel A, children’s ranks are perfectly sorted according to parent education within bins. Each line shows the CDF of child rank, given some father education. The points on the graph correspond to the observations in the data—the value of each CDF is known at each of this points. Children below the 27th percentile are in the lowest observed education bin. Within this bin, the CDF for children with the least educated parents is concave, and the CDF for children with the most educated parents is convex—indicating that children from the best off families have the highest ranks within this bin. This pattern is repeated within each child bin. Panel B presents the high mobility scenario, where children’s outcomes are uniformly distributed within child education bins, and are independent of parent education within child bin.

According to Equation C.2, each of these CDFs corresponds to a single mean child outcome in a given parent bin, or  $Y = E(y|x \in [x_k, x_{k+1}])$ . From these expected values, we can then calculate bounds on any mobility statistic, as in Sections 4 and ?? . Table C1 shows the expected child rank by parent education for the high and low mobility scenario, as well as bounds on the rank-rank gradient and on  $\mu_0^{50}$ . Taking censoring in the child distribution into account widens the bounds on all parameters. The effect is proportionally the greatest on the interval mean measure, because it was so precisely estimated before—the bounds on  $\mu_0^{50}$  approximately double in width when censoring of son data is taken into account.

These bounds are very conservative, as the worst case scenario is unlikely to reflect the true uncensored joint parent-child rank distribution, due to the number and sharpness of kinks in the CDFs in Panel A of Figure C1. A curvature constraint on the CDF would move the set of feasible solutions closer to the high mobility scenario. We next draw on additional data on children, which suggests that the best case mobility scenario is close to the true joint distribution.

### C.3 Estimating the Child Distribution Within Censored Bins

Because we have additional data on children, we can estimate the shape of the child CDF within parent-child education bins using rank data from other outcome variables that are not censored. Under the assumption that latent education rank is correlated with other measures of socioeconomic rank, this exer-

cise sheds light on whether Panel A or Panel B in Figure C1 better describes the true latent distribution.

Figure C2 shows the result of this exercise using wage data from men in the 1960s birth cohort. To generate this figure, we calculate children's ranks first according to education, and then according to wage ranks within each education bin.<sup>59</sup> The solid lines depict this uncensored rank distribution for each father education; the dashed gray lines overlay the estimates from the high mobility scenario in Panel B of Figure C1.

If parent education strongly predicted child wages *within* each child education bin, we would see a graph like Panel A of Figure C1. The data clearly reject this hypothesis. There is some additional curvature in the expected direction in some bins, particularly among the small set of college-educated children, but the distribution of child cumulative distribution functions is strikingly close to the high mobility scenario, where father education has little predictive power over child outcomes after child education is taken into account. The last row of Table C1 shows mobility estimates using the within-bin parent-child distributions that are predicted by child wages; the mobility estimates are nearly identical to the high mobility scenario. This result supports the assumption made in Section ?? that latent child rank within a child rank bin is uncorrelated with parent rank.

Note that there is no comparable exercise that we can conduct to improve upon the situation when parent ranks are interval censored, because we have no information on parents other than their education, as is common in mobility studies. If we had additional information on parents, we could conduct a similar exercise. The closest we can come to this is by observing the parent-child rank distribution in countries with more granular parent ranks, as we did in Section 4. The results in that section suggest that interval censoring of parent ranks does indeed mask important features of the mobility distribution.

An additional factor that makes censoring in the child distribution a smaller concern is the fact that children are more educated than parents in every cohort, and thus the size of the lowest education bin is smaller for children than for parents. This result is likely to be true in many other countries where education is rising. Of course, in other contexts, we may lack additional information

---

<sup>59</sup>We limit the sample to the 50% of men who report wages. Results are similar if we use household income, which is available for all men. Household income has few missing observations, but in the many households where fathers are coresident with their sons, it is impossible to isolate the son's contribution to household income from the father's, which biases mobility estimates downward.

about the distribution of the  $x$  and  $y$  variable within bins, and researchers may prefer to work with conservative bounds as described in C.2.

**Figure C1**  
Best- and Worst-Case Son CDFs  
by Father Education (1960-69 Birth Cohort)

Panel A: Lowest Feasible Mobility

Panel B: Highest Feasible Mobility

---

Figure C1 shows bounds on the CDF of child education rank, separately for each father education group. The lines index father types. Each point on a line shows the probability that a child of a given father type obtains an education rank less than or equal to the value on the X axis in the national education distribution. The large markers show the points observed in the data.

**Figure C2**  
Son Outcome Rank CDF  
by Father Education (1960-69 Birth Cohort)  
Joint Education/Wage Estimates

---

Figure C2 plots separate son rank CDFs separately for each father education group, for sons born in the 1960s in India. Sons are ranked first in terms of education, and then in terms of wages. Sons not reporting wages are dropped. For each father type, the graph shows a child's probability of attaining less than or equal to the rank given on the X axis.

**Table C1**  
Mobility Estimates under Double-Censored CEF

	Upward Interval Mobility ( $\mu_0^{50}$ )	Rank-Rank Gradient ( $\beta$ )
Low mobility scenario	[32.33, 35.90]	[0.55, 0.80]
High mobility scenario	[35.86, 38.80]	[0.45, 0.67]
Wage imputation scenario	[35.79, 38.70]	[0.46, 0.67]

Table C1 presents bounds on  $\mu_0^{50}$  and the rank-rank gradient  $\beta$  under three different sets of assumptions about child rank distribution within child rank bins. The low mobility scenario assumes children are ranked by parent education within child bins. The high mobility scenario assumes parent rank does not affect child rank after conditioning on child education bin. The wage imputation predicts the within-bin child rank distribution using child wage ranks and parent education.

## D Appendix D: Data Construction