Breaking the Caste Barrier: Intergenerational Mobility in India*

Viktoria Hnatkovska†, Amartya Lahiri†, and Sourabh B. Paul†

May 2011

Abstract

Amongst the various inequities typically associated with the caste system in India, probably one of the most debilitating is the perception that one is doomed by birth, i.e., social and economic mobility across generations is difficult. We study the extent and evolution of this lack of mobility by contrasting the intergenerational mobility rates of the historically disadvantaged scheduled castes and tribes (SC/ST) in India with the rest of the workforce in terms of their education attainment, occupation choices and wages. Using household survey data from successive rounds of the National Sample Survey between 1983 and 2005, we find that inter-generational education and income mobility rates of SC/STs have converged to non-SC/ST levels during this period. Moreover, SC/STs have been switching occupations relative to their parents at increasing rates, matching the corresponding switch rates of non-SC/STs in the process. Interestingly, we have found that a common feature for both SC/STs and non-SC/STs is that the sharpest changes in intergenerational income mobility has been for middle income households. This is consistent with the effects of easing credit constraints, a phenomenon that did characterize this period. We conclude that the last twenty years of major structural changes in India have also coincided with a breaking down of caste-based historical barriers to socio-economic mobility.

JEL Classification: J6, R2

Keywords: Intergenerational mobility, wage gaps, castes

*We would like to thank Siwan Anderson, Nicole Fortin, Ashok Kotwal, Thomas Lemieux, Anand Swamy and seminar participants at Delhi School of Economics, Monash University, UBC, the LAEF Conference on "Growth and Development" at UC Santa Barbara and the ISI Growth conference for comments. Special thanks to David Green for extremely helpful and detailed suggestions.

†Department of Economics, University of British Columbia, 997 - 1873 East Mall, Vancouver, BC V6T 1Z1, Canada. E-mail addresses: hnatkovs@interchange.ubc.ca (Hnatkovska), alahiri@interchange.ubc.ca (Lahiri), sbikas@gmail.com (Paul).

1 Introduction

One of oldest and most enduring social arrangements in India dating back thousands of years is the caste system. The system is an offshoot of a method of organizing society into ordered classes such as priests, warriors, traders, workers etc.. A key characteristic of this system is that caste status is inherited (by birth). Given the traditional assignment of jobs/tasks by castes, the social restrictions imposed by the hereditary nature of the system have been viewed as probably the biggest impediment to social mobility for the poor and downtrodden. The traditional narrative – which finds resonance amongst politicians, academics and social activists in India to this day – holds that the son of a poor, uneducated cobbler is likely to also end up as a poor, uneducated cobbler because, independent of his relative skill attributes, it is very hard for the son of a cobbler to find employment in other occupations. Hence, the desire to get educated for such a person is also limited since a large part of the attraction of acquiring education is its value in getting jobs.

This concern was the primary motivation behind the founding fathers of the Indian constitution extending affirmative action protection to the lowest castes in the caste hierarchy via the constitution itself. Specifically, the most disadvantaged castes and tribes were provided with reserved seats in higher educational institutions, in public sector jobs and in state legislatures as well as the Indian parliament. The protected groups were identified in a separate schedule of the constitution and hence called Scheduled Castes and Scheduled Tribes or SC/STs. The reservations were intended as a temporary measure that were expected to help level the playing field for the disadvantaged SC/STs over a few generations.

It has now been over 60 years since the constitution of India came into effect in 1950. Moreover, over the past 25 years India has also experienced rapid and dramatic macroeconomic changes with a sharp rise in aggregate growth, massive structural transformation of the economy, increasing urbanization, etc.. How have the historically disadvantaged castes and tribes – the SC/STs – performed during this period? Has social mobility increased over time or has it stayed relatively unchanged? How does social mobility in India compare with mobility in modern industrialized economies? In this paper we attempt to answer some of these questions.

We use data from five successive rounds of the National Sample Survey (NSS) of India from 1983 to 2004-05 to analyze patterns of education attainment, occupation and industry choices, and wages of both SC/ST and non-SC/ST households. In particular, we contrast the time series behavior of the intergenerational persistence of education, occupation, industry of employment and

wage levels of SC/ST and non-SC/ST households.

We find that intergenerational mobility of SC/STs was lower than that of non-SC/STs at the beginning of our sample in 1983, but has risen faster than that of non-SC/ST households in both education attainment rates and wages. The probability of an SC/ST child changing his level of education attainment relative to the parent was just 42 percent in 1983 but rose sharply to 67 percent by 2004-05. The corresponding probabilities of a change in education attainment for a non-SC/ST child were 57 percent and 67 percent. Hence, there has been a clear convergence of intergenerational *education mobility* rates between SC/STs and non-SC/STs. Correspondingly, the elasticity of wages of children with respect to the wages of their parent has declined from 88 percent to 45 percent for SC/ST households and from 76 to 58 percent for non-SC/ST households, indicating a clear trend towards convergence in intergenerational *income mobility* rates.

Our study finds that intergenerational *occupational and industry mobility* rates have increased for both groups during this period. However, these changes in occupational and industry mobility rates have been relatively similar across the two groups. As a result, children in non-SC/ST households continue to be more likely to work in a different occupation and/or different industry than their parent relative to children from SC/ST households.

A key question of interest to us is whether the gains made by SC/STs during this period were restricted to the relatively well-off sections of SC/STs – the so called "creamy layer"? We study this issue by examining the mobility at different points of the education, occupation, industry and wage distributions. In terms of education attainment, we find that the largest changes for SC/STs were in movements out of illiteracy into middle and primary schools. Similarly, there were significant intergenerational movements from agricultural occupations into blue-collar occupations for both SC/STs and non-SC/STs.

In terms of income mobility, we use the recent approaches of Jäntti et al. (2006) and Bhattacharya and Mazumder (2011) to compute non-parametric measures such as income transition matrices and upward mobility measures. We find an increase in intergenerational income mobility in India and convergence of mobility rates between the SC/STs and non-SC/STs for most income groups. Moreover, the probability of a child improving his rank in his generation's income distribution relative to his father's corresponding rank is higher for SC/STs compared to non-SC/ST households. For both SC/STs and non-SC/STs, on average, the sharpest increases in intergenerational income mobility was in the middle of the income distribution – households in quintiles 2 and 3 of the wage distribution. This result is consistent with the effects of relaxing credit constraints

in the economy and their effects on human capital investment.

Our results indicate that the gains during the past two decades have not been restricted to limited sections of SC/STs. Education mobility has occurred for both low and relatively highly educated SC/ST households. Similarly, income mobility has increased for both low and high-income households amongst SC/STs. Moreover, the increase in mobility for SC/STs has, on average, been faster than for non-SC/STs. Indeed, it has now become far more likely that the son of a poor illiterate SC/ST cobbler would become a machine worker with middle or secondary school education having a much higher rank in his generation's income distribution than his father did in his generation.

In summary, our results suggest that neither the lack of occupational mobility nor the lack of education have been a major impediment toward the SC/STs taking advantage of the rapid structural changes in India during this period to better their economic position.

While there has been considerable work on intergenerational mobility in the US and other western countries (see Becker and Tomes (1986), Behrman and Taubman (1985), Haider and Solon (2006) amongst others), this issue has received remarkably little attention in the work on India. The two notable exceptions are Jalan and Murgai (2009) and Maitra and Sharma (2009) both of which focus on intergenerational mobility in education attainment. The biggest difference between our work and these studies is that we examine intergenerational mobility patterns not just in education attainment but also in occupation choices, industry of employment, and income. We are not aware of any other study that documents intergenerational mobility patterns in education, occupation, industry, and income together. Our work also differs from Jalan and Murgai (2009) and Maitra and Sharma (2009) in two other respects: (a) we use a much larger sample of households due to our use of the NSS data; and (b) by examining multiple rounds of the NSS data we are also able to study the time-series evolution of intergenerational mobility patterns in India.[1]

In the next section we describe the data and our constructed mobility measures as well as some summary statistics. Section 3 presents and discusses the evidence on intergenerational mobility, while the last section concludes.

[1] In related work Munshi and Rosenzweig (2009) document the lack of labor mobility in India. Also, Munshi and Rosenzweig (2006) show how caste-based network effects affect education choices by gender.

2 The Data

Our data comes from the National Sample Survey (NSS) of India Rounds 38 (1983), 43 (1987-88), 50 (1993-94), 55 (1999-2000) and 61 (2004-05). The survey covers the whole country. The number of households surveyed averaged about 121,000 households across the rounds. Our working sample consists of all male households heads and their male children/grandchildren between the ages 16 and 65 who provided their 3-digit occupation code information and their education information.[2] Our focus is on full-time working individuals who are defined as those that worked at least 2.5 days per week, and who are not currently enrolled in any education institution. We conduct all our data work using a sample in which the criteria above are satisfied for both household's head and at least one child or grandchild in that household. This selection leaves us with a sample of about 21,000 households comprising around 43,000-51,000 individuals, depending on the survey round. We refer to this sample as "working" sample.[3]

Our dataset does not contain information on individual's years of schooling. Instead, the education variable is coded into detailed categories ranging from not-literate to postgraduate and above. We aggregate these categories into 5 broader groups: not-literate; literate but below primary; primary education; middle education; and secondary and above education (which includes higher secondary, diploma/certificate course, graduate and above in different professional fields, postgraduate and above). These categories are coded as education categories 1, 2, 3, 4 and 5 respectively. Our dataset also contains information about the three-digit occupation code (based on the 1968 National Classification of Occupation (NCO)) associated with the work that each individual performed over the last year preceding the survey year. Lastly, we also have information on the four-digit industry of employment for each individual. Across rounds we convert industry codes into a uniform classification of 1998 National Industrial Classification (NIC) of the Government of India.

Data on wages are more limited. The sub-sample with complete wage data for both the head of household and at least one child or grandchild in the same household consists of, on average across rounds, about 7,000-9,000 individuals which is considerably smaller than our working sample

[2] We also consider a broader sample in which we do not restrict the gender of the children and find that our results remain robust (in fact, majority of the children working full-time in our sample are male). We choose the restriction to only males for two reasons. First, female led households are few and usually special in that those households are likely to have undergone some special circumstances. Second, since there are a number of societal issues surrounding the female labor force participation decision which can vary both across states and between rural and urban areas, focusing only on males allows us to avoid having to deal with these complications.

[3] Note that the number of individuals included from each household is typically much smaller than the total members of the household due to the restrictions on age, sex, generations etc.

but large enough to facilitate formal analysis. Our wage series is the daily wage/salaried income received for the work done during the week previous to the survey week. We evaluate in-kind wages using current retail prices. Wages are converted into real terms using state-level poverty lines differentiated by rural and urban sectors. All wages are expressed in 1983 rural Maharashtra prices. Details regarding the dataset are contained in the Appendix A.

In order to conduct the intergenerational comparisons, we collect all household heads into a group called "parents" and the children/grandchildren into the group "children". This sorting is done for each survey round and the statistics are computed for each generation for that round. Naturally, this sorting method implies that some individuals who were children in an earlier round could become parents in a subsequent round if there was substantial change in their family dynamics during the intervening period.

Table 1 gives some summary statistics of the data. Panel (a) reports average age, education level, share of males and married individuals among children; while panel (b) reports the corresponding statistics for household heads (parents). Panel (b) also reports the percentage of rural households in our sample, as well as the average household size. Note that "All" refers to the full working sample, while the "Non-SC/ST" and "SC/ST" panels refer to the corresponding caste sub-samples**.**

Household-heads are around 52 years of age while their male working children are typically around 23 years old. Around 81 percent of surveyed households are rural and engaged in farming/pastoral activities. This number is slightly higher for SC/ST households, 88-89 percent of whom live in rural areas on average. Of the working children living with the household-head, 49 percent are married on average. While the percent of married children has declined over time, this change was characteristic of both non-SC/ST and SC/ST children. Finally, the average education level of children is greater than that of parents for both SC/STs and non-SC/STs, and has increased over time. Non-SC/STs are also consistently more educated than SC/ST. The proportion of SC/ST households in the sample across the different rounds is around 24 percent.

2.1 Sample Issues

Before proceeding it is important to discuss some key issues regarding our sample. The ideal sample for addressing intergenerational mobility issues is one that has information on education, occupation and wages for parents as well as their adult children. Another desirable feature of such a sample is that it has wage information for parents and adult children at comparable ages rather than at different phases of their lifecycles. The NSS data unfortunately has some limitations in this regard.

Table 1: Sample summary statistics

	(a) children			(b) parents				
All	age	edu	married	age	edu	married	rural	hh size
1983	22.83	2.58	0.53	51.67	1.79	0.92	0.81	7.18
	(0.04)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.02)
1987-88	23.13	2.69	0.53	51.65	1.88	0.92	0.83	6.98
	(0.04)	(0.01)	(0.00)	(0.06)	(0.01)	(0.00)	(0.00)	(0.02)
1993-94	23.17	2.97	0.48	51.78	2.01	0.94	0.82	6.51
	(0.04)	(0.01)	(0.00)	(0.06)	(0.01)	(0.00)	(0.00)	(0.02)
1999-00	23.43	3.21	0.46	51.60	2.20	0.94	0.81	6.56
	(0.05)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.02)
2004-05	23.38	3.40	0.46	51.57	2.34	0.94	0.80	6.39
	(0.05)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.02)
Non-SC/ST								
1983	23.00	2.78	0.52	52.04	1.93	0.92	0.79	7.29
	(0.05)	(0.01)	(0.00)	(0.08)	(0.01)	(0.00)	(0.00)	(0.03)
1987-88	23.30	2.89	0.51	51.98	2.03	0.93	0.80	7.06
	(0.05)	(0.01)	(0.00)	(0.08)	(0.01)	(0.00)	(0.00)	(0.02)
1993-94	23.36	3.17	0.47	52.10	2.19	0.94	0.79	6.6
	(0.05)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.02)
1999-00	23.76	3.42	0.47	52.01	2.41	0.95	0.78	6.62
	(0.05)	(0.01)	(0.00)	(0.08)	(0.02)	(0.00)	(0.00)	(0.03)
2004-05	24.04	3.56	0.46	52.01	2.52	0.95	0.77	6.42
	(0.06)	(0.01)	(0.01)	(0.08)	(0.02)	(0.00)	(0.00)	(0.03)
SC/ST								
1983	22.30	1.95	0.56	50.59	1.38	0.92	0.89	6.86
	(0.08)	(0.02)	(0.01)	(0.13)	(0.01)	(0.01)	(0.01)	(0.04)
1987-88	22.63	2.06	0.56	50.72	1.45	0.91	0.90	6.76
	(0.08)	(0.02)	(0.01)	(0.12)	(0.01)	(0.00)	(0.00)	(0.04)
1993-94	22.61	2.40	0.49	50.92	1.54	0.92	0.90	6.25
	(0.08)	(0.02)	(0.01)	(0.13)	(0.02)	(0.00)	(0.00)	(0.04)
1999-00	22.85	2.67	0.46	50.61	1.71	0.94	0.88	6.41
	(0.09)	(0.02)	(0.01)	(0.13)	(0.02)	(0.00)	(0.01)	(0.04)
2004-05	23.05	2.99	0.45	50.66	1.87	0.94	0.87	6.3
	(0.09)	(0.03)	(0.01)	(0.14)	(0.02)	(0.00)	(0.01)	(0.05)

Notes: This table reports summary statistics for our sample. Panel (a) gives the statistics for the generational subsample of children, while panel (b) gives the statistics for the household heads (parents). Standard errors are reported in parenthesis.

First, it provides information on parents and their adult children only if the two generations are co-resident in the same household. This immediately raises selection issues as co-resident households may be special and have characteristics that differ systematically from other households. Second, the NSS does not track the same household over time. Hence, for every parent-child pair, we have observations at a point in time which makes wage comparisons between the generations potentially problematic.

How special is our sample? We begin by documenting the incidence of co-resident households in the NSS data. We define co-residence as having multiple adult (16 years of age and above) generations living in the same household: i.e., parents/parents-in-law living with their adult children and/or grandchildren. We find that co-resident households are the norm in India. In contrast to patterns in more industrial and western economies, in India (as well as in a number of other

developing countries), a majority of households tend to co-reside. Thus, in the NSS sample across the rounds, on average, about 62 percent of all sampled households were characterized by multiple adult generations co-residing. Importantly, this fraction of co-resident households has also remained quite stable across the rounds. This stable trend is in contrast to the conventional view that the nuclear family is becoming more and more the norm in India as the economy is growing and modernizing. Joint households are even more prevalent in rural areas where the majority of India still resides. Hence, in the Indian context, drawing inferences from samples reflecting predominantly nuclear households is arguably more problematic due to its unrepresentative nature.

Unfortunately, we cannot directly use the co-resident sample described above because the NSS identification code lumps parents and parents-in-laws together in one category making it problematic for computation of direct intergenerational trends. Hence, we choose to focus instead on households with an adult head of household co-residing with at least one adult of lower generation (child and/or grandchild), both being in the age-group 16-65. This sub-sample of households comprises about 75 percent of co-resident households. Imposing the additional restrictions on sex, education, occupation information and full-time employment status (as we did in our working sample) gives us a sample which covers about 24 percent of the full dataset with the same restrictions. Crucially, this ratio too is stable across the rounds. We refer to this full dataset as the unrestricted sample and contrast the characteristics of the co-resident households with the households from the unrestricted sample. Table 2 reports the results.

Panel (a) of Table 2 reports the household characteristics in our working sample of co-resident households, while panel (b) does the same for the households in the unrestricted sample (no co-residence requirement). The household age column (hh age) reports the average age of all household members. Columns # adults, # kids, # earning mem refer to the number of adult household members (defined as those aged 16 and above), number of kids (below 16 years in age), and the number of earning members in the household (defined as those who reported their employment status as employed during the survey). Column labelled "rural" refers to the share of rural households, and column "# households" reports the number of household in the sample.

As is to be expected, our households are, on average, slightly older, have more adults and earning members, and are more likely to be from rural areas than those in the unrestricted sample. Importantly, however, these differences are small and stable over time. Furthermore, the greater representation of rural households in our sample indicates to us the importance of incorporating

Table 2: Sample comparisons

(a) working sample

round	hh age	# adults	# kids	# earning mem	rural	# households
1983	25.00	4.99	3.15	3.46	0.81	19225
	(0.05)	(0.02)	(0.02)	(0.02)	(0.00)	
1987-88	25.48	5.00	2.96	3.32	0.83	21977
	(0.05)	(0.02)	(0.02)	(0.01)	(0.00)	
1993-94	26.68	4.91	2.48	3.44	0.82	19870
	(0.05)	(0.02)	(0.02)	(0.01)	(0.00)	
1999-00	26.97	5.06	2.52	3.46	0.81	19997
	(0.06)	(0.02)	(0.02)	(0.02)	(0.00)	
2004-05	27.39	5.03	2.35	3.48	0.81	21560
	(0.06)	(0.02)	(0.03)	(0.02)	(0.00)	

(b) unrestricted sample

round	hh age	# adults	# kids	# earning mem	rural	# households
1983	23.12	3.56	2.98	2.29	0.76	87873
	(0.03)	(0.01)	(0.01)	(0.01)	(0.00)	
1987-88	23.39	3.53	2.81	2.18	0.77	94676
	(0.03)	(0.01)	(0.01)	(0.01)	(0.00)	
1993-94	24.16	3.45	2.53	2.24	0.75	87099
	(0.03)	(0.01)	(0.01)	(0.01)	(0.00)	
1999-00	24.46	3.55	2.55	2.22	0.74	87768
	(0.04)	(0.01)	(0.01)	(0.01)	(0.00)	
2004-05	25.38	3.57	2.38	2.29	0.75	87102
	(0.04)	(0.01)	(0.01)	(0.01)	(0.00)	

Note: The overall sample is derived by imposing the same restrictions on sex, age, education, occupation and full-time employment status that were imposed in deriving the working sample. The key difference between the working and unrestricted samples is that the latter does not impose the co-residence requirement. Standard errors are in parenthesis.

controls for rural effects in our empirical analysis below.[4] In summary, we view Table 2 as being indicative of the fact that our sub-sample is a stable representation of the households sampled by the NSS. More generally, the facts above suggest to us that co-residence patterns have not changed significantly during the period under study. Hence the representativeness of the sample under this identification have remained comparable across rounds.

We conduct a further check of the representativeness of our sample by comparing the characteristics of the parents and children generations in our working sample with the counterparts of these generations in the unrestricted sample, separately. This comparison necessarily involves making some assumptions in order to construct the generational counterparts in the unrestricted sample. For the counterpart to the parents generation, we consider the household heads of all households in the unrestricted sample subject to them meeting the age, sex, education, occupation, and full time working status requirement that we imposed on our working sample. Hence, we are essentially comparing the characteristics of household heads in the unrestricted sample with the characteris-

[4] We also examined the daily average real per capita consumption expenditures of the two sets of household and found that those differences too were small, stable and insignificant across the rounds. These results are available from the authors upon request.

tics of household heads in co-resident households. We construct the children's generation in the unrestricted sample by including all non-household head adults whose ages are in a band of plus or minus one standard deviation of the mean age of the children in our working sample.

We report the characteristics of the constructed parents and children generations in the unrestricted sample in Table 3. To contrast their characteritics with those of parents and children generations in the co-resident households we refer the reader to panels (a) and (b) in Table 1. The children in our working sample are quite similar to the children in the unrestricted sample on most characteristics in all the rounds. The parents in our working sample are older, less educated and more rural than those in the unrestricted sample. However, crucially for our goal of determining time trends in mobility patterns, the differences between the two samples are stable over time. Overall, we find no evidence of any time trends in the differences in the characteristics of the generations in the two samples. Hence, our conclusions regarding the time trends in intergenerational mobility patterns remain valid despite the limitations of the dataset.

Table 3: Characteristics of children and parents in the unrestricred sample

	(a) children (unrestricted sample)				(b) parents (unrestricted sample)			
	age	edu	married	rural	age	edu	married	rural
1983	22.38	2.62	0.50	0.77	35.55	2.37	0.79	0.75
	(0.02)	(0.01)	(0.00)	(0.00)	(0.05)	(0.01)	(0.00)	(0.00)
1987-88	22.43	2.71	0.49	0.79	35.81	2.44	0.80	0.77
	(0.02)	(0.01)	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)	(0.00)
1993-94	22.40	3.02	0.43	0.77	36.11	2.66	0.80	0.75
	(0.02)	(0.01)	(0.00)	(0.00)	(0.04)	(0.01)	(0.00)	(0.00)
1999-00	22.59	3.28	0.42	0.76	36.36	2.87	0.80	0.73
	(0.03)	(0.01)	(0.00)	(0.00)	(0.05)	(0.01)	(0.00)	(0.00)
2004-05	22.64	3.44	0.39	0.75	36.99	3.02	0.79	0.73
	(0.03)	(0.01)	(0.00)	(0.00)	(0.05)	(0.01)	(0.00)	(0.00)

Note: This table presents summary statistics for children (panel (a)) and parents (panel (b)) generations in the unrestricted sample. The unrestricted generation of parents is obtained as all co-resident and not co-resident household heads that are males within 16-65 age range, and provided their education, occupation information, are employed full-time and are not enrolled in any education institution. The unrestricted generation of children is obtained as all those individuals whose age lies within 1 std dev band around the mean age of the children in our working (co-resident) sample. Standard errors are in parenthesis.

Our focus on co-resident households potentially misses important intergenerational mobility information that is contained in the decision to move out of the parents' household by younger generations. However, this missing information could bias our mobility measures in either direction. On the one hand, more able and educated children may be more likely to move out of their parents' home. In this case, our sample would underestimate the true intergenerational mobility as it does

not include these children. On the other hand, the less educated and wealthy are the parents, the more likely it is that their children may continue to live in the same household in order to take care of them (the intra-household insurance and risk sharing motive). Since these households are included in our sample, we would tend to overestimate the degree of intergenerational mobility. On balance, the net bias could go either way. Importantly, the stability of the share of co-resident households implies that there would not be any time-series trends in the bias. Hence, our estimates of the changes in intergenerational mobility should remain unaffected by this.

The second issue is about when one observes the wage information for parents and their children. NSS reports the data for both generations at the same point in time rather than at the same point in their lifecycles. This is a perennial problem in intergenerational mobility studies. We address this by using the same approaches and instruments that were developed and implemented in the intergenerational mobility literature by Haider and Solon (2006) and Lee and Solon (2009). We discuss them in greater detail in Section 3.4 below.

3 Intergenerational Mobility

We now turn to the key question that we started with: how have the patterns of intergenerational mobility in India changed between 1983 and 2004-05? Our primary interest is in studying how the occupation and industry choices, education attainment levels and wages of children compare with the corresponding levels for their parents. We shall look at each of these in turn.

In the foregoing analysis we shall define the intergenerational education/ occupation/ industry switch as a binary variable that takes a value of one if the child's or grandchild's education level/ occupation/ industry of employment is different from his parent's (who is the head of the household) education achievements/ occupation/ industry of employment; and zero otherwise. We label the education switch variable as **switch-edu;** the occupation switch variable as **switch-occ;** and the industry switch variable as **switch-ind.** We also distinguish *education improvement*, which is another binary variable equal to one if the child's education is higher than that of his parent and zero otherwise, from *education reduction* which is a binary variable that takes a value of one if the child's education is below his parent's education and zero otherwise.

3.1 Education Mobility

We begin by analyzing intergenerational education switches. Our main interest is in determining the degree to which children are changing their education levels relative to their parents and by how much. We are also interested in determining whether or not the switches reflect increases in educational attainment by the children.

To obtain average probabilities of education switches we posit the following probit model:

$$P_i \equiv \Pr(y_i = 1|x_i) = E\left(y_i|x_i\right) = \psi(x_i\beta),$$

where $\psi(x_i\beta) = \Phi(x_i\beta)$, with $\Phi(.)$ representing the cumulative standard normal distribution function, y_i is a binary variable for education switch as defined above (*switch-edu*), and x_i is a vector of controls. We allow the education switch for individual i to depend on his individual characteristics, such as age, age squared, belonging to an SC/ST group (SC/ST), and religion ($muslim$); household-level characteristics, such as household size (hh_size), his rural location ($rural$); and state-level characteristics, such as state-level reservation quota for SC/STs, and region-specific fixed effects. Thus,

$$\begin{aligned} x_i\beta &= \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 SC/ST_i + \beta_4 muslim_i \\ &\quad + \beta_5 rural_i + \beta_6 hh\_size_i + \beta_7 quota\_scst_j + \sum_{j=1}^{5} \alpha_j region\_dummy_j. \end{aligned} \tag{3.1}$$

We include a Muslim dummy in our regression specification to control for the fact that Muslims, on average, have done poorly in modern India. If included in the non-SC/ST group, the poor outcomes of Muslims may bias out results towards finding more convergence between non-SC/STs and SC/STs. We also control for regional differences by including region dummies in the regression specification, where regions are defined as North, South, East, West, Central and North-East.[5]

The introduction of reservations for SC/STs in public sector employment and in higher education institutions was a key policy initiative in India.[6] Due to their potentially important effects

[5] This grouping reflects similarities across states along their geographic characteristics, and characteristics that are shared based on proximity.

[6] The reservations were provided in proportion to the population shares of SCs and STs. State-level reservations can change over time due to changes in SC/ST population shares. In 1991 the Indian government extended the reservation policy to include other backward castes (OBCs). In our analysis we focus only on the group of SC/STs while OBCs are included in the non-SC/ST reference group. If reservations benefited OBCs then our results potentially understate the true degree of convergence between SC/STs and non-SC/STs (excluding OBCs), especially since the extension of reservations to OBCs in 1991.

on the historical inequities against SCs and STs, reservations in India have been studied in several papers. Thus, Pande (2003) examines the effects of reservations on government policies, while Prakash (2009) studies the effects of reservations on the labor market outcomes of SC/STs. Both authors find evidence of positive effects of reservations on the targeted groups. Hence, it is important to control for state level reservation quotas in the analysis. In the regression analysis we include reservation quotas that were in effect when information on household-head and his children and grandchildren was collected. This allows us to control for differences in reservation policies across states.

We estimate the model for each survey round separately and use it to obtain fitted values for each individual. These fitted values are used to compute the average probability of intergenerational education switch. We compute these probabilities for the overall sample as well as for SC/STs and non-SC/STs separately.

It is worth noting that rather than estimating the probability of education switches using a regression specification we could have instead just computed the frequency distribution of education switches between generations. The two approaches yield very similar computed probabilities. We choose to proceed with the regression approach as we are also interested in the effect of caste on the probability of switching education categories across generations *conditional* on other controls. As we show below, the marginal effects of caste on the estimated probabilities are almost always significant.

Panel (a) of Figure 1 depicts the computed probabilities of intergenerational switches in education attainment together with the $\pm$ 2 standard error confidence bands (dashed lines).[7] There are two features of the Figure worth pointing out. First, intergenerational mobility as reflected by the switch probabilities have increased for both SC/STs and non-SC/STs over the sample period. Second, and possibly more remarkably, the switch probabilities of the two groups have converged at 67 percent by the end of our sample period in 2004-05. This is particularly impressive once one notes that in 1983, the probability of an intergenerational education switch for SC/ST households was a meagre 42 percent relative to the 57 percent corresponding probability of non-SC/ST households.

A related question is about the degree or size of the change in education levels. In particular, amongst the children who switch education levels relative to their parent, how large is the change? How has this evolved over our sample period? Panel (b) of Figure 1 reveals that the average size of

[7] Confidence bands around the probability of education switch are very narrow and do not appear on the graph for that reason.

Figure 1: Intergenerational education switches

(a) (b)

Notes: Panel (a) of this figure presents the average predicted probability of intergenerational education switch, while panel (b) reports the average size of the intergenerational education switches for our overall sample, for SC/STs and non-SC/STs. The numbers are reported for the five NSS survey rounds. Dotted lines are ±2 std error bands.

the switch has been increasing over time for both groups. Crucially, by the end of our sample, the switch sizes for the two groups not only converged but SC/STs were in fact switching education levels by more than non-SC/STs. This again is noteworthy since the average size of a switch for SC/STs was significantly lower at 0.6 in 1983 relative to 0.84 for the non-SC/ST households. Note that positive numbers for the size of the switch indicate improvements in education categories.

We also find that most of the increase in the probability of education mobility over our sample period was due to a fall in the negative effect of the caste, conditional on other attributes. Thus, Table 4 reports the marginal effects associated with the SC/ST dummy from the probit regression for education switches defined in equation (3.1).[8] The Table shows that the caste marginal effect was negative and significant for all but the last round. Crucially, the absolute value of that marginal effect has declined secularly over time culminating in it becoming insignificant in 2004-05. Thus, while being an SC/ST used to have a significant negative effect on the probability of a child switching his education category relative to his parent, by the end of our sample period caste had seemingly lost any independent explanatory power for the switch probability. The bottom panel of Table 4 reports the changes in the SC/ST marginal effect during the entire period 1983-2004/05 as well as the two decadal sub-periods 1983-1993/94 and 1993/94-2004/05. All the changes were

[8] Complete estimation results are included in Appendix B.

highly significant.

Table 4: Marginal effect of SC/ST dummy in probit regression for intergenerational education switches

	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
1-SC/ST, 0-non SC/ST	-0.1412***	-0.1394***	-0.1047***	-0.0665***	-0.0155
	(0.0097)	(0.0088)	(0.0089)	(0.0095)	(0.0105)
Changes		1983 to 1993-94	1993-94 to 2004-05	1983 to 2004-05	
		0.0365***	0.0892***	0.1257***	
		(0.0132)	(0.0138)	(0.0143)	

Notes: This table reports the estimated coefficient on the SC/ST dummy from the probit regression (3.1) in which the dependent variable is whether or not there was an intergenerational education switch. Columns (i)-(v) refer to the survey round. Panel "Changes" reports change in SC/ST dummy coefficient over the successive decades and the entire sample period. Standard errors are in parentheses. * p-value≤0.10, ** p-value≤0.05, *** p-value≤0.01.

We also find that most of the intergenerational education switches are in fact increases in educational attainment levels. The estimated probability of an SC/ST child increasing his level of education attainment relative to the parent was just 36 percent in 1983 but rose sharply to 59 percent by 2004-05. The corresponding probabilities of an increase in education attainment for a non-SC/ST child were 49 percent and 58 percent. The probability of an education reduction is around 9 percent for non-SC/STs and 7 percent for SC/STs. Both these probabilities have remained stable over the sample period. Hence, a majority of the increase in the education switch probability for SC/STs relative to the non-SC/STs is accounted for by an increase in the probability of an improvement in the education attainment level.[9]

3.1.1 Education Transition Matrix

While the overall mobility trends in education are informative, they do not reveal the underlying changes at the disaggregated level. A key question of interest to us is whether there are underlying distributional patterns in the intergenerational education mobility trends of the two groups. In particular, is most of the increase in intergenerational education mobility due to children of the least educated parents moving up the education ladder or is it the upward mobility of the children of the relatively highly educated parents that accounts for the aggregate pattern? Are there differences in the patterns between SC/STs and non-SC/STs?

We explore these issues by computing the education transition matrix for our sample of house-

[9] Estimated probabilities and detailed regression results for education increases and decreases separately are available in Appendix B.

Table 5: Intergenerational education transition probabilities

(a). Average mobility in the 1983 round

Non-SC/ST							**SC/ST**						
	Edu**1**	Edu**2**	Edu**3**	Edu**4**	Edu**5**	size		Edu**1**	Edu**2**	Edu**3**	Edu**4**	Edu**5**	size
Edu**1**	0.85	0.09	0.05	0.01	0.00	0.32	Edu**1**	0.91	0.06	0.02	0.01	0.00	0.56
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)
Edu**2**	0.49	0.37	0.11	0.02	0.01	0.12	Edu**2**	0.63	0.28	0.07	0.01	0.01	0.13
	(0.02)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)		(0.02)	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)
Edu**3**	0.46	0.24	0.22	0.06	0.02	0.20	Edu**3**	0.62	0.22	0.12	0.03	0.02	0.15
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)		(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Edu**4**	0.35	0.24	0.21	0.14	0.05	0.20	Edu**4**	0.55	0.18	0.18	0.07	0.03	0.11
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)		(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Edu**5**	0.24	0.17	0.20	0.17	0.22	0.17	Edu**5**	0.44	0.20	0.17	0.12	0.08	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)		(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.00)

(b). Average mobility in the 2004-05 round

Non-SC/ST							**SC/ST**						
	Edu**1**	Edu**2**	Edu**3**	Edu**4**	Edu**5**	size		Edu**1**	Edu**2**	Edu**3**	Edu**4**	Edu**5**	size
Edu**1**	0.79	0.09	0.06	0.04	0.02	0.13	Edu**1**	0.87	0.06	0.05	0.02	0.01	0.23
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)		(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)
Edu**2**	0.61	0.26	0.07	0.04	0.02	0.10	Edu**2**	0.67	0.22	0.07	0.02	0.01	0.14
	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)		(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Edu**3**	0.45	0.21	0.19	0.10	0.05	0.17	Edu**3**	0.58	0.18	0.16	0.05	0.03	0.21
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)		(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Edu**4**	0.32	0.19	0.21	0.18	0.10	0.28	Edu**4**	0.47	0.17	0.17	0.13	0.06	0.26
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)		(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Edu**5**	0.19	0.11	0.16	0.19	0.36	0.32	Edu**5**	0.34	0.14	0.15	0.16	0.21	0.17
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)		(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)

Notes: Each cell ij represents the average probability (for a given NSS survey round) of a household head with education i having a child with education attainment level j. Column titled 'size' reports the fraction of parents in education category 1, 2, 3, 4, or 5 in a given survey round. Standard errors are in parenthesis.

holds separately for non-SC/STs and SC/STs for the sample years 1983 and 2004-05. For each NSS round we compute p_{ij} where i denotes the education category of the household head and j denotes the education category of the child. Thus, p_{ij} is the probability of a household head with education category i having a child with education category j. A high p_{ij} where $i = j$ reflects low intergenerational education mobility since they give the proportion of children who have the same education as their parent. A relatively large p_{ij} where $i \neq j$, on the other hand, would indicate high mobility.

Table 5 shows the results. Panel (a) shows the mobility matrix for 1983 while panel (b) reports the results for the 2004-05 sample round. Each row of the table shows the education of the parent while columns indicate the education category of the child. Column "size" reports the average share of parents with a given education attainment level in a given round. Thus, the row labelled "Edu1" in the top-left panel of the Table says that in 1983, 85 percent of the adult male children of illiterate non-SC/ST parents remained illiterate, 9 percent acquired some education, 5 percent finished primary school, 1 percent had middle school education, and almost none had secondary school education. The last entry in that row says that 32 percent of non-SC/ST parents were

illiterate in 1983. The associated standard errors are shown in parenthesis below the estimates.

Table 5 reveals some interesting features. For both groups, the intergenerational persistence of illiteracy has declined across the rounds. For non-SC/STs, 85 percent of the children of illiterate parents remained illiterate in the 1983 round. In 2004-05, the persistence of illiteracy had declined to 79 percent. For SC/STs, the corresponding numbers were 91 percent and 87 percent. Moreover, a large part of this upward intergenerational education mobility was children of illiterate parents beginning to acquire middle school or higher education levels. Hearteningly, the shares of illiterate parents also declined sharply across the rounds. For non-SC/STs, the share of illiterate parents declined from 32 to 13 percent while for SC/STs it fell from 56 to 23 percent.

Another positive feature of the time trends in education mobility for both groups was that amongst parents with primary school education and above (categories 3, 4 and 5), there was a significant decline in the share of children with lesser education attainment than their parents. Concurrently, both groups saw an increase in the persistence or improvement of the education status of children of parents with the relatively higher education levels of 4 and 5 (middle school or secondary school and above). Only in households in which the head of the household had below primary level of education (category 2) was there an increase in regress of education attainments of children. Even for these households though, the children that improved over their parents tended to do so by a large margin – they often acquired middle school or secondary and above education levels.

Overall, there was a clear trend of convergence of household education attainment levels of the two groups with sharper movements into categories 4 and 5 for SC/STs. Most importantly, the upward education mobility was not restricted to the more educated households. Rather, this appears to have been a more wide-spread phenomenon during this period.

3.2 Occupation Mobility

We now turn to intergenerational occupation mobility. The conditional probability of an occupation switch is obtained in a similar manner to the education switch probabilities. Now, y_i is a binary variable for occupation switch as defined above (*switch-occ*) while x_i is a vector of controls:

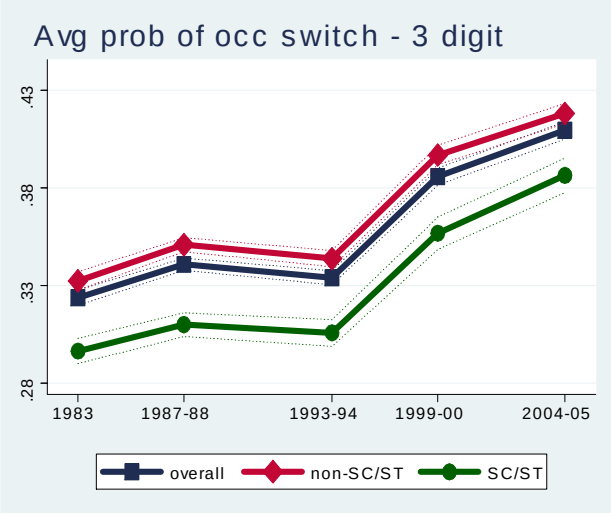
$$\begin{aligned} x_i\beta &= \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 SC/ST_i + \beta_4 muslim_i \\ &\quad + \beta_5 rural_i + \beta_6 hh\_size_i + \beta_7 quota\_scst_j + \sum_{j=1}^{4} \theta_j edu\_dum_j \end{aligned}$$

$$+\sum_{j=1}^{5}\alpha_j region\_dummy_j + \sum_{j=1}^{9}\gamma_j occup\_dummy_j. \tag{3.2}$$

In our model, the occupation switch for individual i depends on three sets of controls. The first set includes individual characteristics such as age, age squared, belonging to an SC/ST group ($SCST$), and religion ($muslim$). Second, we control for household-level characteristics such as household size (hh_size_i), and his rural location ($rural_i$). Third, we allow for occupation-specific fixed effects, region-level fixed effects, and state-level SC/ST reservation quotas.[10]

The model is estimated for each sample round separately and then used to obtain fitted values for each individual. These fitted values provide us with estimates of the probability of occupation switches in each round. We compute this measure of intergenerational occupational mobility for the overall sample as well as for SC/STs and non-SC/STs separately.

Figure 2: Intergenerational occupation switches



Notes: This figure presents the average predicted probability of intergenerational occupation switch for our overall sample, for SC/STs and non-SC/STs. The numbers are reported for the five NSS survey rounds. Dotted lines are ±2 std error bands.

Figure 2 depicts the computed probabilities of occupation switches at the three-digit level (dotted lines plot the ± 2 standard error confidence bands). As the Figure shows, the overall probability of an occupation switch by the next generation relative to the household-head has steadily increased from 32 percent in 1983 to 41 percent in 2004-05. This increase has been mirrored in the two sub-groups with the switch probabilities rising for both. For non-SC/STs the switch probability has risen from 33 to 42 percent while for SC/STs it has gone from 30 to 39 percent.

[10] Occupation fixed effects are defined for one-digit occupation categories.

Crucially, there is no trend towards convergence of these probabilities across the two groups which indicates that differences in intergenerational mobility between them has not changed over this period. We also estimated the occupation switch probabilities at the one-digit and two-digit levels and found that the patterns are similar to the three-digit probabilities. The main difference is that the probability of an occupation switch is universally lower at the two-digit and more so at the one-digit level.[11]

Table A7 in the appendix reports the detailed regression results. The noteworthy feature about those results is that the SC/ST dummy is consistently positive across the rounds even though it is at times insignificant. Hence, after controlling for the covariates of occupation choice, SC/ST effect on the probability of switching occupations was actually non-negative. This indicates that the overall lack of convergence of occupation switch rates between the groups was due to a lack of complete convergence in the other covariates rather than due to caste related factors.

3.2.1 Occupation Transition Matrix

While the overall probability of switches indicates the degree of mobility across occupations, we are also interested in determining the pattern of movements within occupations: children who are switching are most likely to have parents working in which occupation? Which sectors are absorbing most of the intergenerational switchers? Have these trends varied over time? Are there any differences between SC/STs and non-SC/STs in these patterns?

To address these issues, we compute the transition probabilities across occupations. Thus, for each NSS round we compute p_{ij} where i denotes the occupation of the household head and j denotes the occupation of the child. Thus, p_{ij} is the probability of a household head working in occupation i having a child working in occupation j. Clearly, a high p_{ij} where $i = j$ would reflect relatively little intergenerational occupational mobility while large p_{ij} where $i \neq j$ would indicate high mobility.

To facilitate presentation of results, we compute transition probabilities for the three broad occupation categories. In particular, we aggregate the 3-digit occupation codes that individuals report into a one-digit code, leaving us with ten categories. We then group these ten categories further into three broad occupation categories: Occ 1 comprises white collar administrators, executives, managers, professionals, technical and clerical workers; Occ 2 collects blue collar workers such as sales workers, service workers and production workers; while Occ 3 collects farmers, fishermen, loggers, hunters etc.. This grouping reflects the similarity of occupations based on skill

[11] The results for the one- and two-digit occupation categories are available upon request.

Table 6: Intergenerational occupation transition probabilities

(a). Average mobility in the 1983 round

Non-SC/ST	To				**SC/ST**	To			
From	Occ **1**	Occ **2**	Occ **3**	size	From	Occ **1**	Occ **2**	Occ **3**	size
Occ **1**	0.49	0.33	0.18	0.06	Occ **1**	0.29	0.40	0.31	0.03
	(0.02)	(0.01)	(0.01)	(0.00)		(0.05)	(0.06)	(0.05)	(0.00)
Occ **2**	0.06	0.82	0.12	0.26	Occ **2**	0.04	0.77	0.19	0.20
	(0.00)	(0.01)	(0.01)	(0.00)		(0.01)	(0.01)	(0.01)	(0.01)
Occ **3**	0.03	0.10	0.86	0.67	Occ **3**	0.02	0.09	0.90	0.78
	(0.00)	(0.00)	(0.01)	(0.00)		(0.00)	(0.01)	(0.01)	(0.01)

(b). Average mobility in the 2004-05 round

Non-SC/ST	To				**SC/ST**	To			
From	Occ **1**	Occ **2**	Occ **3**	size	From	Occ **1**	Occ **2**	Occ **3**	size
Occ **1**	0.48	0.38	0.14	0.10	Occ **1**	0.35	0.45	0.20	0.05
	(0.01)	(0.01)	(0.01)	(0.00)		(0.03)	(0.03)	(0.03)	(0.00)
Occ **2**	0.07	0.84	0.09	0.30	Occ **2**	0.04	0.85	0.11	0.27
	(0.00)	(0.01)	(0.00)	(0.00)		(0.01)	(0.01)	(0.01)	(0.01)
Occ **3**	0.04	0.19	0.77	0.60	Occ **3**	0.03	0.18	0.79	0.68
	(0.00)	(0.01)	(0.01)	(0.00)		(0.00)	(0.01)	(0.01)	(0.01)

Notes: Each cell ij represents the average probability (for a given NSS survey round) of a household head working in occupation i having a child working in occupation j. Occ 1 collects white collar workers, Occ 2 collects blue collar workers, while Occ 3 refers to farmers and other agricultural workers. Column titled 'size' reports the fraction of parents employed in occupation 1, 2, or 3 in a given survey round. Standard errors are in parenthesis.

requirements.[12]

Table 6 presents the results. Each row of the Table denotes the occupation of the parent while columns indicate the occupation of the child. Thus, going across columns along any row i would indicate the probability of a household head working in occupation i to have a child working in the relevant occupation column. Clearly, off-diagonal elements measure the degree of intergenerational occupational mobility. Column "size" reports the average share of parents employed in each of the occupations in a given round. The Table has two panels: Panel (a) gives the numbers for 1983 and Panel (b) for 2004-05. Bracketed numbers below the estimated probabilities are the associated standard errors.

Table 6 reveals a few interesting features. First, the diagonal elements of both Panel (a) and (b) are quite high, indicating relatively little intergenerational occupation mobility over this period. The highest persistence rates (or the least mobility) in 1983 was in occupation 3 (agriculture) for both SC/STs and non-SC/STs with the persistence rate being slightly higher for SC/STs. In 2004-05, the persistence rate in occupation 3 was significantly lower for both caste groups, though the SC/ST rate remained larger. The intergenerational persistence in occupation 2, in contrast, increased, and significantly so for SC/STs. In fact, in the 2004-05 round, occupation 2 shows the

[12] We confirm that our occupation groupings are plausible by examining education attainments and wages of the three groups. Indeed, Occ 1 is characterized by the highest education attainments and wages, followed by Occ 2, and Occ 3. See Appendix A for more details on the definitions of occupation categories.

most intergenerational persistence among all occupations. Interestingly, SC/STs also experienced a large increase in intergenerational persistence in occupation 1, while non-SC/STs saw a reduction in that persistence. These trends imply a dramatic convergence in the intergenerational persistence of all occupations between the two caste groups.

Second, the probability of the son of a farmer (working in occupation 3) switching to occupations 1 or 2 has risen for both groups. This probability is of interest as it indicates an improvement in the quality of jobs across generations. In 1983 the probability of an intergenerational switch from occupation 3 to occupations 1 or 2 was 13% for non-SC/STs and 11% for SC/STs. By 2004-05 these numbers had risen to 23% for non-SC/STs and 21% for SC/STs. We interpret these findings as evidence of convergence in upward occupation mobility of both caste groups, with SC/STs experiencing larger positive changes.

Third, the probability of a child working in occupation 3 conditional on his father being employed in occupation 1 or 2 has declined from 50% to 31% for SC/STs and from 30% to 23% for non-SC/STs over our sample period. We believe that this reflects a significant reduction in regress prospects of SC/ST households during this period.

Lastly, an interesting feature of this period has been a slight increase in the probability of an intergenerational switch from occupation 1 to occupation 2 for both groups, i.e., children switching from the white collar occupations of their father to working in blue-collar jobs. This mostly reflects an increase in the share of the sales and service sectors during the 1990s after the reforms – an outcome of the key changes that the economy was undergoing in its industrial composition during this period. We turn next to examining more generally the intergenerational mobility implications of the large-scale sectoral transformation of the economy during this period.

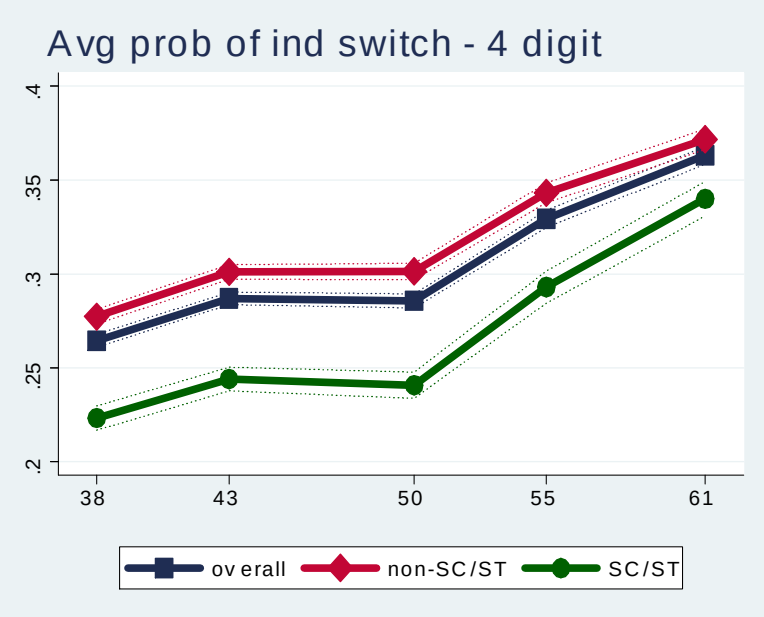
3.3 Industry Mobility

Given the large sectoral changes in India during the period under study, an issue of independent interest is the degree of industry mobility in India between 1983 and 2004-05. We define intergenerational industry switch in the same manner as occupation switches and estimate the conditional probability of industry switches using equation symmetric to equation (3.2), except we included industry-specific fixed effects instead of occupation-specific fixed effects.[13]

Figure 3 presents the overall probability of industry switches at the four-digit level as well as the probability of switches for SC/STs and non-SC/STs (dotted lines plot the ± 2 standard error

[13] Industry fixed effects are defined for one-digit industry categories.

Figure 3: Intergenerational industry switches



Notes: This figure presents the average predicted probability of intergenerational industry switch for our overall sample, for SC/STs and non-SC/STs. The numbers are reported for the five NSS survey rounds. Dotted lines are ±2 std error bands.

confidence bands). The figure shows that the overall probability of children switching the industry of employment relative to their parent has risen from 26 percent in 1983 to 36 percent in 2004-05 period. The industry mobility trends of both SC/STs and non-SC/STs have converged marginally although the level of the switching probability has remained significantly higher for non-SC/STs. We also estimated the probability of industry switching at the three-, two-, and one-digit levels and found similar time-series trends, with little convergence across the two groups. As with the occupation mobility estimates, the main difference when considering more aggregated industry categories is that the probability of an industry switch is universally lower.[14] Lastly, in Table A8 we report the detailed regression results for the industry switches. The key point to note from those results is that the SC/ST dummy is positive but insignificant for all rounds except 1999-2000. Hence, for the most part, the differences in the probability of industry switches between the castes have been driven by differences in the other covariates of industry choice.

[14] The results for three-, two- and one-digit industry categories are available upon request.

Table 7: Intergenerational industry transition probabilities

(a). Average mobility in the 1983 round

Non-SC/ST	To				**SC/ST**	To			
From	Ind **1**	Ind **2**	Ind **3**	size	From	Ind **1**	Ind **2**	Ind **3**	size
Ind **1**	0.87	0.04	0.09	0.67	Ind **1**	0.90	0.04	0.06	0.78
	(0.00)	(0.00)	(0.00)	(0.00)		(0.01)	(0.00)	(0.00)	(0.01)
Ind **2**	0.09	0.75	0.16	0.11	Ind **2**	0.16	0.68	0.16	0.08
	(0.01)	(0.01)	(0.01)	(0.00)		(0.02)	(0.03)	(0.02)	(0.00)
Ind **3**	0.15	0.12	0.73	0.22	Ind **3**	0.22	0.09	0.69	0.14
	(0.01)	(0.01)	(0.01)	(0.00)		(0.02)	(0.01)	(0.02)	(0.01)

(b). Average mobility in the 2004-05 round

Non-SC/ST	To				**SC/ST**	To			
From	Ind **1**	Ind **2**	Ind **3**	size	From	Ind **1**	Ind **2**	Ind **3**	size
Ind **1**	0.77	0.05	0.18	0.60	Ind **1**	0.79	0.05	0.17	0.68
	(0.01)	(0.00)	(0.01)	(0.01)		(0.01)	(0.00)	(0.01)	(0.01)
Ind **2**	0.07	0.70	0.23	0.11	Ind **2**	0.09	0.66	0.25	0.08
	(0.01)	(0.01)	(0.01)	(0.00)		(0.02)	(0.03)	(0.03)	(0.00)
Ind **3**	0.12	0.12	0.76	0.29	Ind **3**	0.13	0.14	0.73	0.24
	(0.01)	(0.01)	(0.01)	(0.00)		(0.01)	(0.01)	(0.02)	(0.01)

Notes: Each cell ij represents the average probability (for a given NSS survey round) of a household head working in industry i having a child working in industry j. Ind 1 refers to agriculture, Ind 2 collects manufacturing and mining&quarrying, while Ind 3 refers to services. Column titled 'size' reports the fraction of parents employed in industry 1, 2, or 3 in a given survey round. Standard errors are in parenthesis.

3.3.1 Industry Transition Matrix

We now turn to the industry choices of the two groups. Using the same approach that we employed to evaluate occupation mobility, we compute industry transition probabilities. As with occupations, in order to facilitate the presentation, we aggregate the 4-digit industry code that individuals report into a one-digit code. This gives us seventeen categories. We then group these seventeen categories into three broad industry categories: Ind 1, Ind 2 and Ind 3. Ind 1 comprises the Agricultural sector, Ind 2 collects Manufacturing and Mining and Quarrying, while Ind 3 comprises all Service industries. Our grouping reflects the traditional industrial classification according to United Nations classification system.[15]

The results are summarized in Table 7. As with occupation transition probabilities, each row of the Table denotes the industry of the parent's employment while columns indicate the industry of the child's employment. Thus, going across columns along any row i would indicate the probability that a household-head working in industry i has a child working in the relevant industry column. Off-diagonal elements measure the degree of intergenerational industry mobility. Column "size" reports the average share of parents employed in each of the industries in a given round. Panel (a) gives the numbers for 1983 and panel (b) for 2004-05. The standard errors are reported within brackets under the relevant estimated probability.

[15] See Appendix A for more details on the industry grouping.

Not surprisingly, Ind 1 (agriculture) has remained the primary industry of employment for both SC/STs and non-SC/STs throughout, although its share has declined significantly between 1983 and 2004-05. Ind 1 also has the highest persistence of the three industry groups. The numbers indicate that intergenerational industry persistence has decreased sharply for Ind 1. Children are switching from agriculture into other industries more frequently in 2004-05 in comparison with 1983. While most of this move is primarily into service industries, the probability of moving into manufacturing has increased, especially for SC/STs. At the same time, the probabilities of moving from Ind 2 or Ind 3 into Ind 1 have declined and more so for SC/STs. We interpret these results as evidence of upward industry mobility, especially for SC/STs.

3.4 Income Mobility

Our fourth, and probably the most typical, measure of intergenerational mobility is on income. We proxy income with the individual's wage. Before describing our results we should note that the sample size for the wage data is, on average, a third of the sample size for the education and occupation distribution data due to a large number of households with missing wage observations. The missing wage observations are mostly accounted for by the segment of the rural population who identify themselves as being self-employed and therefore do not report any wage data. Across the rounds, on average, about 65 percent of the sample are self-employed with 76 percent of them residing in rural areas. The missing wage data raises sample selection concerns. In particular, if non-SC/ST rural households are more likely to be land-owning and hence self-employed, then the wage data (particularly for rural households) would be skewed towards landless SC/ST households. The problem would be compounded by the fact that the wage earning non-SC/ST households may also be the most worse off amongst the non-SC/STs who may have the lowest mobility rates. In this event we would be biasing our results toward finding low wage mobility gaps between the two groups.

We examined this issue in two ways. First, on average, 21 percent of the self-employed belong to SC/ST households. This is comparable to the 24 percent share of SC/STs in our working sample. Clearly, SC/STs are not disproportionately under-represented amongst the self-employed. Second, to assess the seriousness of the potential sample selection problem, we computed the per capita household consumption expenditure of non-SC/STs relative to SC/STs for self-employed households and wage earning households separately. Stable across rounds, the ratio was 1.24 for both. Hence, self-employed households do not appear to be distinctly different from wage earning households.

Based on these two findings, we feel that the sample selection issues raised by the missing wage observations are not too serious and that the patterns of inter-group welfare dynamics indicated by the wage data are likely to generalize to the self-employed as well.

We should also note one important anomaly in the 1987-88 round of the survey. We find that the number of observations for wages in this round falls substantially relative to the other rounds. This occurs due to a very large and disproportionate decline in the rural wage observations for this round. We could not find any explanations in the data documentation or in conversation with NSS officials as to the reasons for this sudden decline in the number of wage observations in the 43th round. For the sake of completeness though, we report the results for all rounds. However, the results for the 43th round should be treated with caution on account of the missing rural wage observations.

The goal of measuring income mobility is to provide a measure of the degree to which the long run income of a child of a family is correlated with the long run income of his father. One such commonly used measure is the intergenerational elasticity (IGE). IGE of long run income is typically estimated as the slope coefficient in a regression of the log of the long run income (relative to the mean) of the child on the log of the parents' long run income (relative to the mean for the parents' generation). The estimated coefficient indicates the degree to which income status in one generation gets transmitted to the next generation. More precisely, IGE provides a measure of intergenerational persistence in income, while one minus IGE measures intergenerational mobility.

The typical problem surrounding income mobility regression specifications is the absence of measures of long run income. The standard procedure is to use short run measures of income as proxies for long run income. We face the same problem since our income data is the daily wage during the census period. Clearly, the daily wage may be a very noisy measure of long run income with significant associated measurement error. Moreover, as pointed out by Haider and Solon (2006), an additional problem with using short run measures for children's income is the systematic heterogeneity in income growth over the life cycle. In particular, individuals with higher lifetime income also tend to have steeper income trajectories. As a result, early in the lifecycle, current income gaps between those with high lifetime incomes and those with low lifetime incomes tend to understate their lifetime income differences while current income gaps later in the lifecycle overstate the lifetime income gaps.

We follow Lee and Solon (2009) to address these issues by (a) introducing controls for children's age to account for the stage of the life-cycle at which the income is observed; (b) introduce an

interaction between parents's income and children's age to account for the systematic heterogeneity in the profiles; and (c) by instrumenting parents's income with household consumption expenditure and household size to mitigate the measurement error associated with using daily wage data. Hence, our regression specification is

$$\begin{aligned} w_{ic} &= \alpha + \beta w_{ip} + \gamma_1 A_{ip} + \gamma_2 A_{ip}^2 + \gamma_3 A_{ip}^3 + \delta_1 \tilde{A}_{ic} + \delta_2 \tilde{A}_{ic}^2 + \delta_3 \tilde{A}_{ic}^3 \\ &\quad + \theta_1 w_{ip} \tilde{A}_{ic} + \theta_2 w_{ip} \tilde{A}_{ic}^2 + \theta_3 w_{ip} \tilde{A}_{ic}^3 + \varepsilon_i \end{aligned} \tag{3.3}$$

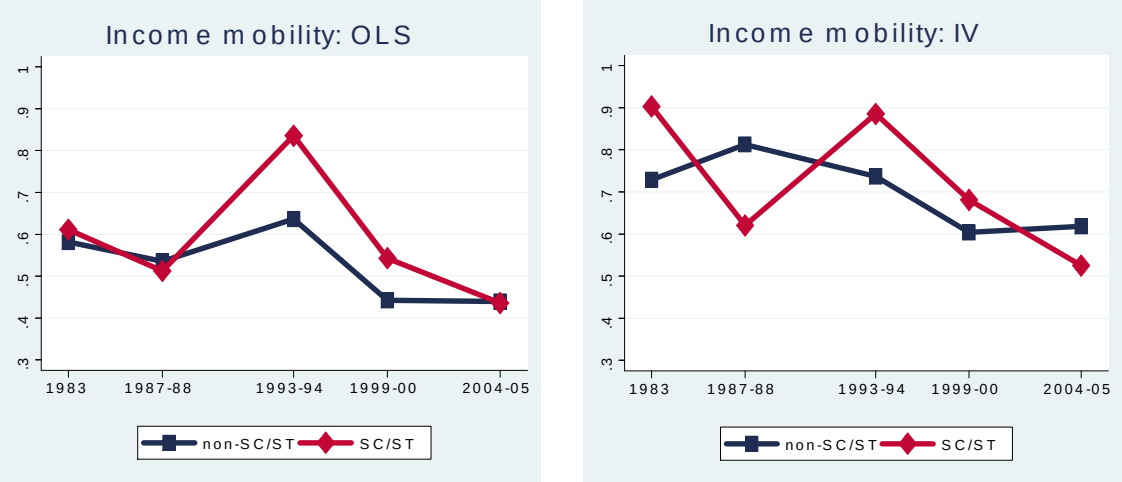
where w_{ic} denotes the log daily wage of the child of household i and w_{ip} is the log daily wage of the male head of the same household. A_{ip} denotes the head of household i's age while $\tilde{A}_{ic}$ is the child's age, which we normalized to equal zero at age 23 which is the mean age of children in our sample.

The control for a cubic in parents' age is to account for differences in the ages of parents in the sample at the time of observing their child's income. As pointed out in Haider and Solon (2006), the short run proxy for long run income of parents will bias the estimated β downward. However, as long as the bias is stable over time it will not alter the interpretation of how the intergenerational elasticity of income has evolved over time.

We run this regression separately for each NSS sample year and for each caste group. The key parameter of interest is β. We estimate a different β for each NSS round and focus on how the estimated β's have changed over the sample period. The detailed estimation results are reported in the Appendix E. Table A9 reports the OLS regressions while Table A10 reports the results of our Instrumental Variable (IV) regressions where we instrument parent's income with household consumption expenditure and household size.

We plot the OLS estimates in panel (a) of Figure 4 below, while panel (b) of the Figure presents our estimates from an instrumental variable (IV) regression. We should note that all the point estimates in both figures are significant at the 1 percent level except for the OLS estimate for 1987-88 which is significant at the 5 percent level. There are three features of the results worth noting. First, the income persistence across generations has declined sharply over the period 1983 and 2004-05 for both SC/STs and non-SC/STs. In fact by the end of our sample period the estimates are much closer to the typical numbers around 0.45 that are reported for the USA by a number of different studies (see Solon, 2002). Second, there has been a clear convergence in intergenerational income persistence across the two groups.

Figure 4: Intergenerational income mobility



(a) (b)

Notes: Figures (a) and (b) present the results from the OLS and IV regressions, respectively, of child's per day log real wage on parent's per day log real wage and a set of controls. The figure plot the coefficients on the parent's wage from those regressions estimated separately for non-SC/STs and SC/STs. All estimated coefficients are statistically significant. Detailed estimation results are presented in the Appendix.

Third, the IV estimates are uniformly higher than the OLS estimates. This is similar to the findings of Solon (1992) for the US. More importantly, they confirm our findings from the OLS estimation. In fact, IV estimates suggest that SC/STs' intergenerational income persistence has declined from a whopping 0.87 to 0.45 and, by the end of our sample period, was below that for non-SC/STs.

One drawback of IGE when comparing intergenerational mobility of subpopulations is that each group's mobility measure only captures the persistence of that group relative to its mean, not the mean of the entire distribution. For instance, the IGE coefficient for SC/STs tells us the rate at which income of an SC/ST child regresses to the mean of the SC/ST income distribution. To the extent that SC/STs and non-SC/STs mean earnings are different and are changing over time, the IGE coefficients of the two groups will be of limited comparability. Furthermore, if mobility patterns are different at various points in the income distribution, IGE will not be able to capture these differences.

To account for both shortcomings, two alternative approaches to measuring intergenerational mobility have been proposed in the literature (see Black and Devereux (2010) for a review of the literature). The first approach consists of computing mobility matrices which summarize transition probabilities of child's earnings conditional on father's earnings for different quantiles. Transition

probabilities for each social group are obtained using distributions for the entire generation comprising both social groups. This facilitates meaningful comparisons of mobility patterns across subpopulations.

Another approach to measuring intergenerational mobility has been developed recently by Bhattacharya and Mazumder (2007, 2011). They criticize the existing transition probability approach as being sensitive to the choice of quantiles, that is it predicts different mobility patterns depending on whether the researcher used quintiles, quartiles, etc. Instead, they propose to compute upward mobility measure which measures the probability that son's relative standing in his generational distribution exceeds the relative standing of the father in his generational distribution. A key advantage of this approach is that it accounts for even small upward movements in son's relative position, thus providing a more forgiving measure of mobility. In contrast, mobility matrices require son's income to improve sufficiently to jump the specified quantile. Given that SC/STs are typically poorer than non-SC/STs for every quantile, SC/STs sons would be required to make larger income gains than non-SC/ST sons in order to record an improvement in income mobility.

We conduct both evaluations next. Following Jäntti et al. (2006) we begin by computing mobility matrices for SC/STs and non-SC/STs based on income quintiles. The results are presented in Table 8. Each row j of the table reports the probability of child's income being in quintile $k = 1..5$, conditional on father's income being in quintile j. These matrices are reported separately for SC/STs and non-SC/STs, but are computed using the entire income distribution of SC/STs and non-SC/STs for each generation. Panel (a) reports the results for 1983, while panel (b) does the same for 2004-05 survey round.[16]

Several feature of the data stand out from the table. First, in 1983 the intergenerational income persistence, as captured by the diagonal entries in the mobility matrices, was substantially larger for SC/STs relative to non-SC/STs located in the bottom quintiles of income distribution; while it was significantly smaller in the top quintiles of income distribution. That is, the son of low income SC/ST was *more* likely to remain in the bottom income quintiles than the son of low income non-SC/ST. At the same time, the son of a high income SC/ST was *less* likely to remain in the high income quintiles relative to non-SC/ST sons. The situation changes a lot by 2004-05. In particular, the intergenerational income persistence has declined for both social groups for all quintiles.[17]

[16] Standard errors are computed using bootstrap procedure in which we accounted for the complex survey design of the NSS data. In particular, in our procedure we use adjusted sampling weights. The variance is estimated using the resulting replicated point estimates (see Rao and Wu (1988), and Rao et al. (1992)).

[17] The only exception was the 5th quintile of SC/STs, in which the persistence has increased over time.

Table 8: Intergenerational income transition probabilities

(a). Average mobility in the 1983 round

Non-SC/ST							**SC/ST**					
	q1	q2	q3	q4	q5	size	q1	q2	q3	q4	q5	size
q1	0.51	0.36	0.08	0.04	0.01	0.17	0.57	0.29	0.08	0.03	0.02	0.24
	(0.05)	(0.05)	(0.02)	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.03)	(0.01)	(0.01)	(0.01)
q2	0.18	0.44	0.30	0.06	0.02	0.18	0.13	0.51	0.30	0.04	0.02	0.23
	(0.02)	(0.05)	(0.04)	(0.02)	(0.01)	(0.01)	(0.02)	(0.05)	(0.04)	(0.02)	(0.01)	(0.01)
q3	0.14	0.17	0.44	0.20	0.05	0.18	0.07	0.14	0.45	0.30	0.04	0.22
	(0.02)	(0.02)	(0.04)	(0.03)	(0.01)	(0.01)	(0.02)	(0.02)	(0.04)	(0.04)	(0.01)	(0.01)
q4	0.09	0.06	0.11	0.49	0.25	0.20	0.06	0.04	0.08	0.45	0.37	0.21
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.01)
q5	0.07	0.04	0.08	0.17	0.64	0.27	0.05	0.07	0.16	0.24	0.48	0.10
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.04)	(0.05)	(0.06)	(0.01)

(b). Average mobility in the 2004-05 round

Non-SC/ST							**SC/ST**					
	q1	q2	q3	q4	q5	size	q1	q2	q3	q4	q5	size
q1	0.52	0.33	0.08	0.04	0.03	0.20	0.58	0.33	0.03	0.04	0.02	0.22
	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.01)	(0.01)	(0.02)	(0.01)
q2	0.15	0.40	0.34	0.06	0.05	0.20	0.15	0.32	0.42	0.05	0.06	0.20
	(0.03)	(0.04)	(0.05)	(0.02)	(0.01)	(0.01)	(0.03)	(0.05)	(0.05)	(0.01)	(0.03)	(0.01)
q3	0.07	0.15	0.35	0.35	0.08	0.19	0.07	0.08	0.34	0.45	0.06	0.24
	(0.01)	(0.03)	(0.05)	(0.05)	(0.02)	(0.01)	(0.02)	(0.02)	(0.05)	(0.05)	(0.02)	(0.02)
q4	0.09	0.12	0.11	0.35	0.32	0.19	0.09	0.05	0.11	0.34	0.41	0.20
	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.01)
q5	0.11	0.08	0.09	0.18	0.54	0.22	0.11	0.11	0.10	0.16	0.52	0.14
	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.01)	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)	(0.01)

Note: Each cell i j reports the probability (for a given NSS survey round) of a household head with income in quintile i having his child earning income in quintile j. q1-q5 refer to the quintile of the generational income distribution (fathers' in the columns; kids' in the rows). Column "size" refers to the fraction of parents falling in a given income quintile in that round. Bootstrapped standard errors are in parenthesis.

Second, the decline in persistence was accompanied by an increase in upward intergenerational income mobility of both social groups, with SC/STs often experiencing more dramatic improvements. In fact, by 2004-05 SC/STs have surpassed the non-SC/STs in terms of upward income mobility for all quintiles except the very bottom quintile (q1 in the table). Interestingly, Table 8 also shows that for non-SC/STs, the highest income quintile households (q5 in the table) actually experienced an increase in the probability of intergenerational regress for quintiles 4 and 5 whereas the corresponding SC/ST households had the opposite or more muted trends.

A third noteworthy feature is that the biggest movements in the intergenerational income transitions for both groups have occurred in the middle of the distribution – quintiles 2 and 3 for SC/STs and quintile 3 for non-SC/STs. Thus, the probability of a SC/ST father in the 2nd income quintile having a son in the third and higher quintiles was 0.36 in 1983 and increased to 0.52 in 2004-05. The same probability for non-SC/ST sons was 0.38 in 1983 and 0.44 in 2004-05. For the SC/ST fathers in the third income quintile, the upward transitions probabilities for their sons were 0.35 in 1983 and 0.52 in 2004-05; for non-SC/ST the corresponding probabilities were 0.26 in 1983 and 0.42 in 2004-05.

What factors might account for these large changes in the middle of the income distribution rather than at the very top or bottom? One reason may be the evolution of credit constraints during this period. There is a large literature going back to Becker and Tomes (1986) that has studied the effects of credit constraints on intergenerational mobility. While there are many nuances to the relationship between the two (see Grawe and Mulligan (2002) for an in-depth review), one argument due to Corak and Heisz (1999) holds that the wealthiest households can self-finance education while the constraint is unlikely to bind for the lowest income groups as long as ability and earnings are correlated on average. Hence, credit constraints are most likely to bind for the middle income groups. Corak and Heisz (1999) find support for this view in the Canadian data. Clearly, this link works through the education channel, i.e., changes in investments in human capital of children affect their earnings. One of the key characteristics of the reforms in India during our sample period of 1983 to 2005 was the progressive liberalization of the financial sector. This has made credit access in India both wider and deeper. In as much as a relaxation of binding credit constraints may have facilitated greater investments in children's education, this may be an explanation for the large increases in intergenerational mobility in the middle of the income distribution in India. However, whether this can account for the sharper changes amongst SC/ST households is a question that would require a more detailed study that is beyond the scope of this paper.[18]

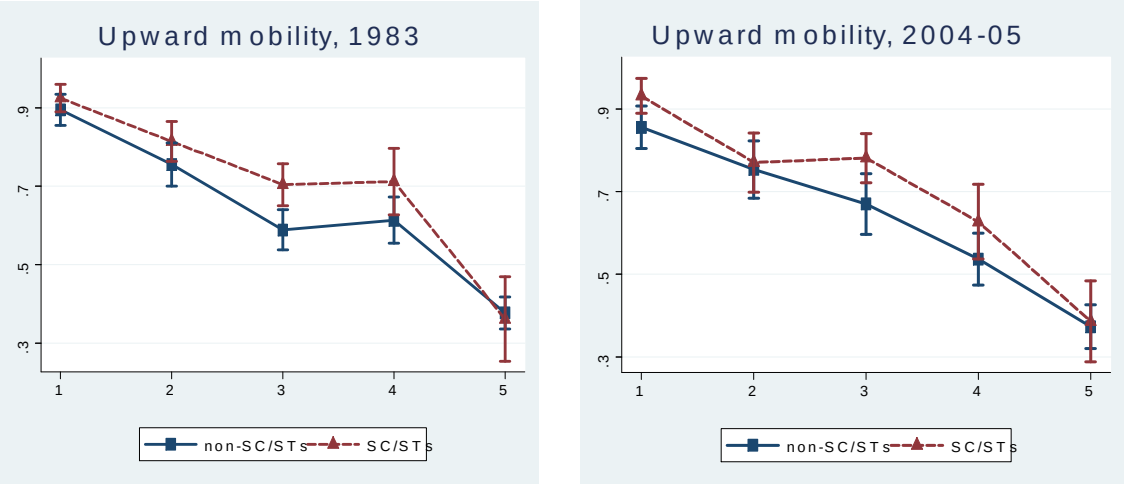
Next we compute a more direct measure of upward mobility, as proposed by Bhattacharya and Mazumder (2011). It is computed as the probability that a son's income rank in his generational income distribution exceeds the income rank of his father in the income distribution of father's generation. Figure 5 reports the estimated probabilities for SC/STs and non-SC/STs conditional on father's quintile. Panel (a) is for 1983, while panel (b) is for 2004-05 survey round.

Interestingly, according to this measure, SC/ST kids show higher upward mobility than non-SC/ST kids for all quintiles of fathers' distribution, and this is so in both 1983 and 2004-05 survey rounds.[19] This result is further confirmed by examining a measure of upward mobility obtained by conditioning on father's income being below some threshold, rather than falling within a particular quintile. As noted in Bhattacharya and Mazumder (2011), this approach avoids the aggregation

[18] We should note that easier credit conditions need not affect earnings only through greater enrollment in higher education categories. If parents supplement the school education of their children with private lessons or switch from public to better quality private education, then children could have greater earnings due to the higher quality of their human capital. In the Indian context where the state provides widespread education subsidies but there are limits on the number of available seats in post-secondary education institutions, easier credit also allows families to enrol their children in private institutions where there are no state subsidies.

[19] The differences however are mostly insignificant.

Figure 5: Intergenerational upward mobility in income



(a) (b)

Note: Figures present upward income mobility conditional on parent's quintile (see text for details). Panel (a) is for 1983, while panel (b) is for 2004-05 survey round. Bootstrapped 95% pointwise confidence intervals are shown as bands.

bias that may arise because of income heterogeneity within a given quintile. We report the results in Table 9, where we use quintiles of father's distribution as the threshold for computing upward mobility. These results indicate that upward mobility for SC/STs is higher than for non-SC/STs in both periods. In combination with the transition probabilities obtained in Table 8, we deduce that SC/ST children were more likely to improve their relative standing in income distribution as compared to non-SC/STs even though the size of the improvements tended to be smaller than for non-SC/STs. The magnitude of these improvements, however, increased over time, especially for SC/STs.

Overall, our results suggest that there has indeed been an upward trend in the degree of intergenerational mobility in education, occupation, industry and income of SC/STs with a significant convergence in the intergenerational educational and wage mobility to non-SC/ST levels.

4 Conclusion

In this paper we have contrasted the evolution of intergenerational mobility rates in education attainment rates, occupation and industry choices and wages of scheduled castes and tribes (SC/STs) between 1983 and 2005 with the corresponding mobility rates of non-SC/STs. Using successive rounds of the National Sample Survey (NSS), we have shown that this period has been marked

Table 9: Intergenerational upward mobility in income

	(a) 1983		(b) 2004-05	
	non-SC/ST	SC/ST	non-SC/ST	SC/ST
q1	0.90	0.93	0.86	0.93
	(0.0193)	(0.0162)	(0.0276)	(0.0214)
q2	0.83	0.87	0.81	0.86
	(0.0163)	(0.0150)	(0.0240)	(0.0196)
q3	0.75	0.82	0.76	0.83
	(0.0187)	(0.0151)	(0.0186)	(0.0185)
q4	0.71	0.79	0.71	0.78
	(0.0172)	(0.0122)	(0.0152)	(0.0177)
q5	0.62	0.75	0.63	0.73
	(0.0165)	(0.0142)	(0.0138)	(0.0159)

Note: Table upward income mobility conditional on parent income being below a threshold. Quintiles are used as thresholds (see text for details). Panel (a) is for 1983, while panel (b) is for 2004-05 survey round. Bootstrapped standard errors are in parenthesis.

by a remarkable convergence in the intergenerational mobility rates of SC/STs to non-SC/ST levels in both education attainment and wages. SC/STs have also been switching occupations and industry of employment relative to their parents at increasing rates during this period and have matched non-SC/STs in this regard. Interestingly, we have found that a common feature for both SC/STs and non-SC/STs is that the sharpest changes in intergenerational income mobility has been for middle income households. This is consistent with the effects of easing credit constraints, a phenomenon that did characterize this period.

While we have focused here on disparities in inter-generational social mobility, in related work in Hnatkovska et al. (2011) we have also studied *intra*-generational disparities between SC/STs and non-SC/STs *within* age cohorts. In findings mirroring those in this paper, we found intra-generational gaps in education attainment levels, occupation choices, wages and consumption also declined between 1983 and 2004-05. The two sets of results combined suggest to us that the past three decades of major macroeconomic changes in India have also coincided with a rapid and significant reduction in caste-based restrictions to socioeconomic mobility.

We believe that there are three candidate explanations for the observed pick-up in the socioeconomic mobility rates of SC/STs in India. First, the past 25 years have been a period of major economic reforms in India. The reforms have unleashed strong competitive pressures on vast segments of a previously protected economy. As has been argued by Becker (1957), increasing competition could reduce discrimination by making it more expensive for businesses to pursue discriminatory labor market practises. This could reduce caste-based discrimination in both hiring and wages and thereby induce a faster rise in the intergenerational mobility rates of SC/STs. In

as much as a decline in wage and employment discrimination raises the perceived returns from getting educated for SC/STs, this could also induce a faster increase in intergenerational education mobility of SC/STs.

Second, a strengthening of caste-based networks of SC/STs could have been at play during this period. As has been recently shown in Munshi (2010), caste-based networks can often form quickly amongst the more disadvantaged groups in order to help them escape low-skill occupation traps. Specifically, these types of community based networks can often substitute for low endowments of inherited human capital stocks. The increasing political empowerment of the lower castes over the past 30 years may have been a contributing factor as well in accelerating this process.

Lastly, reservations policy in place for higher education and public sector employment over the past 60 years could also have played a significant *albeit* delayed role in reducing the wage and employment gaps between SC/STs and others. This, in turn, could have increased the intergenerational mobility rates of SC/STs relatively more. We intend to study these alternative explanations in greater detail in subsequent work.

References

Becker, G. (1957). *The Economics of Discrimination*. University of Chicago Press, Chicago.

Becker, G. and Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, 4(3):1–39.

Behrman, J. and Taubman, P. (1985). Intergenerational earnings mobility in the United States: some estimates and a test of Becker's intergenerational endowments model. *The Review of Economics and Statistics*, 67(1):144–151.

Bhattacharya, D. and Mazumder, B. (2007). Nonparametric analysis of intergenerational income mobility with application to the united states. Working Paper Series WP-07-12, Federal Reserve Bank of Chicago.

Bhattacharya, D. and Mazumder, B. (2011). A nonparametric analysis of black-white differences in intergenerational income mobility in the united states. Working paper series, Oxford University.

Black, S. E. and Devereux, P. J. (2010). Recent developments in intergenerational mobility. NBER Working Papers 15889, National Bureau of Economic Research, Inc.

Corak, M. and Heisz, A. (1999). The intergenerational earnings and income mobility of canadian men: Evidence from longitudinal income tax data. *Journal of Human Resources*, 34(3):504–533.

Grawe, N. D. and Mulligan, C. B. (2002). Economic interpretations of intergenerational correlations. *Journal of Economic Perspectives*, 16(3):45–58.

Haider, S. and Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, 96(4):1308–1320.

Hnatkovska, V., Lahiri, A., and Paul, S. (2011). Castes and labor mobility. Working papers, University of British Columbia.

Jalan, J. and Murgai, R. (2009). Intergenerational mobility in education in india. Working papers, Indian Statistical Institute, Delhi.

Jäntti, M., Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Österbacka, E., Björklund, A., and Eriksson, T. (2006). American exceptionalism in a new light: A comparison of intergenerational earnings mobility in the nordic countries, the united kingdom and the united states. IZA Discussion Papers 1938, Institute for the Study of Labor (IZA).

Lee, C. and Solon, G. (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics*, 91(4):766–772.

Maitra, P. and Sharma, A. (2009). Parents and children: Education across generations in india. Working papers, Monash University, Department of Economics.

Munshi, K. (2010). Strength in numbers: Networks as a solution to occupational traps. *Review of Economic Studies, forthcoming.*

Munshi, K. and Rosenzweig, M. (2006). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. *American Economic Review*, 96(4):1225–1252.

Munshi, K. and Rosenzweig, M. (2009). Why is mobility in india so low? social insurance, inequality, and growth. Working papers, Brown University, Department of Economics.

Pande, R. (2003). Can mandated political representation increase policy influence for disadvantaged minorities? theory and evidence from india. *American Economic Review*, 93(4):1132–1151.

Prakash, N. (2009). The impact of employment quotas on the economic lives of disadvantaged minorities in india. Working papers, Dartmouth College, Department of Economics.

Rao, J. N. K. and Wu, C. F. J. (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association*, 83:231–241.

Rao, J. N. K., Wu, C. F. J., and Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey Methodology*, 18:209Ű217.

Solon, G. (1992). Intergenerational income mobility in the united states. *The American Economic Review*, 82(3):393–408.

Solon, G. (2002). Cross-country differences in intergenerational earnings mobility. *The Journal of Economic Perspectives*, 16(3):59–66.

Appendix

A Data Appendix

A.1 National Sample Survey (NSS)

The National Sample Survey Organization (NSSO), set up by the Government of India, conducts rounds of sample surveys to collect socioeconomic data. Each round is earmarked for particular subject coverage. We use the latest five large quinquennial rounds – 38(Jan-Dec 1983), 43(July 1987-June 1988), 50(July 1993-June 1994), 55(July 1999-June 2000) and 61(July 2004-June 2005) on Employment and Unemployment (Schedule 10). The survey covers the whole country except for a few remote and inaccessible pockets. The NSS follows multi-stage stratified sampling with villages or urban blocks as first stage units (FSU) and households as ultimate stage units. The field work in each round is conducted in several sub-rounds throughout the year so that seasonality is minimized. The sampling frame for the first stage unit is the list of villages (rural sector) or the NSS Urban Frame Survey blocks (urban sector) from the latest available census. We describe the broad outline of sample design – stratification, allocation and selection of sample units - with a caveat that the details have changed from round to round.

The whole country is divided politically into states and union territories, and each state is further divided into districts for administrative purpose. The NSSO also constructs regions by grouping contiguous districts within a state which are similar in population density and crop pattern for the sampling purpose. Two different stratification methods are used for rural and urban sector in each state. In the rural sector, each district is generally counted as a separate stratum (populous districts are split into two or more strata) whereas in the urban sector, strata are formed within the NSS region based on population size of cities. For example, all towns with population less than 50,000 in a region will form stratum 1 and so on. In the 61st round, the stratification method was changed substantially. For this round, each district is divided into two basic strata – rural and urban. Then the rural and urban strata are further divided into sub-strata.

The total sample size of first stage unit (villages/urban blocks) is allocated to the states and union territories in proportion to population. The subsequent allocations to rural and urban sector and at stratum level within a state are based on population size as well. In rural sectors, sample FSUs are selected with probability proportional to population from each stratum (sub-stratum for

61st round). In urban sectors, they are selected by simple random sampling without replacement in 38th and 61st round and circular systematic sampling with equal probability in the 43rd, 50th and 55th round. Within each stratum (sub-stratum for 61st round), samples are drawn in the form of two independent sub-samples for both rural and urban sectors. Once the FSUs are randomly drawn, the large FSUs are subdivided into certain number of parts (hamlet-group/sub-block) with approximately equal population and one of them selected randomly for listing of households. Complex second stage stratification based on "means of livelihood class" is implemented to select households randomly from the sample frame of households in each FSU (or hamlet-group/sub-block).

As the sample design changes over the rounds, estimation without considering the complex design may be misleading. The NSSO supplies household level multipliers with the unit record data for each round to help minimize estimation errors on the part of researchers. The questionnaire collects demographic details like age, sex, marital status, education, etc. and information about occupation, industry, activity, time disposition in reference week, wage, etc. of household members. It also collects monthly total household expenditure along with other household level characteristics.

The data are given in fixed format text files with a list of variable names and byte positions. We have checked the validity of our data extraction process by comparing the statistics on a number of the variables with numbers reported in published works by other authors. However, there is some miscoding which is typical for any survey data and we tried our best to clean it. Other notable changes over the rounds are formation of new states, deletion of the social group called "Neo-Buddhist" and formation of new social group called "Other Backward Class" or "OBC" (see below), and changes in coding for education, enrolment in educational institution, activity status and industry. We recoded all these changes to make it uniform and consistent over time.

A.2 Sample Selection

We drop all households for which we have no information on social group or whose social group is miscoded (3/ 120706 households in 38th round, 43/ 129060 households in 43rd round, none for 50th and 55th rounds (115409 and 120386 households, respectively), and 86/124680 households for 61st round are dropped). The classification of Scheduled Castes (SC) and Scheduled Tribes (ST) groups remain unchanged over the rounds. However, there is a new classification of "Other Backward Classes" (OBC) from the 55th round while the "Neo-Buddhist" classification was discontinued from the 50th round. We club these groups with non-SC/ST so that the scheduled caste and scheduled tribe groups (SC/ST) remain uniform throughout the period.

In our data work, we only consider individuals that report their 3-digit occupation code and education attainment level. Occupation codes are drawn from the National Classification of Occupation (NCO) – 1968. We use the "usual" occupation code reported by an individual for the usual principal activity over the previous year (relative to the survey year). The dataset does not contain information on the years of schooling for the individuals. Instead it includes information on general education categories given as (i) not literate -01, literate without formal schooling: **EGS/ NFEC/ AEC -02, TLC** -03, others -04; (ii) literate: below primary -05, primary -06, middle -07, secondary -08, higher secondary -10, diploma/certificate course -11, graduate -12, postgraduate and above -13. We aggregate those into five similarly sized groups as discussed in the main text. We are also interested in studying the patterns of industry employment for different social groups. We employ 4-digit National Industry Classification (NIC) – 1998 industry code that is reported for each individual over the previous year (relative to the survey year).

In our analysis we dedicate a lot of attention to studying wage dynamics. NSS only reports wages from activities undertaken by an individual over the previous week (relative to the survey week). Household members can undertake more than one activity in the reference week. For each activity we know the "weekly" occupation code, number of days spent working in that activity, and wage received from it. We identify the main activity for the individual as the one in which he spent maximum number of days in a week. If there are more than one activities with equal days worked, we consider the one with paid employment (wage is not zero or missing). Workers sometimes change the occupation due to seasonality or for other reasons. To minimize the effect of transitory occupations, we only consider wages for which the weekly occupation code coincides with usual occupation (one year reference). We calculate the daily wage by dividing total wage paid in that activity over the past week by days spent in that activity.

Lastly, we identify full time workers in our dataset. We assume that an individual is a full time worker if he is employed (based on daily status code) for at least two and half days combined in all activities during the reference week.[20] We drop observations if total number of days worked in the reference week is more than seven.

To summarize, our working sample imposes the following restrictions on the data:
1) The overall sub-sample includes all households with a male head of household in the 16-65 age group with at least one other directly related male member of a younger generation (son or grandson) also in the 16-65 age group, where neither is enrolled in an educational institution, both

[20] Based on daily status code we can classify all individuals into employed, unemployed and not in labor force.

have education and occupation information and are working full-time. Within included households, we only consider the head of the household and his direct male descendants.
2) The wage sub-sample includes only those households from the overall sample for which wage data for head and at least one of his descendants are non-missing and non-zero, and for whom the occupation code reported for the last year (relative to the survey year) coincided with the occupation code for which wages over the last week (relative to the survey week) were collected.

The working sample is further subdivided into two generational groups – children and parents. Only household heads are considered as parents in our analysis. Any members from younger generations are considered as children (therefore it includes grandchildren).

A.3 Occupation and Industry Categories

Table A1 summarizes the one-digit occupation categories in our dataset and presents our grouping of these categories into the Occ 1 - "white collar", Occ 2 - "blue collar" and Occ 3 - "agriculture" groups that we used in the text.

Table A1: Occupation categories

Occupation code	Occupation description	Group
0-1	Professional, technical and related workers	Occ 1
2	Administrative, executive and managerial workers	Occ 1
3	Clerical and related workers	Occ 1
4	Sales workers	Occ 2
5	Service workers	Occ 2
6	Farmers, fishermen, hunters, loggers and related workers	Occ 3
7-8-9	Production and related workers, transport equipment operators and labourers	Occ 2

Table A2 summarizes one-digit industry codes in our dataset. In the presentation in the text we group these codes further into three broad industry categories: Ind 1 refers to Agriculture, Hunting, Forestry and Fishing; Ind 2 collects all tradable industries; while Ind 3 refers to all non-tradable industries. These groupings are detailed in Table A2.

B Intergenerational education mobility

Table A3 reports estimation results from a probit regression of education switches in education attainments of children relative to their parents, given by (3.1).

Table A4 presents average probabilities of education improvements (panel (a)) and education reductions (panel (b)) for the overall sample and separately for non-SC/STs and SC/STs over

Table A2: Industry categories

Industry code	Industry description	Group
A	Agriculture, Hunting and Forestry	Ind 1
B	Fishing	Ind 1
C	Mining and Quarrying	Ind 2
D	Manufacturing	Ind 2
E	Electricity, Gas and Water Supply	Ind 3
F	Construction	Ind 3
G	Wholesale and Retail Trade; Repair of Motor Vehicles, motorcycles and personal and household goods	Ind 3
H	Hotels and Restaurants	Ind 3
I	Transport, Storage and Communications	Ind 3
J	Financial Intermediation	Ind 3
K	Real Estate, Renting and Business Activities	Ind 3
L	Public Administration and Defence; Compulsory Social Security	Ind 3
M	Education	Ind 3
N	Health and Social Work	Ind 3
O	Other Community, Social and Personal Service Activities	Ind 3
P	Private Households with Employed Persons	Ind 3
Q	Extra Territorial Organizations and Bodies	Ind 3

different survey rounds. These probabilities were estimated following the procedure we used to obtain average conditional probabilities of education switches, which is described in details in the main text.

Table A5 reports estimation results from a probit regression for education switches associated with improvements in education attainments of children relative to their parents.

Table A6 reports estimation results from a probit regression for education switches associated with reductions in education attainments of children relative to their parents.

C Intergenerational occupation mobility

Table A7 reports estimation results from a probit regression (3.2) for occupation switches at the 3-digit level.

D Intergenerational industry mobility

Table A8 reports estimation results from a probit regression (3.2) for industry switches at the 4-digit level.

E Intergenerational income elasticity

The following tables report the results of the intergenerational wage regressions (3.3). The top panel gives the estimates for the OLS regression while the bottom panel reports the IV estimates.

Table A3: Intergenerational education switches: Estimation results

	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
age	0.0737***	0.0550***	0.0220**	0.0014	-0.0104
	(0.0130)	(0.0109)	(0.0123)	(0.0133)	(0.0152)
age sqr	-0.0012***	-0.0010***	-0.0004**	-0.0001	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
1-SC/ST, 0-non SC/ST	-0.3693***	-0.3659***	-0.2786***	-0.1787***	-0.0314
	(0.0258)	(0.0234)	(0.0238)	(0.0255)	(0.0282)
hh size	0.0112***	0.0130***	0.0123***	0.0030	0.0126***
	(0.0028)	(0.0027)	(0.0030)	(0.0031)	(0.0034)
1-rural, 0-urban	-0.1782***	-0.0916***	0.1033***	0.1983***	0.2430***
	(0.0244)	(0.0225)	(0.0232)	(0.0255)	(0.0280)
1-muslim, 0-other	-0.2660***	-0.2552***	-0.2211***	-0.0614**	0.0278
	(0.0323)	(0.0301)	(0.0320)	(0.0356)	(0.0362)
quota SC/ST	-0.0064***	-0.0068***	-0.0067***	-0.0008	0.0084***
	(0.0022)	(0.0017)	(0.0018)	(0.0019)	(0.0019)
N	24119	28149	25716	25994	29098

Notes: This table reports estimation results from the probit regression (3.1) in which the dependent variable is whether or not there was an intergenerational education switch. Columns (i)-(v) refer to the survey round. Panel "Changes" reports change in SC/ST dummy coefficient over the successive decades and the entire sample period. Standard errors are in parentheses.

Table A4: Intergenerational education improvements and reductions

	(a) education improvements			**(b) education reductions**		
	overall	non-SC/STs	SC/STs	overall	non-SC/STs	SC/STs
1983	0.4557	0.4874	0.3552	0.0863	0.0915	0.0700
	(0.0008)	(0.0008)	(0.0011)	(0.0003)	(0.0003)	(0.0004)
1987-88	0.4684	0.5014	0.3676	0.0915	0.0949	0.0811
	(0.0007)	(0.0007)	(0.001)	(0.0002)	(0.0003)	(0.0004)
1993-94	0.5234	0.5449	0.4621	0.0827	0.0885	0.0664
	(0.0006)	(0.0006)	(0.0008)	(0.0002)	(0.0003)	(0.0004)
1999-00	0.5363	0.5488	0.5035	0.0900	0.0951	0.0767
	(0.0006)	(0.0007)	(0.0012)	(0.0003)	(0.0003)	(0.0004)
2004-05	0.5806	0.5779	0.5880	0.0884	0.0921	0.0785
	(0.0006)	(0.0007)	(0.0011)	(0.0003)	(0.0004)	(0.0005)

Notes: This table presents average probabilities of education improvements (Panel (a)) and education reductions (Panel (b)) for the overall sample and separately for SC/STs and non-SC/STs. These probabilities were estimated using equation (3.1), except we used a binary variable denoting education improvements or education reductions as the left-hand-side variable. Standard errors are in parenthesis.

Table A5: Intergenerational education improvements: Estimation results

	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
age	0.1132	0.0946	0.0639	0.0380	0.0089
	(0.0136)	(0.0115)	(0.0122)	(0.0129)	(0.0149)
age sqr	-0.0017	-0.0015	-0.0011	-0.0006	-0.0001
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
1-SC/ST, 0-non SC/ST	-0.3356	-0.3402	-0.2111	-0.1220	0.0086
	(0.0262)	(0.0234)	(0.0238)	(0.0251)	(0.0273)
hh size	0.0049	0.0057	0.0055	-0.0006	0.0033
	(0.0028)	(0.0026)	(0.0029)	(0.0030)	(0.0033)
1-rural, 0-urban	-0.0882	-0.0348	0.1506	0.2758	0.3137
	(0.0241)	(0.0223)	(0.0230)	(0.0251)	(0.0280)
1-muslim, 0-other	-0.3344	-0.3114	-0.2401	-0.0532	0.0077
	(0.0326)	(0.0301)	(0.0318)	(0.0364)	(0.0352)
quota SC/ST	-0.0069	-0.0099	-0.0073	0.0011	0.0092
	(0.0022)	(0.0017)	(0.0018)	(0.0019)	(0.0019)
N	24119	28149	25716	25994	29098

Notes: This table reports estimation results from the probit regression (3.1) in which the dependent variable is whether or not there was an intergenerational education improvement. Columns (i)-(v) refer to the survey round. Standard errors are in parentheses.

Table A6: Intergenerational education reductions: Estimation results

	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
age	-0.0866	-0.0805	-0.0983	-0.0796	-0.0352
	(0.0163)	(0.0142)	(0.0181)	(0.0202)	(0.0242)
age sqr	0.0012	0.0010	0.0014	0.0011	0.0002
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0005)
1-SC/ST, 0-non SC/ST	-0.1121	-0.0784	-0.1699	-0.1314	-0.0924
	(0.0360)	(0.0352)	(0.0357)	(0.0375)	(0.0392)
hh size	0.0151	0.0164	0.0168	0.0090	0.0196
	(0.0039)	(0.0037)	(0.0039)	(0.0041)	(0.0045)
1-rural, 0-urban	-0.2045	-0.1266	-0.1284	-0.1960	-0.1954
	(0.0332)	(0.0291)	(0.0314)	(0.0352)	(0.0389)
1-muslim, 0-other	0.1315	0.1184	0.0494	-0.0186	0.0405
	(0.0418)	(0.0389)	(0.0444)	(0.0451)	(0.0461)
quota SC/ST	0.0005	0.0077	0.0022	-0.0042	-0.0025
	(0.0027)	(0.0023)	(0.0025)	(0.0027)	(0.0026)
N	24119	28149	25716	25994	29098

Notes: This table reports estimation results from the probit regression (3.1) in which the dependent variable is whether or not there was an intergenerational education reduction. Columns (i)-(v) refer to the survey round. Standard errors are in parentheses.

Table A7: Intergenerational occupation switches: Estimation results

	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
age	-0.0168	-0.0353	-0.0053	0.0057	0.0040
	(0.0140)	(0.0114)	(0.0145)	(0.0145)	(0.0162)
age sqr	0.0001	0.0005	0.0000	-0.0002	-0.0002
	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
1-SC/ST, 0-non SC/ST	0.0798	0.0438	0.1060	0.0970	0.0148
	(0.0283)	(0.0267)	(0.0275)	(0.0293)	(0.0295)
edu-2 dummy	-0.0796	-0.0386	0.0391	0.0990	0.0232
	(0.0402)	(0.0379)	(0.0387)	(0.0472)	(0.0510)
edu-3 dummy	-0.0145	0.0516	0.0214	0.1830	0.0946
	(0.0311)	(0.0305)	(0.0354)	(0.0417)	(0.0440)
edu-4 dummy	-0.0230	0.0548	0.0649	0.2047	0.1244
	(0.0328)	(0.0303)	(0.0331)	(0.0369)	(0.0412)
edu-5 dummy	0.0039	0.0922	0.0218	0.2773	0.1861
	(0.0382)	(0.0338)	(0.0343)	(0.0375)	(0.0423)
hh size	0.0219	0.0142	0.0170	0.0060	0.0168
	(0.0029)	(0.0028)	(0.0033)	(0.0034)	(0.0035)
1-rural, 0-urban	0.0753	0.0932	0.2559	0.1161	0.2266
	(0.0318)	(0.0292)	(0.0302)	(0.0337)	(0.0317)
1-muslim, 0-other	0.0514	-0.0299	0.0026	-0.0005	0.0435
	(0.0359)	(0.0333)	(0.0365)	(0.0401)	(0.0369)
quota SC/ST	-0.0071	-0.0022	-0.0094	-0.0065	0.0002
	(0.0023)	(0.0018)	(0.0020)	(0.0022)	(0.0020)
N	24119	28149	25716	25994	29098

Notes: This table reports estimation results from the probit regression (3.2) in which the dependent variable is whether or not there was an intergenerational occupation switch (at the 3 digit). Columns (i)-(v) refer to the survey round. Standard errors are in parentheses.

Table A8: Intergenerational industry switches: Estimation results

	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
age	-0.0340	-0.0504	-0.0207	-0.0040	-0.0026
	(0.0135)	(0.0118)	(0.0156)	(0.0155)	(0.0167)
age sqr	0.0004	0.0008	0.0003	0.0000	-0.0001
	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
1-SC/ST, 0-non SC/ST	0.0197	0.0330	0.0573	0.0880	0.0301
	(0.0298)	(0.0285)	(0.0292)	(0.0313)	(0.0300)
edu-2 dummy	-0.0272	-0.0021	0.1037	0.1888	-0.0268
	(0.0404)	(0.0416)	(0.0415)	(0.0499)	(0.0529)
edu-3 dummy	0.0353	0.0982	0.1233	0.2068	0.0597
	(0.0328)	(0.0324)	(0.0380)	(0.0453)	(0.0454)
edu-4 dummy	0.0423	0.1146	0.1464	0.2655	0.1263
	(0.0352)	(0.0323)	(0.0355)	(0.0393)	(0.0425)
edu-5 dummy	0.0957	0.1466	0.2311	0.3053	0.1882
	(0.0393)	(0.0356)	(0.0364)	(0.0397)	(0.0434)
hh size	0.0203	0.0096	0.0163	-0.0009	0.0129
	(0.0031)	(0.0031)	(0.0034)	(0.0037)	(0.0034)
1-rural, 0-urban	0.1746	0.1912	0.3161	0.2209	0.2882
	(0.0334)	(0.0304)	(0.0313)	(0.0361)	(0.0334)
1-muslim, 0-other	0.0453	-0.0485	-0.0357	-0.0009	0.0069
	(0.0364)	(0.0347)	(0.0388)	(0.0419)	(0.0385)
quota SC/ST	-0.0102	-0.0051	-0.0077	-0.0088	-0.0012
	(0.0024)	(0.0019)	(0.0021)	(0.0022)	(0.0020)
N	23989	28031	25549	25994	29098

Notes: This table reports estimation results from the probit regression (3.2) in which the dependent variable is whether or not there was an intergenerational industry switch (at the 4 digit). Columns (i)-(v) refer to the survey round. Standard errors are in parentheses.

Table A9: Estimated intergenerational income elasticities, OLS

Non-SC/STs	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
w_{ip}	0.6206***	0.5417***	0.6481***	0.4647***	0.4400***
	(0.0184)	(0.0371)	(0.0267)	(0.0177)	(0.0243)
$\tilde{A}_{ic}$	-0.0045	-0.0190	-0.0052	-0.0146	-0.0019
	(0.0083)	(0.0178)	(0.0150)	(0.0111)	(0.0147)
$\tilde{A}^2_{ic}$	-0.0013*	0.0011***	-0.0033***	-0.0025**	-0.0015
	(0.0008)	(0.0004)	(0.0013)	(0.0012)	(0.0015)
$\tilde{A}^3_{ic}$	0.0000	0.0000	0.0000	0.0000	0.0001
	0.0000	0.0000	0.0000	0.0000	0.0000
A_{ip}	-0.0554	-0.0849	-0.1038***	-0.2228***	-0.2516***
	(0.0513)	(0.0740)	(0.0426)	(0.0753)	(0.1300)
A^2_{ip}	0.0010	0.0014	0.0018***	-0.0049***	0.0047***
	(0.0010)	(0.0016)	(0.0009)	(0.0015)	(0.0025)
A^3_{ip}	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000
$w_{ip} * \tilde{A}_{ic}$	0.0170***	0.0288***	0.0153***	0.0190***	0.0121***
	(0.0033)	(0.0065)	(0.0055)	(0.0038)	(0.0051)
$w_{ip} * \tilde{A}^2_{ic}$	0.0002	-0.0004	0.0002	0.0005*	0.0012***
	(0.0002)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
$w_{ip} * \tilde{A}^3_{ic}$	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000
R-sqr	0.55	0.45	0.54	0.44	0.39
N	4641	2490	5117	5662	5137
SC/STs	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
w_{ip}	0.6473***	0.5398***	0.8379***	0.5581***	0.4359***
	(0.0360)	(0.0695)	(0.0219)	(0.0300)	(0.0340)
$\tilde{A}_{ic}$	-0.0136	0.0260	0.0025	-0.0273	0.0077
	(0.0127)	(0.0412)	(0.0101)	(0.0113)	(0.0143)
$\tilde{A}^2_{ic}$	0.0011	-0.0050	-0.0010	-0.0014	-0.0048
	(0.0010)	(0.0025)	(0.0007)	(0.0009)	(0.0011)
$\tilde{A}^3_{ic}$	0.0000	0.0001	0.0000	0.0000	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
A_{ip}	-0.1449***	-0.1376	0.1194***	0.0079	-0.1149
	(0.0725)	(0.2121)	(0.0619)	(0.0455)	(0.1082)
A^2_{ip}	0.0030***	0.0025	-0.0026**	-0.0002	0.0022
	(0.0015)	(0.0043)	(0.0013)	(0.0010)	(0.0022)
A^3_{ip}	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$w_{ip} * \tilde{A}_{ic}$	0.0159***	0.0022	0.0033	0.0167***	0.0008
	(0.0068)	(0.0182)	(0.0044)	(0.0050)	(0.0058)
$w_{ip} * \tilde{A}^2_{ic}$	0.0001	0.0016	0.0004*	0.0004	0.0018
	(0.0005)	(0.0011)	(0.0003)	(0.0004)	(0.0004)
$w_{ip} * \tilde{A}^3_{ic}$	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R-sqr	0.42	0.46	0.73	0.46	0.34
N	2729	751	2988	3443	3365

Notes: This table presents OLS estimation results from a regression equation (3.3) for five NSS survey rounds ((i)-(iv)). The coefficient estimates for intergenerational income persistence from these regressions are also plotted in Panel (a) of Figure 4. Standard errors are in parenthesis. * p-value≤0.10, ** p-value≤0.05, *** p-value≤0.01.

Table A10: Estimated intergenerational income elasticities, IV

Non-SC/STs	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
w_{ip}	0.7576***	0.6717**	0.7442***	0.6271***	0.5649***
	(0.0347)	(0.1964)	(0.0568)	(0.0358)	(0.0339)
$\tilde{A}_{ic}$	0.0401***	0.0580	0.0368**	0.0206**	0.0103
	(0.0087)	(0.0528)	(0.0141)	(0.0080)	(0.0088)
$\tilde{A}^2_{ic}$	-0.0017	-0.0026	-0.0016	0.0000	0.0008
	(0.0005)	(0.0030)	(0.0009)	(0.0005)	(0.0006)
$\tilde{A}^3_{ic}$	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
A_{ip}	-0.0654***	-0.1071	-0.0659*	-0.0366***	-0.0093
	(0.0177)	(0.1253)	(0.0318)	(0.0195)	(0.0225)
A^2_{ip}	0.0031***	0.0054	0.0035*	0.0001	-0.0024
	(0.0011)	(0.0069)	(0.0020)	(0.0013)	(0.0015)
A^3_{ip}	0.0000	-0.0001	-0.0001	0.0000	0.0001
	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
$w_{ip} * \tilde{A}_{ic}$	0.0324**	0.0537	0.0470***	0.0794***	0.0770***
	(0.0054)	(0.0459)	(0.0076)	(0.0066)	(0.0084)
$w_{ip} * \tilde{A}^2_{ic}$	-0.0009***	-0.0017	-0.0015***	-0.0024***	-0.0019**
	(0.0002)	(0.0011)	(0.0003)	(0.0003)	(0.0003)
$w_{ip} * \tilde{A}^3_{ic}$	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R-sqr	0.51	0.4	0.53	0.4	0.37
N	2447	1312	2720	3016	2750
SC/STs	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
w_{ip}	0.8712***	0.6732***	0.8683***	0.7016***	0.4547***
	(0.0578)	(0.0803)	(0.0593)	(0.0483)	(0.0506)
$\tilde{A}_{ic}$	0.0376***	0.0277	0.0043	0.0138	-0.0078
	(0.0139)	(0.0197)	(0.0074)	(0.0095)	(0.0115)
$\tilde{A}^2_{ic}$	-0.0022***	-0.0008	0.0003	0.0001	0.0025***
	(0.0010)	(0.0012)	(0.0006)	(0.0006)	(0.0008)
$\tilde{A}^3_{ic}$	0.0000	0.0000	0.0000	0.0000	-0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
A_{ip}	-0.0548***	-0.0347	0.0006	-0.0214	0.0291
	(0.0264)	(0.0435)	(0.0161)	(0.0219)	(0.0292)
A^2_{ip}	0.0033*	0.0009	-0.0009	-0.0006	-0.0064***
	(0.0019)	(0.0027)	(0.0014)	(0.0014)	(0.0021)
A^3_{ip}	-0.0001	0.0000	0.0000	0.0000	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$w_{ip} * \tilde{A}_{ic}$	0.0058	0.0307***	0.0164*	0.0403***	0.0753***
	(0.0102)	(0.0182)	(0.0092)	(0.0086)	(0.0098)
$w_{ip} * \tilde{A}^2_{ic}$	-0.0001	-0.0007	-0.0005	-0.0009***	-0.0015***
	(0.0003)	(0.0007)	(0.0003)	(0.0003)	(0.0003)
$w_{ip} * \tilde{A}^3_{ic}$	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R-sqr	0.38	0.4	0.73	0.43	0.34
N	2729	734	2967	3443	3365

Notes: This table presents IV estimation results from a regression equation (3.3) for five NSS survey rounds ((i)-(iv)). We used household per capital consumption expenditure and household size as instruments for parent's log wage. The coefficient estimates for intergenerational income persistence from these regressions are also plotted in Panel (b) of Figure 4. Standard errors are in parenthesis. * p-value≤0.10, ** p-value≤0.05, *** p-value≤0.01.