Castes and Labor Mobility*

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

September 2010

Abstract

How have the poorest sections of society in India responded to the rapid changes in the Indian economy over the past 30 years? We examine this question by contrasting the fortunes 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. We study the period 1983-2005 using household survey data from successive rounds of the National Sample Survey. The key message is that education attainment rates and wages have been converging across the two groups while SC/STs have also been switching occupations at increasing rates during this period. Moreover, inter-generational education and income mobility rates of SC/STs have converged to non-SC/ST levels. Clearly, the last twenty years of major structural changes in India have seen a sharp improvement in the relative economic fortunes of these historically disadvantaged social groups. In fact, the median wages of SC/STs relative to non-SC/STs in India have surpassed the median wages of blacks relative to whites in the US.

JEL Classification: J6, R2

Keywords: Intergenerational mobility, wage gaps, castes

*We would like to thank Nicole Fortin, Ashok Kotwal, Thomas Lemieux and seminar participants at UBC and Monash University 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).

"If there is no enthusiasm, life becomes drudgery - a mere burden to be dragged. Nothing can be achieved if there is no enthusiasm. The main reason for this lack of enthusiasm on the part of a man is that an individual loses the hope of getting an opportunity to elevate himself. Hopelessness leads to lack of enthusiasm. The mind in such cases becomes deceased...When is enthusiasm created? When one breathes an atmosphere where one is sure of getting the legitimate reward for one's labor, only then one feels enriched by enthusiasm and inspiration"

"This condition obtains even where there is no slavery in the legal sense. It is found where as in caste system, some persons are forced to carry on the prescribed callings which are not their choice..."

B. R. Ambedkar (Chief architect of the Indian Constitution.)

1 Introduction

The past 25-30 years have been a period of massive changes in the Indian economy. Average annual GDP growth rates have climbed rapidly from the anaemic 3-3.5 percent that characterized the first 35 years since 1947 to between 8 and 10 percent. Accompanying this growth takeoff has been a hastening process of structural transformation of the economy. The agricultural sector, which historically had the largest employment and output share, has rapidly lost ground in both during this period. Such rapid structural changes often deeply affect the lives of people in these economies by redistributing income and economic opportunities from some groups to others. As an example, a particularly emotive issue that has been debated energetically with regard to the Indian experience is the effect of this economic take-off on the fortunes of the poor.

In this paper we study the impact of the rapid transformation of the Indian economy since the 1980s on one such historically disadvantaged group: the Scheduled Castes and Scheduled Tribes (SC/STs). SC/STs were historically economically backward, mostly very poor, concentrated in low-skill (mostly agricultural) occupations and primarily rural. Moreover, they were also subject to centuries of systematic caste-based discrimination both economically and socially. This was so endemic that the constitution of India aggregated these castes into a Schedule of the constitution and provided them with affirmative action cover in both education and public sector employment.

Indeed, this was viewed as a key component of attaining the ultimate policy goal of raising the social and economic mobility of the SC/STs to the levels of the non-SC/STs.

A key goal of the reservations policy was to make it easier for, say, the child of an illiterate SC or ST farm worker living below the poverty line to get educated and find productive employment in a better paying occupation. How have the tectonic changes in India since the early 1980s affected this goal? Has the rapid growth percolated down to the SC/STs in terms of tangible changes in their economic and social conditions? Is the primary reason for the economic deprivation of these underprivileged castes the types of occupations they tend to work in, i.e., do successive generations of SC/STs tend to get stuck in low wage jobs? Alternatively, is the key impediment the lack of education, i.e., do they get stuck in low wage jobs due to the lack of education? Or, is ongoing discrimination in occupations and wages the primary problem facing these groups? This paper attempts to answer some of these questions.

We use data from five successive rounds of the National Sample Survey (NSS) from 1983 to 2004-05 to analyze patterns of occupation choices, education attainment and wages of both SC/ST and non-SC/ST households. We conduct our analysis along two dimensions. First, we contrast the time-series evolution of education, occupation and industry choices and wages of SC/STs with their non-SC/ST counterparts from the same age cohort. We conduct this cohort-level analysis both at an aggregated generation level of parents and children as well as at a more disaggregated level of five different age cohorts.[1] Second, 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.

Our analysis yields four main results. First, while SC/ST households are, on average, less educated than their non-SC/ST counterparts throughout the sample period and across cohorts, the education attainment levels of SC/STs have been converging toward the level of their non-SC/ST cohort. This trend is particularly pronounced for SC/ST children. Moreover, the trend towards education convergence of the two groups emerges both in rural and urban sectors but is sharper in urban areas. The trend also shows up clearly across occupations.

Second, there have been similar compositional changes in the occupational distributions of SC/STs and non-SC/STs between 1983 and 2004-05. Children and parents of both SC/STs and

[1] We look at aggregated cohorts of parents and children to set the stage for the intergenerational mobility analysis.

non-SC/STs have been moving out of low skill agrarian occupations into relatively higher skill occupations. However, these changes have occurred slightly faster for SC/STs. As a result, the occupation distribution of the two groups appear to be converging during this period. We also study trends in industry mobility of the two groups and find that the results for industry mobility are broadly similar to those for occupation mobility.

Third, we find a clear trend of convergence of the relative wage of the two groups towards one, i.e., the median wage premium of non-SC/STs relative to SC/STs has declined systematically from 25 percent in 1983 to 6 percent 2004-05. The trend is particularly strong for children where the wage premium of non-SC/STs has declined from 14 percent to approximately zero. For the parent cohort, the non-SC/ST wage premium has declined from 35 percent to 19 percent during this period. This pattern of relative wage convergence also emerges in mean wages and across more disaggregated age cohorts. Using conditional wage regressions we find that relative education attainment levels account for over 50 percent of the mean wage gaps. To put these wage gaps in perspective, the median white male to black male wage premium in the US has hovered stubbornly between 25 and 40 percent over the past 35 years, which makes the SC/ST relative wage behavior in India even more striking.

Fourth, we find that *intergenerational mobility* of SC/STs 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 41 percent in 1983 but rose sharply to 66 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 60 percent to 40 percent for SC/ST households and from 53 to 39 percent for non-SC/ST households. Clearly, the intergenerational *income mobility* rates have also converged. Lastly, 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.

In summary, these results suggest some uplifting answers to the questions we set out to answer. Over the last 30 years SC/STs have sharply narrowed both their education and wage gaps relative to non-SC/STs. The fact that these trends are sharpest amongst younger age-cohorts and amongst urban households suggests that the overall statistics are likely to improve even more sharply in their favor in the coming years as these cohorts become older and as the country becomes more urbanized. Moreover, children from the historically disadvantaged SC/ST households are increasingly raising their education attainments levels, switching occupations and improving their income positions relative to their parents. Crucially, intergenerational income and educational mobility of SC/ST households has been rising faster than for non-SC/STs. Overall, we conclude 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. This period of rapid structural changes appears to have been very beneficial for SC/STs who have used this period to rapidly narrow their huge historical economic disparities with non-SC/STs.

To the best of our knowledge, our's is the first study to jointly analyze caste differences in education, occupation, industry and wage outcomes in a single study, track the time series evolution of these outcomes, and do so using data that covers the entire country. It is worth reiterating that we do this using the NSS data which has the broadest coverage for India both spatially and over time.

There exists a large literature which has investigated the existence and extent of labor market discrimination in India. Amongst others, Banerjee and Knight (1985) and Madheswaran and Attewell (2007) have studied the extent of wage discrimination faced by SC/STs in the urban Indian labor market. Borooah (2005) has studied the extent of discrimination in employment in the urban labor market. Ito (2009) studies both wage and employment discrimination simultaneously by examining data from two Indian states – Bihar and Uttar Pradesh. Our study differs from these in that we examine the data for all states and for both rural and urban areas. Moreover, as opposed to most of these studies, our study controls for the presence of occupation and industry effects on wage outcomes. Lastly, by using data for five rounds of the National Sample Survey of households we are also able to provide a time series perspective on the evolution of SC/ST fortunes in India, a feature that other studies have typically not examined.

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 other 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 occupation, industry, and income. 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.[2]

In the next section we describe the data and our constructed measures as well as some summary statistics. Section 3 contrasts SC/STs with their non-SC/ST cohorts in terms of the evolution of the distributions of education attainment rates, occupations, industry of employment and wages. Section 4 presents and discusses the evidence on intergenerational mobility, while the last section concludes.

2 The Data

Our data comes from the National Sample Survey (NSS) of India and its various rounds. In particular, we use the NSS Rounds 38 (1983), 43 (1987-88), 50 (1993-94), 55 (1999-2000) and 61 (2004-05). Our working sample consists of all households heads and their children who provided their 3-digit occupation code information. We further restrict the sample by considering only those households whose head's age is between 16 and 65. We consider only male-led households and male children, but do not restrict the birth order of the children.[3] Our sample is further restricted to 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. This selection leaves us with a sample

[2] 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.

[3] We also consider a broader sample in which we do not retrict 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).

of around 43,000-51,000 individuals, depending on the survey round.[4] Details regarding the dataset are contained in the Appendix.

Data on wages are more limited. The sub-sample with complete wage and occupation data for the head of household and working children consists of, on average across rounds, about 7,000-9,000 individuals which is considerably smaller than our basic sub-sample but large enough to facilitate formal analysis. Wages are obtained as the daily wage/salaried income received for the work done by respondents during the previous week. Wages can be paid in cash or kind, where the latter are evaluated by the current retail prices. We convert wages into real terms using state-level poverty lines that differ for rural and urban sectors. We express all wages in 1983 Maharashtra prices.

We conduct all our data work on the restricted sample of households with both a full-time working male head and at least one male child in the admissible age group working full-time. The restriction is necessitated by our interest in examining inter-generational mobility trends. We choose to work with the restricted sample for our intra-generational exercises as well in order to retain comparability of the samples and the results. If we do not restrict the sample to households with working heads and at least one working child, the sample size grows to between 136,000-151,000 individuals, depending on the round, if we do not include the wage data. If we also include the wage data then the sample size in this larger sample declines to an average of about 50,000 across rounds. As we show later in this section, the convergence results on education and wages reported here for SC/STs and non-SC/STs are robust to these alternative sample restrictions.

Our education variable contains 5 categories: 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 occupation choices of individuals. In particular, we know the three-digit occupation code associated with the work that each individual performed over the

[4] We conduct all our data work on the restricted sample of households with both a full-time working male head and at least one male child in the admissible age group working full-time. The restriction is necessitated by our interest in examining inter-generational mobility trends. We choose to work with the restricted sample for our intra-generational exercises as well in order to retain comparability of the samples and the results. If we do not restrict the sample to households with working head's and at least one working child, the sample size grows to around 150,000 individuals on average across the five rounds if we do not include the wage data. If we also include the wage data then the sample size averages about 55,000 across rounds. As we show in the Data Appendix, the convergence results on education and wages reported here for SC/STs and non-SC/STs are robust to these alternative sample restrictions.

last year (relative to the survey year). We use only those individuals for whom the occupation code reported for the last year coincided with the occupation code for which wages over the last week were collected (relative to the survey week). Our dataset also contains information on the four-digit industry of employment for each individual. For those individuals reporting multiple activities during the reference period, we consider the activity in which she spent maximum days in a week. If there are more than one activities with equal days, we consider the one with paid employment (wage is not zero or missing).

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 "Overall" refers to the full sample, while the "Non-SC/ST" and "SC/ST" panels refer to the corresponding 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, although this number is higher for SC/ST households, of which 88-89 percent 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.

3 Intragenerational Cohort Comparison: How Have the Scheduled Castes Fared?

We start our analysis by comparing SC/STs with non-SC/STs across age and generational cohorts. As mentioned before, we focus on working males in the age group 16-65. We construct cohorts in two ways. First, for every round of the census we split our sample into five broad age cohorts: 16-25, 26-35, 36-45, 46-55 and 56-65. Second, in each round we split the sample into household heads and working male children in co-residence with a working male household head. We call these

Table 1: Sample summary statistics

	(a) children			(b) parents				
Overall	age	edu	%married	age	edu	%married	%rural	hh size
1983	22.83	2.58	0.53	51.67	1.79	0.92	0.81	8.31
	(0.04)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.03)
1987-88	23.13	2.69	0.53	51.65	1.88	0.92	0.83	8.05
	(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	7.59
	(0.04)	(0.01)	(0.00)	(0.06)	(0.01)	(0.00)	(0.00)	(0.02)
1999-00	23.43	3.21	0.46	51.59	2.20	0.94	0.81	7.93
	(0.05)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.03)
2004-05	23.76	3.40	0.46	51.63	2.34	0.94	0.80	7.51
	(0.05)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.03)
Non-SC/ST								
1983	23.00	2.78	0.52	52.04	1.93	0.92	0.79	8.4
	(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	8.11
	(0.05)	(0.01)	(0.00)	(0.08)	(0.01)	(0.00)	(0.00)	(0.03)
1993-94	23.36	3.17	0.47	52.10	2.19	0.94	0.79	7.65
	(0.05)	(0.01)	(0.00)	(0.07)	(0.01)	(0.00)	(0.00)	(0.03)
1999-00	23.68	3.42	0.46	52.00	2.41	0.95	0.78	7.95
	(0.05)	(0.01)	(0.00)	(0.08)	(0.02)	(0.00)	(0.00)	(0.04)
2004-05	24.03	3.55	0.46	52.01	2.53	0.95	0.77	7.48
	(0.06)	(0.01)	(0.01)	(0.08)	(0.02)	(0.00)	(0.00)	(0.04)
SC/ST								
1983	22.30	1.95	0.56	50.59	1.38	0.92	0.89	8.07
	(0.08)	(0.02)	(0.01)	(0.13)	(0.01)	(0.01)	(0.01)	(0.05)
1987-88	22.63	2.06	0.56	50.72	1.45	0.91	0.90	7.87
	(0.08)	(0.02)	(0.01)	(0.12)	(0.01)	(0.00)	(0.00)	(0.05)
1993-94	22.61	2.40	0.49	50.92	1.54	0.92	0.90	7.44
	(0.08)	(0.02)	(0.01)	(0.13)	(0.02)	(0.00)	(0.00)	(0.05)
1999-00	22.77	2.67	0.46	50.59	1.71	0.94	0.88	7.90
	(0.09)	(0.02)	(0.01)	(0.13)	(0.02)	(0.00)	(0.01)	(0.05)
2004-05	23.04	2.98	0.45	50.65	1.87	0.94	0.87	7.58
	(0.09)	(0.03)	(0.01)	(0.14)	(0.02)	(0.00)	(0.01)	(0.06)

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.

generational cohorts "Parents" and "Children", respectively. For each age and generational cohort we compute the occupation distribution, the average education attainment level and the average daily wage earned for the entire group as well as for SC/STs and non-SC/STs separately. Issues of particular interest to us are: (a) whether the education attainment levels of SC/ST children and parents are converging to the levels of their non-SC/ST cohorts? (b) whether their occupation choices are converging over time; and (c) whether wages of SC/STs are converging to non-SC/ST levels.

A few notes on our generational cohort classification are in order. First, we refer to household heads as parents. In a literal sense household heads are not always the parents of younger working members in the household since there are a few households with a grandparent as the head of a household that also contains his working children and grandchildren. More generally, our termi-

nology is meant as a stand-in for parent-figures. Second, since we evaluate the performances of parents and children in successive rounds of the census, there will definitely be cases where children in one round become household heads, and therefore "parents", in later rounds. However, across the different census rounds the mean age of parents remains relatively stable at around 52 years while the mean age of children remains around 23 years. Thus, all children under the age of 30 in 1983 would still be less than the mean parent age in the last round of the sample in 2004-05. Hence, while there definitely will be some movement of people from one cohort into another over time, this doesn't appear to be a large share of the sample.

Third, we choose to work with age cohorts rather than birth cohorts. This is a deliberate choice which reflects our interest in determining the effects of changing aggregate conditions over time and how they change the incentives of agents over time. Thus, the age-cohort approach allows us to compare and contrast the behavior of 16-25 year olds in 2004-05 with 16-25 year olds in 1983. If the behavior is deifferent then it would indicate that the incentives underlying the choices being made by this age cohort have changed over this period. The alternative of examining birth cohorts and tracking them over time makes it harder to make this deduction since some of their changes over time would also reflect the ageing process.

3.1 Education Attainment

We start with the record on education attainment rates. Panel (a) of Table 2 shows the average education attainment level of the overall population as well as those for working children and parents separately. Both generational groups increased their education attainment levels over the sample period. Panel (b) of Table 2 shows the relative education gap between non-SC/STs and SC/STs, computed as the ratio of their corresponding education attainments. In 1983, non-SC/STs had three quarters of a category more education than SC/STs.[5] However, over the sample period, there is a clear trend towards convergence in education levels of SC/STs toward their non-SC/ST counterparts. This trend is particularly pronounced for the cohort of children. While the difference in 1983 was almost one full education category, by 2004-05 this had narrowed to half a category.[6] Both groups increased their education attainment levels over the period with the SC/STs rising

[5] The level of education attainments for SC/STs in 1983 was 1.68, while for non-SC/STs it was 2.41.

[6] In 1983, education attainment levels of SC/ST and non-SC/ST children were 1.93 and 2.77, respectively. By 2004-05 these levels have increased to 2.97 and 3.55, respectively.

faster.

Table 2: Education attainment levels and gaps

	(a). Levels			(b). Gaps		
	overall	children	parents	overall	children	parents
1983	2.23	2.58	1.79	1.43	1.42	1.40
	(0.01)	(0.01)	(0.01)			
1987-88	2.33	2.69	1.88	1.41	1.40	1.40
	(0.01)	(0.01)	(0.01)			
1993-94	2.55	2.97	2.01	1.36	1.32	1.42
	(0.01)	(0.01)	(0.01)			
1999-00	2.77	3.21	2.20	1.33	1.28	1.41
	(0.01)	(0.01)	(0.01)			
2004-05	2.94	3.40	2.34	1.25	1.19	1.35
	(0.01)	(0.01)	(0.01)			

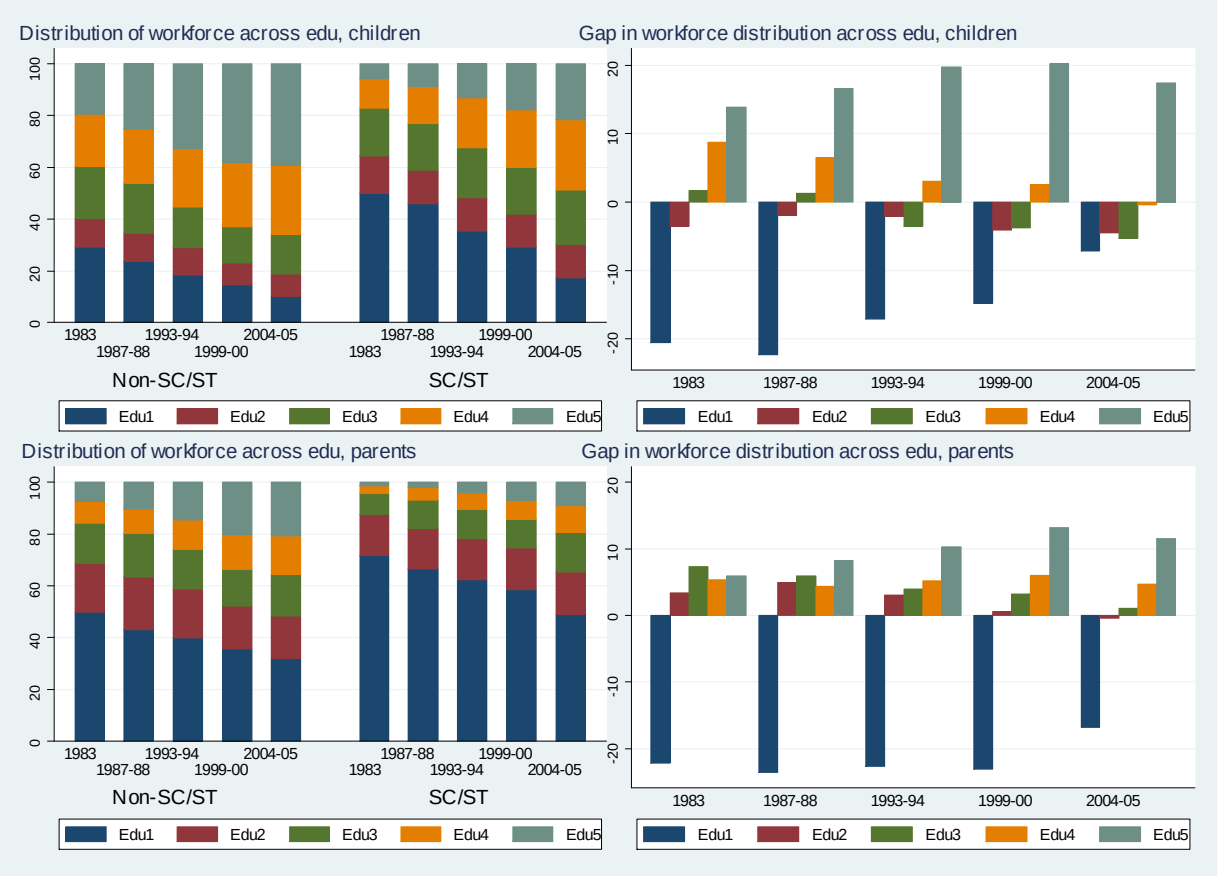
Notes: This table presents the average education attainment levels for our overall benchmark sample and separately for two generational groups – parents and children. Gaps refer to the ratio of average education attainment levels of non-SC/STs to SC/STs for the same three groups. The reported statistics are obtained for each NSS survey round which is shown in the first column. Standard errors are in parenthesis.

The picture is a bit different for parents. First, the increase in average education levels has been more tepid for parents relative to their children. Second, there is no trend toward convergence in the average levels across the parents: SC/STs parents start and end the sample period with about a half category lower education level than non-SC/STs.

A related issue of interest is the distribution of parents and children across the five education categories. In particular, is the change in the average attainment level due to more illiterates beginning to go to primary school or is it primarily due to more people going on to middle school or higher? We answer this question using Figure 1.

Panels (a) and (b) of Figure 1 show the distribution of the workforce across education categories and the corresponding non-SC/ST–SC/ST differences for children and parents respectively. The top graph of Panel (a) shows the distribution of non-SC/ST children across the five education categories (left set of bars) and the corresponding distribution of SC/ST children (right set of bars). It is clear that SC/ST children are systematically less educated than their non-SC/ST counterparts. The difference is most glaring in the lowest and highest categories. In category 1 (the illiterate groups) SC/STs are hugely over-represented while in category 5 (secondary education or above) they are strongly under-represented. The scale of the lack of education in India, both in general

Figure 1: Education distribution of children and parents



(a) (b)

Notes: Panel (a) of this figure presents the distribution of workforce of children and parents across five education categories across different NSS rounds. The left set of bars on each figure refers to non-SC/STs, while the right set is for SC/STs. Panel (b) presents absolute gaps in the distribution of non-SC/STs relative to SC/STs across five education categories. The gaps are also reported for children and parents. See the text for the description of how education categories are defined (category 1 is the lowest education level - illiterate).

and amongst SC/STs, is probably best summarized by the fact that as recently as in 1983, over 65 percent of SC/ST children were either illiterate or had below primary level education while the corresponding number for non-SC/STs was 40 percent. These numbers declined to 31 percent for SC/ST children and 19 percent for non-SC/ST children by 2004-05.

The figure also makes clear that there has been a sustained decrease over time in the share of illiterates amongst both SC/STs and non-SC/STs. Thus, by 2004-05 the proportion of SC/ST children with less than primary education (categories 1 and 2) fell to 33 percent. This was by far the sharpest change amongst all education categories. Correspondingly, the shares of all the other categories increased for SC/ST children though the sharpest increases occurred in education

categories 4 (middle school) and 5 (secondary or above). The pattern for non-SC/STs is broadly similar except for the fact that their sharpest increase occurred in the secondary or above education category 5.

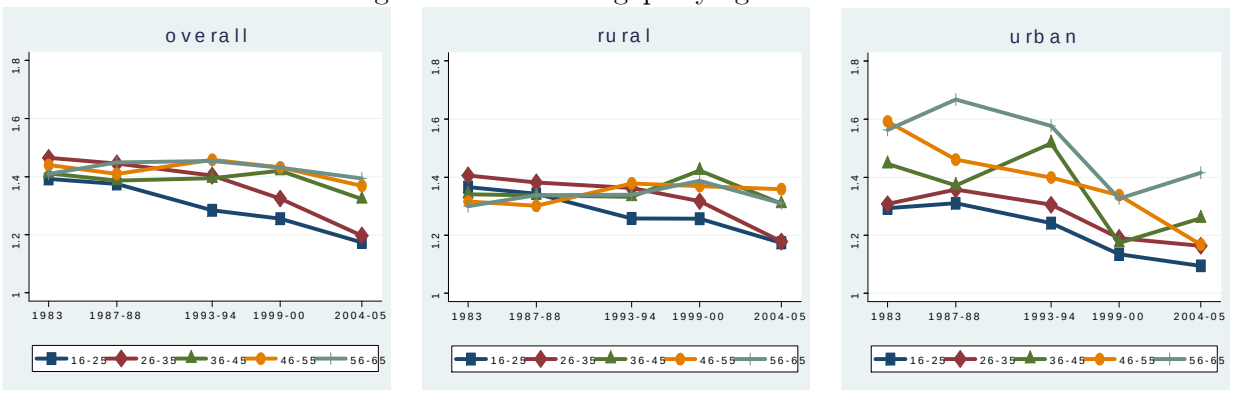
The top graph of Panel (b) of Figure 1 shows the difference between the percentage of non-SC/ST children and SC/ST children within each education category. Thus, the first bar from the left in the top graph of Panel (b) shows that the percentage of all SC/ST children who belonged to education category 1 in 1983 exceeded the corresponding percentage of non-SC/ST children in category 1 in that year by over 20 percentage points. This panel captures, to some extent, the tendency toward convergence of patterns across the two groups. With one exception, the differences in the proportion of children in the other education categories either stayed constant or tended towards convergence for the two groups. The exception was in category 5 (secondary or higher). In 1983 about 19 percent of non-SC/ST children had secondary school or higher levels of education while the number was just around 6 percent for SC/STs. By 2004-05, 39 percent of non-SC/ST children had secondary or higher levels of education while the number for SC/STs had risen to 21 percent. Clearly, for both groups there has been a significant increase in the share of children with secondary or higher education but the difference between the two groups has continued to remain very high.

The bottom panel of Figure 1 shows the same information for parents. There are a few key differences between the education distribution patterns for parents and their children. First, the share of illiterates and those with less than primary education (education categories 1 and 2) is higher for both SC/ST and non-SC/ST parents throughout and declined at a slower rate than that of their children. Thus, in 1983 the combined share of categories 1 and 2 was 69 percent for non-SC/ST parents and 88 percent for SC/ST parents. These numbers fell over time but still remained at a very high 48 percent and 66 percent, respectively, in 2004-05. Second, at the high end of the education distribution the changes have been much more tepid for parents of both groups relative to that of their children. The share of secondary or higher educated parents amongst non-SC/STs rose from 7 percent in 1983 to 21 percent in 2004-05. Correspondingly, the share for SC/ST parents rose from 1 percent to 9 percent. Third, in contrast to the pattern amongst children, there is no clear trend towards homogenization of the two groups in their educational composition. This feature is clearly brought out in Panel (b) of Figure 1, in particular, by the heights of the bars depicting the

shares for education categories 1 and 5.

Do these trends in aggregated generational cohorts mask key differences in the relative movements within more disaggregated age cohorts? To investigate this we compute the education attainment levels of non-SC/STs relative to SC/STs within five age-cohorts for each census round. Figure 2 plots the result. Panel (a) reveals a clear pattern of education convergence across the different age-cohorts over time. Importantly, the convergence appears to be the sharpest amongst the younger cohorts. Given the large concentration of households in rural areas, a related question is whether the trends in education attainment rates are different between rural and urban households. To address this, we split the different age-cohorts into rural and urban households and then plot the education attainment gaps for the two sectors separately in panels (b) and (c) of Figure 2. Three features of the figure are noteworthy. First, the variation in education levels across cohorts is much smaller in rural areas than in urban areas. Second, except for the oldest rural cohorts, education attainment levels have been converging in both rural and urban areas. Third, the convergence rates are, on average, faster in urban areas and for younger cohorts.

Figure 2: Education gaps by age cohorts



(a) (b) (c)

Notes: The figures show the evolution of the relative education gap between non-SC/STs and SC/STs over time for different age groups. Panel (a) presents the results for the overall sample, while panels (b) and (c) report the results for rural and urban households separately.

Overall, the data suggests that there has been a universal trend toward convergence in education levels of SC/STs toward the levels of non-SC/STs. While this trend is common across generations, ages and rural-urban locations, it is sharpest amongst the younger cohorts and in the urban areas. These trend patterns are likely to get sharper in the future as more uneducated parents drop out and more educated children become parents.

3.2 Occupation Choices

We now turn to the occupation choices of the two groups. In order to facilitate ease of presentation, we aggregate the 3-digit occupation codes that individuals report into a one-digit code. This leaves us with ten categories which are then grouped further into three broad occupation categories.[7] Our groupings, while subjective, are based on combining occupations with similar skill requirements. Thus, 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.. The groupings also reflect differences in the returns to skills in the Indian economy: Occ 1 is characterized by the highest mean wage in our sample, followed by Occ 2, and Occ 3.

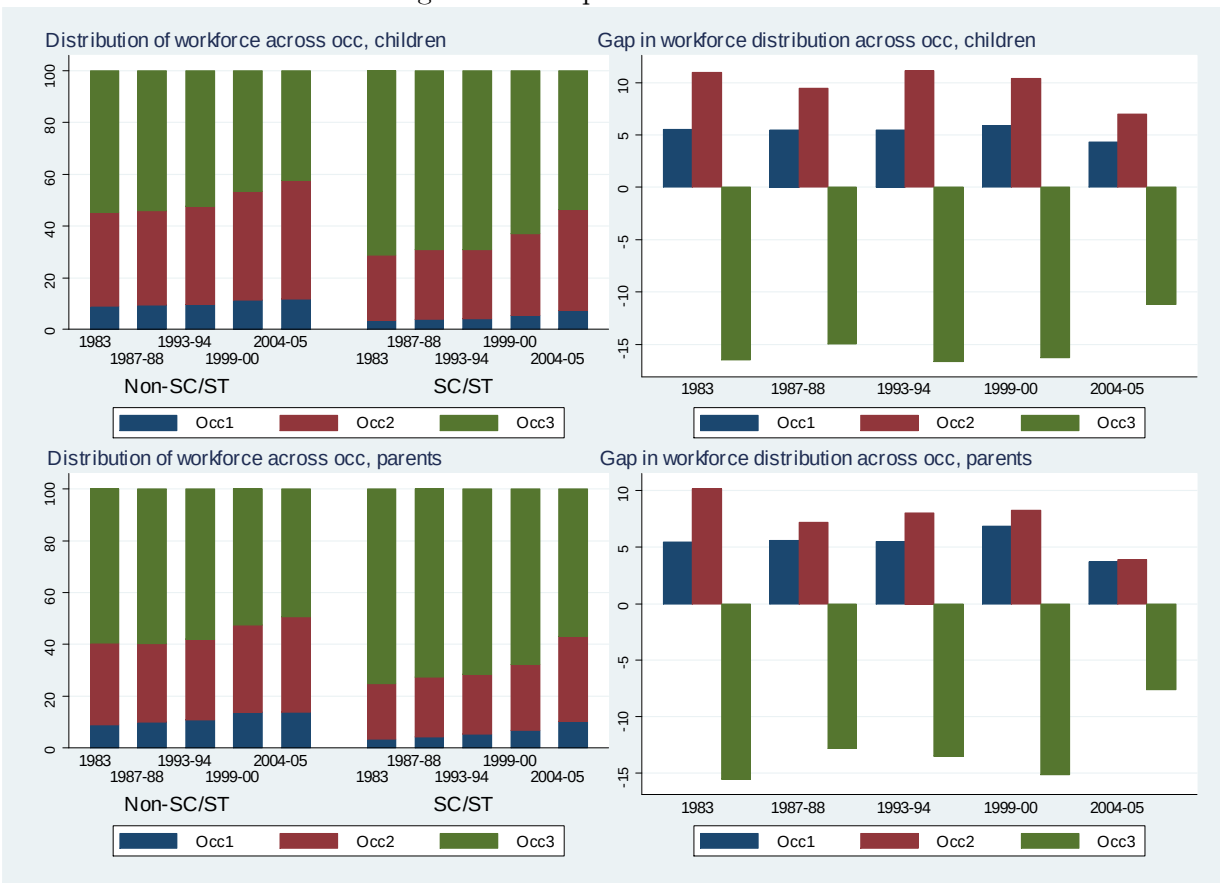
Figure 3 shows the occupation distribution for the working population in our sample, and the differences between non-SC/STs and SC/STs in this distribution. The top panel of Figure 1 refers to children, while the bottom panel refers to parents. There are three features to note. First, there has been a systematic decline in Occ 3 (farming/pastoral activities) between 1983 and 2004-05 across all groups. This decline has been marginally sharper for SC/STs – both children and parents. This reflects the structural transformation at the aggregate level for India wherein there has been a gradual decline in the output and employment share of the agricultural sector. Second, the largest expansion in the employment share has been in Occ 2 which comprises mostly low skill blue collar and service sector jobs. This phenomenon too has been common to both groups. Third, the share of Occ 1 (white collar/high skill) has risen slightly faster for SC/STs than non-SC/STs. This is possibly a sign of increasing mobility for SC/STs and an indicator of possibly faster future improvements.

Since SC/STs were over-represented in Occ 3 and under-represented in Occ 1 and 2 in 1983, the trends in occupation shares of the two groups imply that the overall occupation distribution has become more similar over the sample period for both parents and children, i.e., the distributions have been converging. We should point out though that the occupation distribution appears to be converging marginally faster for parents than for children.

The large changes in the sectoral distribution of occupations along with the previously documented differences in educational attainment levels of SC/STs and non-SC/STs raises a key issue.

[7] See Appendix A.1 for more details on the data.

Figure 3: Occupation distribution



(a) (b)

Notes: Panel (a) of this figure presents the distribution of workforce of children and parents across three occupation categories for different NSS rounds. The left set of bars on each figure refers to non-SC/STs, while the right set is for SC/STs. Panel (b) presents absolute gaps in the distribution of non-SC/STs relative to SC/STs across three occupation categories. The gaps are also reported for children and parents. Occ 1 collects white collar workers, Occ 2 collects blue collar workers, while Occ 3 refers to farmers and other agricultural workers.

Has this period of rapid structural transformation of the Indian economy with a contraction of some traditional sectors and expansion of other sectors favoured one group versus the other? To answer this we look at two aspects of these occupations and the people that work in them. First, how different are these occupation categories in terms of their educational requirements? Second, are there systematic differences in the educational levels of SC/STs and non-SC/STs even within occupations?

Panel (a) of Table 3 shows the average educational attainment level of children and parents working in each of the occupations. Clearly, children working in occupation 1 have the highest education level while occupations 2 and 3 employ children with progressively lesser education, on

average. Moreover, the average level of education in all occupations has risen throughout the period with the sharpest increase in education levels being in blue collar low-skill jobs (Occ 2) and farming/agricultural jobs (Occ 3). The pattern of average education attainment levels of parents across occupations is similar to that for their children – occupation 1 employs parents with the highest education, with occupations 2 and 3 following in that order. Similar to the pattern for children, the average education levels have been rising in all occupations. However, the rise in education levels of parents in occupations 2 and 3 have been much more muted than the corresponding increase for children.

Panel (b) of Table 3 shows the relative gap in the average education levels of non-SC/STs and SC/STs within the same occupation. Two features are noteworthy here. First, SC/ST children are less educated than non-SC/STs of the same cohort even within the same occupation. Second, gaps in education attainment levels are lowest in the high-skill white collar occupations. Third, education gaps amongst children within the same occupations have declined over time in all occupations. The trends in gaps for parents are broadly similar to those we uncovered for children with one key difference. The education gaps between non-SC/ST and SC/ST parents are much larger in white collar, high-skill occupations (occupation 1). This is in sharp contrast to the pattern for children where the difference are the smallest in these occupations.

Finally, we study the more disaggregated age cohorts and document the evolution of education attainments within each occupation. Panel (a) of Figure 4 shows the relative education gap between non-SC/STs and SC/STs working in occupation 1 for different age cohorts. Similarly, panel (b) summarizes the corresponding gap for those employed in occupation 2; and panel (c) for those working in occupation 3.

These results confirm our earlier findings: education attainment levels are converging between non-SC/STs and SC/STs. They are converging faster for younger age cohorts and for higher-skill occupations 1 and 2. Education gaps in occupation 3 have declined for the youngest age cohorts while remaining relatively unchanged or even increasing slightly for the older age cohorts.[8]

[8] One other interesting feature of our data is that the dispersion in education gaps across age cohorts is the highest in occupation 1 with lower dispersions in occupations 2 and 3. This probably reflects the heterogeneity of skills underlying the occupation groups we constructed. Occ 1 combines a variety of high-skill occupations that can lead to more heterogeneity in education gaps. Such skill heterogeneity is lower in occupations 2 and 3.

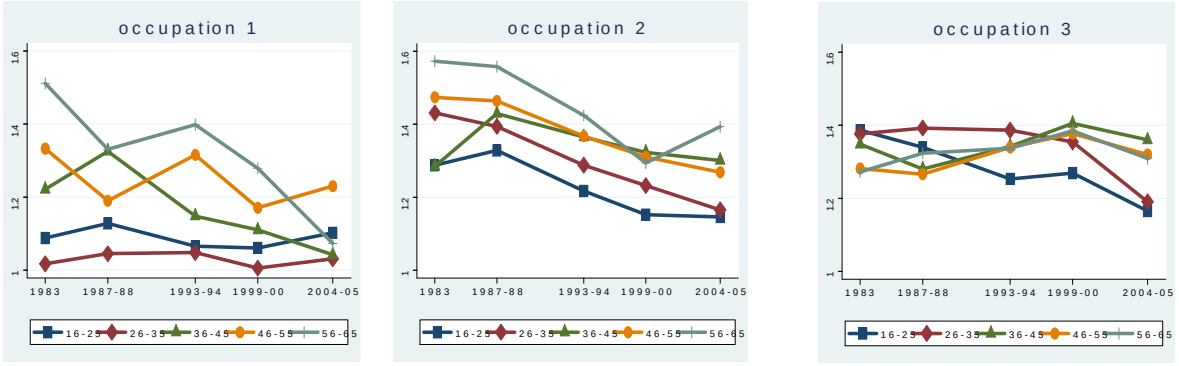
Table 3: Education attainment by occupations

	Panel (a). Education attainment levels					
	children			parents		
	Occ 1	Occ 2	Occ 3	Occ 1	Occ 2	Occ 3
1983	4.41	2.93	2.26	3.58	2.11	1.53
	(0.05)	(0.02)	(0.01)	(0.05)	(0.02)	(0.01)
1987-88	4.41	2.97	2.39	3.71	2.11	1.63
	(0.03)	(0.02)	(0.01)	(0.04)	(0.02)	(0.01)
1993-94	4.51	3.28	2.67	3.92	2.28	1.72
	(0.03)	(0.02)	(0.01)	(0.04)	(0.02)	(0.01)
1999-00	4.49	3.44	2.93	3.87	2.44	1.89
	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.01)
2004-05	4.53	3.52	3.13	3.90	2.56	2.02
	(0.02)	(0.02)	(0.02)	(0.05)	(0.03)	(0.02)

	Panel (b). Education attainment gaps					
	children			parents		
	Occ 1	Occ 2	Occ 3	Occ 1	Occ 2	Occ 3
1983	1.07	1.33	1.39	1.35	1.41	1.28
1987-88	1.09	1.35	1.36	1.26	1.46	1.27
1993-94	1.06	1.24	1.29	1.28	1.36	1.32
1999-00	1.03	1.18	1.29	1.19	1.29	1.36
2004-05	1.07	1.17	1.18	1.17	1.29	1.31

Notes: Panel (a) of this Table presents the average education attainment levels for the two generational groups – parents and children – by occupations. Panel (b) summarizes the relative education gaps for parents and children computed as a ratio of education attainments levels of non-SC/STs to SC/STs. The reported statistics are obtained for each NSS survey round which is shown in the first column. Occ 1 collects white collar workers, Occ 2 collects blue collar workers, while Occ 3 refers to farmers and other agricultural workers. Standard errors are in parenthesis.

Figure 4: Education gaps by age cohorts and occupations



(a) (b) (c)

Notes: The figures show the evolution of the relative education gap between non-SC/STs and SC/STs over time for different age and occupation groups. Occ 1 collects white collar workers, Occ 2 collects blue collar workers, while Occ 3 refers to farmers and other agricultural workers.

3.3 Industry choices

Next, we look at the industry of employment choices of households. 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 tradable industries while Ind 3 comprises non-tradable sectors. Our grouping reflects the tradition of classifying industries into tradables and non-tradables. Incorporating Agriculture as a distinct category is intended to take account of the traditional reliance of the Indian economy on agriculture. See Appendix A.1 for more details on the industry grouping.

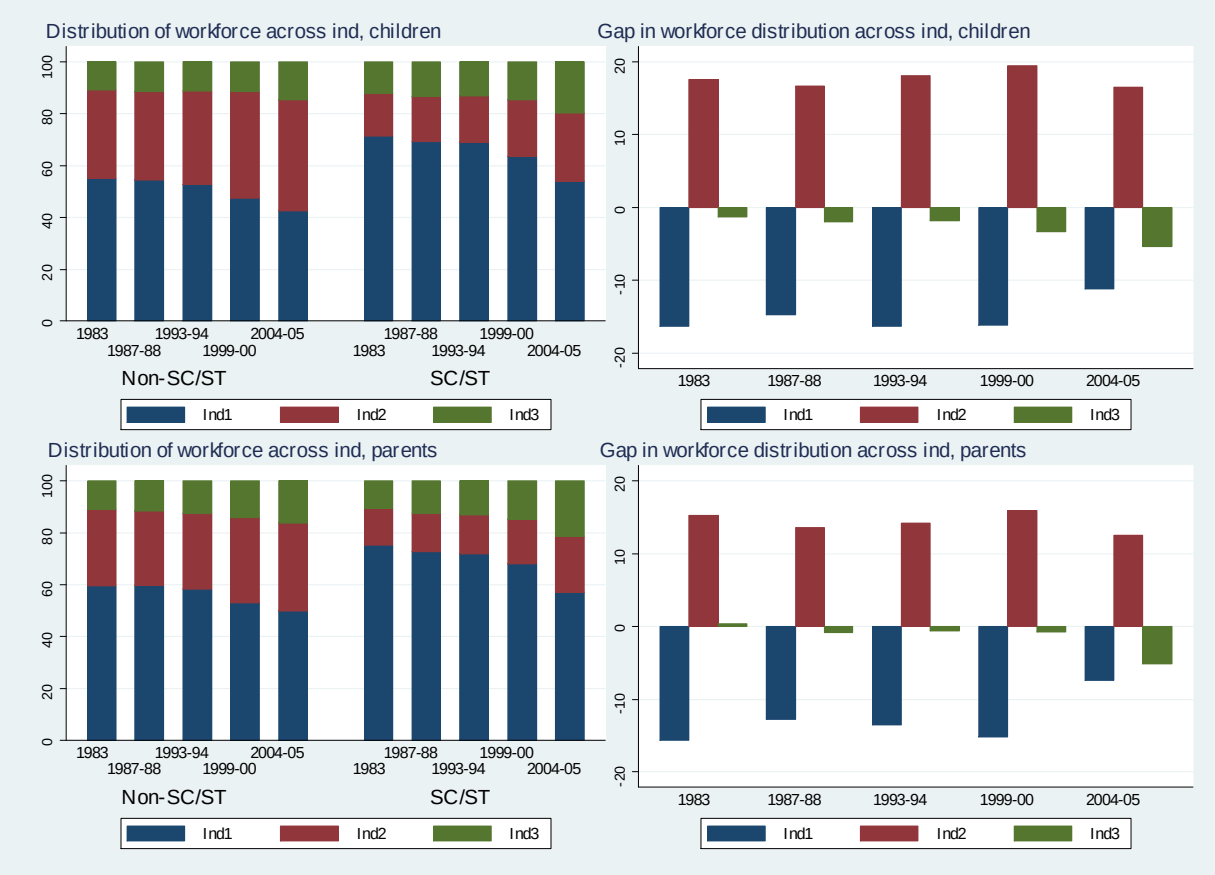
Consistent with the results for occupation choice, SC/STs were and remain more likely to be employed in agriculture and other farming activities (Ind 1) than non-SC/STs, however the gap has somewhat narrowed in the last ten years of our sample. Interestingly, SC/STs are also more likely to work in non-tradable industries (Ind 3) relative to non-SC/STs - with the gap becoming more pronounced over the past twenty years for children and emerging in 2004-05 round for parents. In contrast, a larger share of non-SC/STs population is employed in tradable industries (Ind 2), but the data suggests that SC/STs are gradually moving into those industries as well.

3.4 Wages

The third issue of interest is the evolution of wages of SC/STs and their non-SC/ST cohorts. We are particularly interested in determining whether the rising educational attainment rates and changing occupation distribution of SC/STs towards relatively higher skilled occupations have also resulted in a change in the wage gap relative to non-SC/STs.

Before describing our results on wages it is important to reiterate 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 primarily due to the large segment of the rural population who identify themselves as being self-employed and correspondingly do not report any wage data. Across the census 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 a natural concern about sample selection. 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. In this event we would be biasing our results

Figure 5: Industry distribution



(a) (b)

Notes: Panel (a) of this Figure presents the distribution of workforce of children and parents across three industry categories for different NSS rounds. The left set of bars on each Figure refers to non-SC/STs, while the right set is for SC/STs. Panel (b) presents absolute gaps in the distribution of non-SC/STs relative to SC/STs across three industry categories. The gaps are also reported for children and parents. Ind 1 refers to Agriculture and other farming activities, Ind 2 collects tradable industries, while Ind 3 refers to nontradable industries.

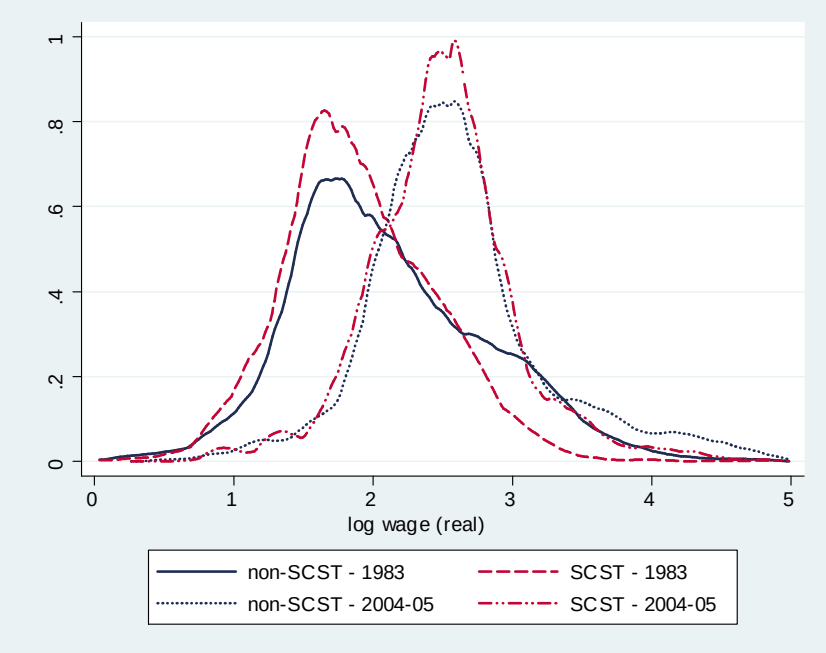
toward finding low wage gaps between the two groups.

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 the overall 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. Averaged 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 miss-

ing 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.

It is instructive to start our analysis of the wage data by presenting the distribution of wages for the first and last rounds of our sample, i.e., for 1983 and for 2004-05. Figure 6 plots the kernel densities of the wage distribution for SC/STs and non-SC/STs separately for both these rounds. Two features emerge clearly from the figure. First, for both groups the wage distribution has shifted sharply to the right. This is to be expected as the period 1983-2005 coincides with the rapid takeoff of the Indian economy. Second, the density functions for the two groups have come much closer together in 2004-05 relative to 1983. Specifically, in 1983 the wage distribution of non-SC/STs appears to exhibit first order stochastic dominance over the distribution for SC/STs. This first-order stochastic dominance disappears in 2004-05 when the two distributions appear to be much more closely aligned.[9]

Figure 6: Wage density for non-SC/STs and SC/STs in 1983 and 2004-05



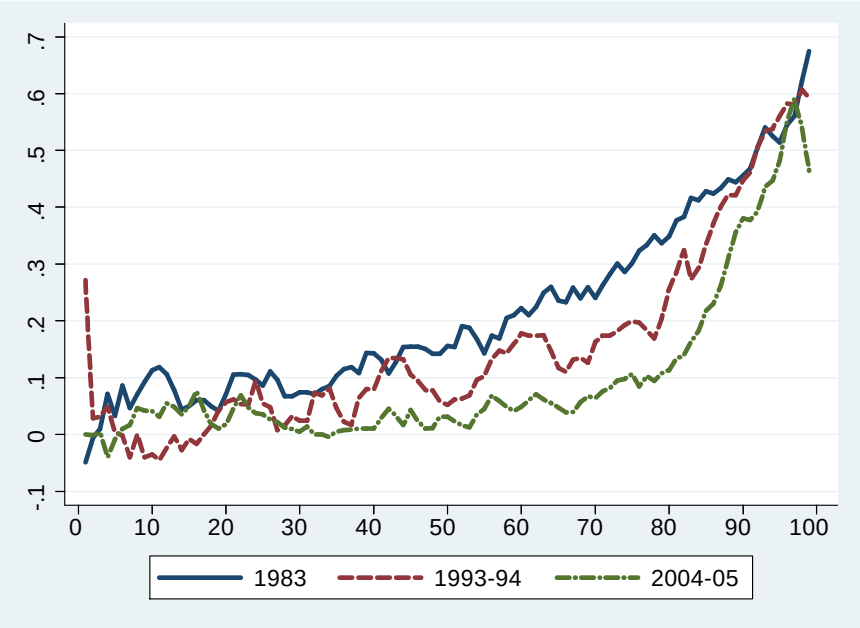
Notes: This figure shows the estimated kernel density of log real wages for non-SC/STs and SC/STs over 1983 and 2004-05 NSS rounds.

We can also examine the changes in wage inequality by looking at the differences in log wages

[9] We should note though that a formal Kolmogorov-Smirnov test of the equality of SC/ST and non-SC/ST wage distributions rejects the null hypothesis of equality both for 1983 and 2004-05. Moreover, the test also rejects the null hypothesis of the SC/ST distribution in 1983 being the same as the SC/ST distribution in 2004-05. This conclusion carries over to a comparison of the non-SC/ST wage distributions in these two rounds as well.

between non-SC/STs and SC/STs for different percentiles of their wage distributions. Figure 7 contrasts three differences for three survey rounds: 1983, 1993-94 and 2004-05.

Figure 7: Differences in log wage percentiles for non-SC/STs and SC/STs in 1983, 1993-94 and 2004-05



Notes: This figure shows the difference in percentiles of log-wages between non-SC/STs and SC/STs plotted against the percentile. The plots are for 1983, 1993-94 and 2004-05 NSS rounds. The line that slopes upward and to the right indicates more unequal distribution for non-SC/STs compared to SC/STs. The lines that are above the horizontal axis indicate stochastic dominance in non-SC/STs wage distribution.

Several features worth pointing out from the Figure 7. First, the figures indicate the presence of the first-order stochastic dominance in the non-SC/ST wage distribution as wages are almost uniformly higher for non-SC/STs than for SC/STs. The degree of the stochastic dominance has declined over time. Second, all three lines slope up and to the right, indicating that the wage distribution of non-SC/STs is more unequal that the wage distribution of SC/STs. This inequality has declined over time, as the lines have become flatter, but significant diffences in inequality remain at the right end of the wage distributions. Overall, the plot confirms our earlier finding of convergence in the two distributions over time as the line for 2004-05 round is well below the line for 1983 round.

The wage distributions plotted in Figure 6 appear to indicate a decline in wage inequality between SC/STs and non-SC/STs between 1983 and 2004-05. We now examine this impression

more closely by contrasting the wage evolution of SC/STs with non-SC/STs over finer sub-groups of age and generation cohorts as well as for all the census rounds under study.

We start with the wage evolution of the generational cohorts of parents and children across the census rounds. Table 4 shows the daily wage earned by working children and parents of non-SC/ST households relative to their SC/ST cohorts over the period 1983 to 2004-05. Overall, the wage premium of non-SC/STs has declined from 47 percent to 35 percent during this period. The Table reveals a contrast between the children and parents in terms of the evolution of the wage gap between SC/STs and non-SC/STs during this period. There has been a clear convergence of wages between children of these two groups. The wage premium of non-SC/ST children has secularly declined from around 34 percent in 1983 to 14 percent by 2004-05. For parents too the non-SC/ST wage premium has fallen but more modestly from 57 percent in 1983 to 45 percent in 2004-05. In fact, the wage premium actually increased in the middle of this period.

The trends we uncover are even more dramatic is we look at wage gaps computed using median wages. In particular, we find that the median wage premium of non-SC/STs relative to SC/STs has declined from 25 percent in 1983 to 6 percent in 2004-05. This decrease is especially pronounced for children for whom the relative wage premium of non-SC/STs relative to SC/STs essentially disappears during this period – falling from 14 percent in 1983 to approximately zero in 2004-05. For parents too the premium fell from 35 percent in 1983 to 19 percent in 2004-05. Clearly, both mean and median wages have been converging across the two groups during this period.

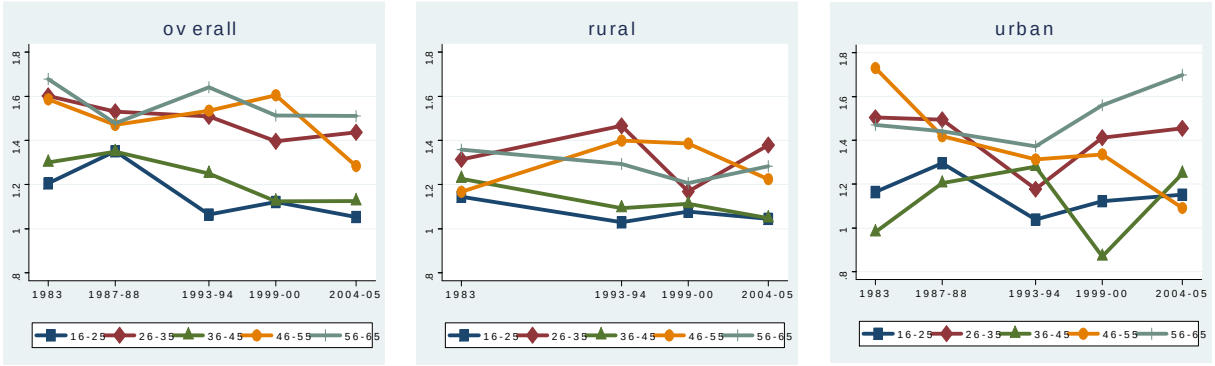
Table 4: The relative wage gap

	overall		children		parents	
	mean	median	mean	median	mean	median
1983	1.42	1.17	1.34	1.14	1.51	1.25
1987-88	1.47	1.26	1.45	1.11	1.48	1.33
1993-94	1.32	1.05	1.18	1.08	1.47	1.16
1999-00	1.34	1.13	1.20	1.09	1.45	1.11
2004-05	1.23	1.03	1.14	0.99	1.31	1.10

Notes: This Table presents the relative mean and median wage gaps for our overall benchmark sample (columns "overall") and separately for the two generational groups – children (columns "children") and parents (columns "parents"). The mean gaps are obtained as the ratios of average real wages of non-SC/STs to SC/STs; while median wage gaps are computed as the ratios of median real wages of the two groups. The reported statistics are obtained for each NSS survey round which is shown in the first column.

Next, we switch from the aggregated generational cohorts to the more disaggregated age-cohorts. Our principal interest here is to determine whether the relative wage gap behavior at the aggregate generational level is masking significant variation across different age cohorts. Figure 8 plots the wage gaps for our five age cohorts for the different census rounds. Panel (a) reveals a general pattern of wage convergence across the cohorts with the younger cohorts, on average, closer to parity than the older ones. Do these overall wage gaps between non-SC/STs and SC/STs reflect significant differences between rural and urban areas? Panels (b) and (c) of Figure 8 shows that the evidence is mixed on this. Relative wages have tended to converge for younger cohorts in both sectors but have often widened for the older ones. Thus, in both urban and rural areas, the 56-65 age group has witnessed a significant increase in the wage premium of the non-SC/STs. Overall, we view this evidence to be along the same lines as the evidence on education, albeit more volatile due to the smaller sample size.

Figure 8: Wage gaps by age cohorts



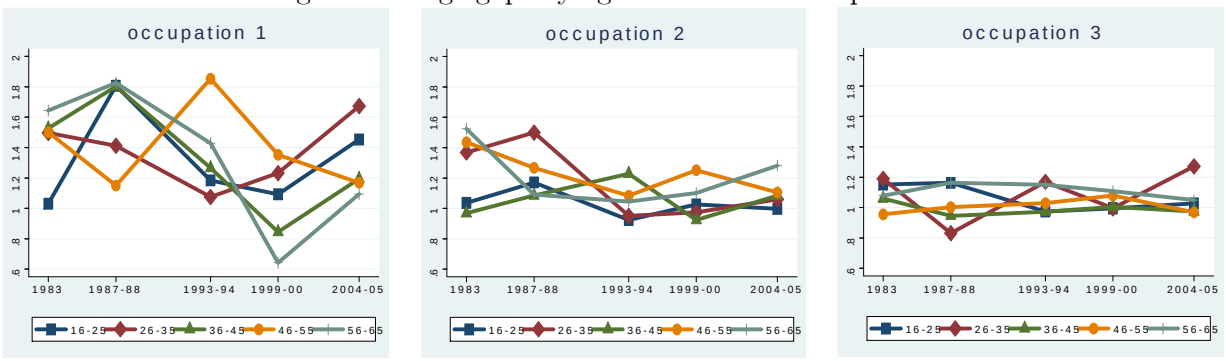
(a) (b) (c)

Notes: The figures show the evolution of the relative wage gap between non-SC/STs and SC/STs over time for different age groups. Panel (a) presents the results for the overall sample, while panels (b) and (c) report the results for rural and urban households separately.

We also examine the behavior of relative wages of non-SC/STs relative to SC/STs by age cohorts across different occupations. Figure 9 presents the results. Panel (a) is for occupation 1, panel (b) is for occupation 2, while panel (c) is for occupation 3. It can be seen that the relative wage premia in occupations 1 and 2 have declined across most age cohorts. In agricultural jobs (occupation 3) the convergence is a bit muted, but the wage gaps there are very small to begin with. As with educations gaps, we see that wage gaps are the most spread out in occupation 1, and to a smaller extent in occupations 2 and 3. Overall, the data suggests that SC/ST wages have been converging

toward non-SC/ST levels, and this trend is most pronounced for higher-skill occupations.

Figure 9: Wage gaps by age cohorts and occupations



(a) (b) (c)

Notes: The figures show the evolution of the relative wage gap between non-SC/STs and SC/STs over time for different age and occupation groups. Occ 1 collects white collar workers, Occ 2 collects blue collar workers, while Occ 3 refers to farmers and other agricultural workers.

The evolution of the wage gaps between SC/STs and non-SC/STs provides an interesting counterpoint to the racial wage gaps that are typically reported in the USA. Thus, during the period 1980-2006 the median wage of black males relative to white male workers has fluctuated between 70 and 80 percent with an average of around 75 percent. During the same period the median wage of Hispanic men relative to white men has declined from 71 percent to under 60 percent.[10] In contrast, our computations above imply that the median wage of SC/STs relative to non-SC/STs has increased from 80 percent in 1983 to 95 percent in 2004-05.[11] Amongst the younger cohorts the wage catch-up has been even faster with the relative wages of SC/ST children having risen from 88 percent to 101 percent during this period. Clearly, the rate of wage convergence for SC/STs since 1983 has been quite striking both at an absolute level as well as in comparison to historically disadvantaged minority groups in more developed countries like the USA.

3.4.1 Conditional Wages

The trends we documented above suggest that the wage gap between SC/STs and non-SC/STs has been declining over the past 22 years. We now examine this impression more closely using more formal statistical tests. In particular, for each census round we estimate a linear log wage regression on the following characteristics: individual age and age squared, his education (*edu*)

[10] These numbers are from US Current Population Survey.

[11] For ease of comparison with the typical wage gap numbers reported for the USA, the SC/ST wage gaps reported here are the inverses of the non-SC/ST to SC/ST relative wage gaps we reported above.

and education squared (*edu sqr*), rural/urban dummy, SC/ST dummy, Muslim dummy, state and occupation specific dummies. We allow the returns to education to be different in rural and urban sectors by including interaction terms between individual education attainment (and education attainment squared) and the rural/urban dummy (*eduXrural* and *edusqrXrural*) in our regression specification.

We also control for differences in reservation policies across states by including state-level SC/ST reservation quotas (*quota SC/ST*). The introduction of reservations for SC/STs in public sector employment and in higher education institutions was a key policy initiative in India. The reservations were provided in proportion to the population shares of SCs and STs.[12]

We control for regional differences by grouping states into five regions – North, South, East, West, Central and Northe-East – and include region dummies in the regression specification. In combination with the state-level reservation policy, this allows us to decompose state-level differences into those attributable to reservations policy, and those due to other time-invariant factors that are common to all state within a given region. The identifying assumption behind this strategy is that the states within a region only differ by the reservation quota they implement.

Table 5 reports the results. We find that the coefficient on the SC/ST dummy variable is negative and significant throughout except for the 1999-2000 round. The negative estimates for the SC/ST dummy indicate that the conditional wages of SC/STs were significantly lower than similarly endowed non-Muslim non-SC/STs. Interestingly, except for the last round, the size of this negative SC/ST effect has become smaller over the sample period. We also find that the lower wages of SC/STs are not due to their rural concentration since the rural dummy becomes insignificant in the last three rounds. Lastly, our results suggest that reservations have tended to increase the average wage for all groups.[13]

As we saw earlier, relative wage gaps have been declining and this decline has coincided with a decline in the gaps in education attainment levels between SC/STs and non-SC/STs. So, how

[12] 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 increased OBC relative wages 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.

[13] In separate regressions in which we controlled for the interaction between the quota variable and the SC/ST dummy, we found that reservations did not have any systematic effect on the wage gap of SC/STs relative to non-SC/STs. This result is in line with Prakash (2009) who also finds broadly insignificant effects of reservations on wages of all except the very poorly educated SC/STs. These results are available from the authors upon request.

Table 5: Conditional wage regressions

	1983	1987-88	1993-94	1999-00	2004-05
	(i)	(ii)	(iii)	(iv)	(v)
age	0.0300***	0.0684***	0.0361***	0.0324***	0.0286***
	(0.0032)	(0.0074)	(0.0047)	(0.0032)	(0.0035)
age sqr	-0.0003***	-0.0007***	-0.0004***	-0.0003***	-0.0002***
	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
1-SC/ST, 0-non SC/ST	-0.0508***	-0.0673**	-0.0207	0.0092	-0.0571***
	(0.0150)	(0.0357)	(0.0212)	(0.0149)	(0.0161)
1-rural, 0-urban	-0.0120	0.1802***	0.0698***	0.0955***	0.2460***
	(0.0247)	(0.0768)	(0.0284)	(0.0228)	(0.0234)
edu 2 dummy	0.1200***	0.1872***	0.1112***	0.1169***	0.1290***
	(0.0201)	(0.0457)	(0.0281)	(0.0193)	(0.0208)
edu 3 dummy	0.2136***	0.2571***	0.1671***	0.1697***	0.1096***
	(0.0236)	(0.0415)	(0.0300)	(0.0213)	(0.0218)
edu 4 dummy	0.2677***	0.2951***	0.2193***	0.1653***	0.1277***
	(0.0305)	(0.0461)	(0.0344)	(0.0217)	(0.0253)
edu 5 dummy	0.5801***	0.6705***	0.3472**	0.3868***	0.3568***
	(0.0406)	(0.0522)	(0.0387)	(0.0288)	(0.0310)
1-muslim, 0-other	-0.0044	-0.0485	-0.0345	0.0389*	-0.0795***
	(0.0246)	(0.0330)	(0.0372)	(0.0220)	(0.0226)
quota SC/ST	-0.0191***	-0.0005	-0.0035*	-0.0129***	-0.0094***
	(0.0014)	(0.0029)	(0.0020)	(0.0013)	(0.0013)
R-sqr	0.4031	0.4026	0.2233	0.3721	0.3179
N	7370	3241	8059	9076	8502

Notes: This table presents estimation results from a regression of log real wages on a set of individual-level, household-level and aggregate control variables for five NSS survey rounds ((i)-(iv)). Refer to the text for model specification details. Standard errors are in parenthesis. * p-value≤0.10, ** p-value≤0.05, *** p-value≤0.01.

much of the wage gap between the two groups arises due to differences in education? We answer this question by using the Oaxaca-Blinder decompositions.

We employ a two-fold Oaxaca-Blinder procedure which involves running wage regressions separately for the two groups on a list of controls including education levels. One then decomposes the wage gaps into the part coming from the different coefficients on the controls for the two groups, and the part due to differences in endowments between the two groups. To obtain the reference coefficients we use a pooled approach which allows for a group membership indicator (as in Fortin, 2006). Our controls are the same as in the regression specification above. Table 6 reports the results for the overall sample.

Columns (i) and (ii) report average log wages of non-SC/STs and SC/STs, respectively; while column (iii) reports the wage gap for the two groups over different survey rounds.[14] Column (iv) attributes a fraction of this gap to group differences in measured endowments, while column (v) reports the size of the gap attributable to discrimination or potentially to group differences in

[14] Wage gap reported in Oaxaca-Blinder decomposition is computed as a difference between average log wages of non-SC/STs and SC/STs. The relative wage gaps we reported earlier were obtained as the ratios of average wages (in levels) of non-SC/STs to SC/STs. As a result, the magnitudes of the gaps from these two approaches are different, but the trends are the same.

Table 6: Oaxaca-Blinder decomposition

	non-SC/ST	SC/ST	difference	explained	unexplained	fraction to edu
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
1983	2.12	1.90	0.22	0.17	0.05	0.56
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	
1987-88	2.50	2.25	0.25	0.19	0.06	0.78
	(0.02)	(0.04)	(0.04)	(0.03)	(0.04)	
1993-94	2.28	2.10	0.18	0.16	0.02	0.36
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	
1999-00	2.52	2.34	0.18	0.19	-0.01	0.35
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	
2004-05	2.60	2.50	0.10	0.04	0.06	0.95
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	

Notes: This Table presents a two-fold Oaxaca-Blinder decomposition of the log-wage gap between non-SC/STs and SC/STs for the five NSS survey rounds, as identified in the first column. Columns (i) 'non-SC/ST', (ii) 'SC/ST' and (iii) 'difference' present the average real log-wages for non-SC/STs, SC/STs and the gap between them, respectively. Columns (iv) 'explained' and (v) 'unexplained' refer to the size of the wage gap attributable to differences in endowments between non-SC/STs and SC/STs, and to the differences in the returns to those endowments, respectively. Column (vi) 'fraction to edu' reports the share of the explained wage gap coming from education attainment differences between the two groups. Standard errors are in parenthesis.

unobserved characteristics. Finally, the last column of Table 6 reports the fraction of the overall log wage difference that is accounted for by differences in education endowments alone. The column shows that differences in education accounted for 38 percent of the wage gap in 1983 which increased to 57 percent in 2004-05.

We also conduct an analogous non-SC/ST – SC/ST wage gap decomposition for parents and children separately and find that the results are similar. While the detailed results are reported in the supplemental tables available from http://faculty.arts.ubc.ca/vhnatkovska/research.htm, the results are broadly similar to those reported here for the overall sample. Specifically, we have three main findings: (i) in accordance with our earlier, unconditional results, the wage gaps for children are significantly smaller than those for parents; and decline by a larger amount than the gaps for parents over our sample period; (ii) majority of the wage gap for both parents and children can be attributed to the differences in endowments (with an average of 82% of the overall gap for parents and 85% for children); (iii) majority of the gap in endowments is due to the differences in education attainments of the two groups.

These results, in conjunction with the facts that both wage and education differences have been

declining over time, suggest that the major part of the decline in the wage differences between SC/STs and non-SC/STs between 1983 and 2004-05 is due to a decline in the education differences between them. Moreover, the wage gap between SC/STs and non-SC/ST has essentially disappeared after one controls for age, education, occupation, location (state and rural-urban).

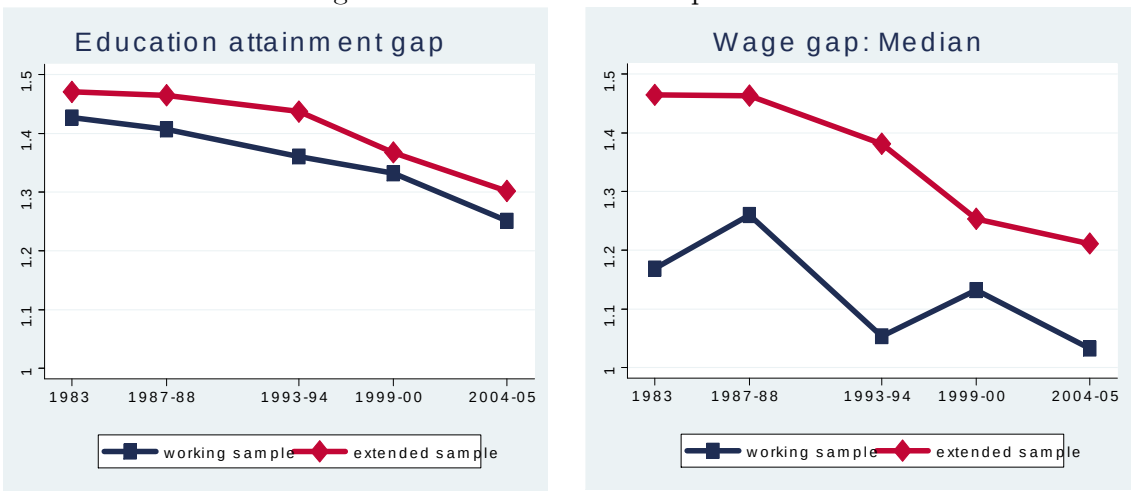
3.5 Sample and Robustness

A key sample restriction underlying our data is that we have only considered joint households with a full-time working male head of household and at least one full-time working male child between the age of 16-65. This restriction allowed us to construct the groups Parents and Children whose fortunes we tracked over time. Moreover, this restriction allowed us to study inter-generational mobility patterns within SC/ST and non-SC/ST households. The imposed restriction did however reduce the sample size by dropping brothers/cousins of the head of the household as well as their children. Moreover, the restriction also dropped households with only one generation of full-time working males. Does this sample restriction matter for our results? To check this we compared the relative education and wage gaps from the unrestricted sample with the results reported above which used the restricted sample.

Panel (a) of Figure 10 shows the average education attainment levels of non-SC/STs relative to SC/STs for both the unrestricted sample and the restricted sample.[15] Clearly, the same pattern of convergence emerges in both samples. Similarly, Panel (b) of Figure 10 plots the median wages of non-SC/STs relative to median SC/ST wages for both samples. Once again, the convergence patterns are similar for both samples. Both figures show, however, that the non-SC/ST to SC/ST gaps are larger for all years in the unrestricted sample. This is due to the fact that the larger sample includes a number of older individuals (like brothers of the head of household living in a joint household) who get dropped in the restricted sample. Recall that the measured gaps in both education and wages are larger for older cohorts. We conclude from this exercise that our results on education and wage convergence are not an artifact of the restricted sample size imposed by our joint household condition.

[15] Both samples are restricted to males between the ages of 16-65 who are working full-time, are not enrolled in any education institution, and for whom there is education and occupation information (and wages if we are computing relative wage gaps). The difference between the two samples is that the unrestricted sample does not impose the restriction of considering only households with at least two working males of different generations.

Figure 10: Robustness to Sample Choice



(a) (b)

Notes: Panel (a) depicts the ratio of the average education attainment level of non-SC/STs to SC/STs for the sample with no joint household restriction. Panel (b) presents the ratio of the median wages of non-SC/STs to SC/STs for the same unrestricted sample.

4 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? Are children changing occupation, industry, education and income status relative to their parents more frequently than before? 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 to be a binary variable that takes a value of one if the child'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 education switch variable as **switch-edu;** occupation switch variable as **switch-occ;** and 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.

4.1 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:

$$P_i \equiv \Pr(y_i = 1|x_i) = E\left(y_i|x_i\right) = \psi(x_i\beta), \tag{4.1}$$

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), per capita consumption expenditure for the household ($mpce$), and his rural location ($rural$); and state-level characteristics, such as state-level reservation quota for SC/STs, and region-specific fixed effects. Thus,

$$\begin{aligned} x_i\beta &= \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 SC/ST_i + \beta_4 muslim_i \\ &\quad + \beta_5 rural_i + \beta_6 hh\_size_i + \beta_7 quota\_scst_j + \sum_{j=1}^{6} \alpha_j region\_dummy_j. \end{aligned}$$

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 conditional probability of intergenerational education switch. We compute these probabilities for the overall sample as well as for SC/STs and non-SC/STs separately.[16]

Panel (a) of Figure 11 depicts the computed probabilities of intergenerational switches in education attainment together with the ± 2 std error confidence bands (dashed lines).[17] The remarkable feature highlighted by the Figure is that 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

[16] The detailed regression results are available in supplemental tables from http://faculty.arts.ubc.ca/vhnatkovska/research.htm.

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

one notes that in 1983, the probability of an intergenerational education switch for SC/ST households was a meagre 41 percent relative to the 57 percent corresponding probability of non-SC/ST households.

Figure 11: 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.

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 11 reveals that the average size of the switch has been increasing over time for both groups. Moreover, by the end of our sample, the switch sizes for the two groups not only converged at the overall level of 1.1, 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 education improvements.

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 34 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 48 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. Detailed summary of these results is available in the Appendix Table A3.

4.2 Occupation Mobility

We now turn to intergenerational occupation switches. 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:

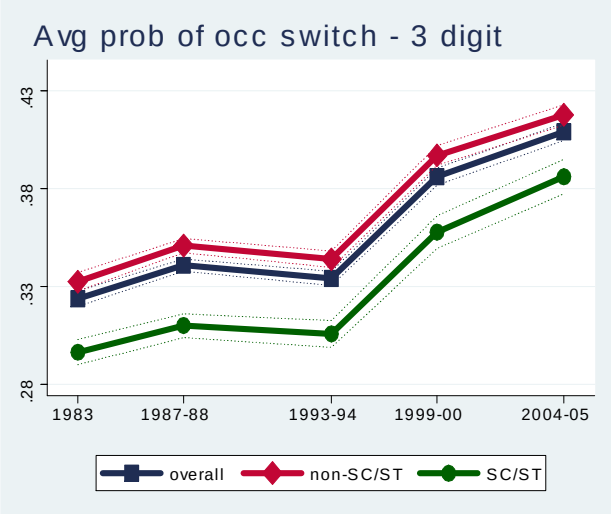
$$\begin{aligned} x_i\beta &= \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 SC/ST_i + \beta_4 muslim_i \\ &\quad + \beta_5 rural_i + \beta_6 hh\_size_i + \beta_7 quota\_scst_j + \sum_{j=1}^{4} \theta_j edu\_dum_j \\ &\quad + \sum_{j=1}^{6} \alpha_j region\_dummy_j + \sum_{j=1}^{9} \gamma_j occup\_dummy_j. \end{aligned} \tag{4.2}$$

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 ($SCST$), religion ($muslim$) and whether or not the individual switched education levels relative to his parent (*switch-edu*). We also include an interaction between education switch and the SC/ST dummy ($switch\text{-}edu_i * SC/ST_i$) to allow for a different effect of an educational switch by SC/STs on their probability of occupational mobility. Second, we control for household-level characteristics such as household size (hh_size_i), per capita consumption expenditure for the household ($mpce_i$), and his rural location ($rural_i$). Third, we allow for occupation-specific fixed effects and state-level fixed effects using state-level SC/ST reservation quotas and state-specific fixed effects.

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.[18]

Figure 12: 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 12 depicts the computed probabilities of occupation switches at the three-digit level (dotted lines plot the ± 2 std 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 42 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 43 percent while for SC/STs it has gone from 30 to 39 percent. 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.[19]

4.2.1 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 sector? Which sectors are absorbing most

[18] The detailed estimation results are available in Supplemental Tables from http://faculty.arts.ubc.ca/vhnatkovska/research.htm.

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

Table 7: 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.

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, high p_{ii} would reflect relatively little intergenerational occupational mobility while large p_{ij} where $i \neq j$ would indicate high mobility.

We report the results for the three broad occupation categories we defined earlier in Table 7. 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.

Table 7 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 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 to us as it indicates an improvement in the quality of jobs across generations. In 1983 the conditional probability of an intergenerational switch from occupation 3 to occupations 1 or 2 was 15% for non-SC/STs and 11% for SC/STs. By 2004-05 these numbers had risen to 24% for non-SC/STs and 22% 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.

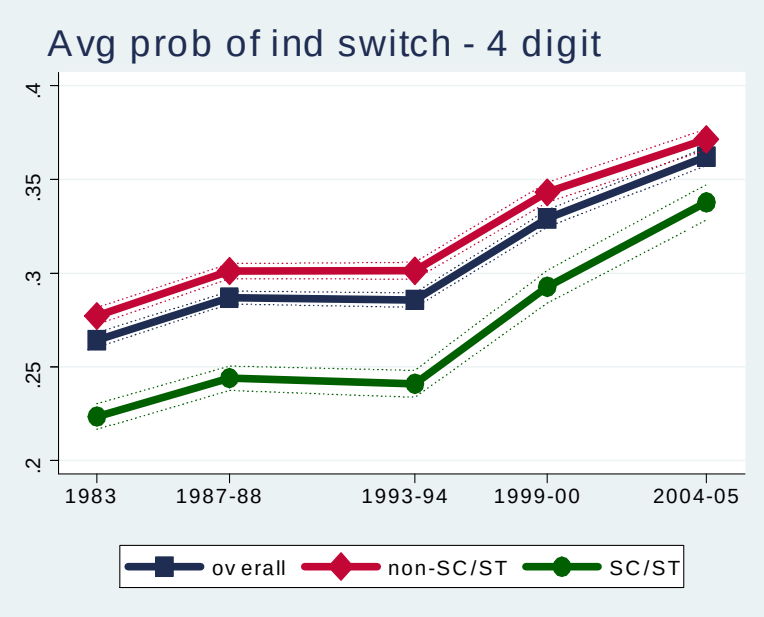
Third, the probability of a child working in occupation 3 conditional on his father being employed in occupation 1 or 2 has declined from 50% to 32% for SC/STs and from 30% to 23% for non-SC/STs over our sample period. We believe that this reflects a significant reduction in regress prospects of SC/ST households during this period.

4.3 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 2005. We define intergenerational industry switch in the same manner as occupation switches and estimate the conditional probability of industry switches using equation (4.2).

Figure 13 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 $\pm$ 2 std error 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 37 percent in 2004-05 period. The industry mobility trends of both SC/STs and non-SC/STs have been similar although

Figure 13: 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.

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.[20]

4.3.1 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 and summarize them in Table 8. 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

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

Table 8: 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	0.08	0.05	0.67	Ind **1**	0.90	0.06	0.04	0.78
	(0.00)	(0.00)	(0.00)	(0.00)		(0.01)	(0.01)	(0.00)	(0.01)
Ind **2**	0.11	0.83	0.06	0.24	Ind **2**	0.17	0.75	0.08	0.13
	(0.01)	(0.01)	(0.00)	(0.00)		(0.02)	(0.02)	(0.01)	(0.01)
Ind **3**	0.19	0.32	0.49	0.08	Ind **3**	0.24	0.18	0.58	0.09
	(0.01)	(0.01)	(0.01)	(0.00)		(0.02)	(0.02)	(0.03)	(0.00)

(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	0.17	0.07	0.60	Ind **1**	0.79	0.12	0.09	0.68
	(0.01)	(0.00)	(0.00)	(0.00)		(0.01)	(0.01)	(0.01)	(0.01)
Ind **2**	0.08	0.82	0.10	0.29	Ind **2**	0.13	0.72	0.15	0.18
	(0.00)	(0.01)	(0.01)	(0.00)		(0.01)	(0.02)	(0.01)	(0.01)
Ind **3**	0.16	0.36	0.49	0.11	Ind **3**	0.12	0.29	0.60	0.14
	(0.01)	(0.01)	(0.01)	(0.00)		(0.01)	(0.02)	(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 all traded industries, while Ind 3 refers to to all nontraded industries. 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.

employed in each of the industries in a given round. Panel (a) gives the numbers for 1983 and Panel (b) for 2004-05.

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 tradable industries, the probability of moving into non-tradable industries 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.

4.4 Income Mobility

Our fourth, and probably the most typical, measure of intergenerational mobility is on income. We turn to this issue next. The goal of this exercise 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. The intergenerational elasticity 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.

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

$$\begin{aligned} w_{ic} &= \alpha + \beta w_{ip} + \gamma_1 A_{ip} + \gamma_2 A_{ip}^2 + +\gamma_3 A_{ip}^3 + \delta_1 \tilde{A}_{ic} + \delta_2 \tilde{A}_{ic}^2 \\ &\quad +\delta_3 \tilde{A}_{ic}^3 + \theta_1 w_{ip} \tilde{A}_{ic} + \theta_2 w_{ip} \tilde{A}_{ic}^2 + \theta_3 w_{ip} \tilde{A}_{ic}^3 + \varepsilon_i \end{aligned} \tag{4.3}$$

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.

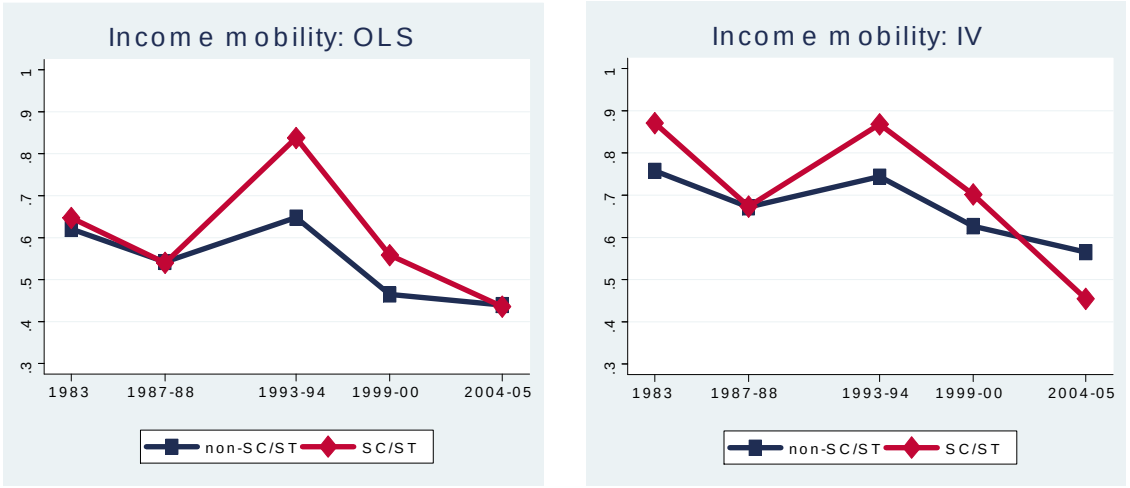
We run this regression separately for each NSS sample year. Note that the constant α picks up

any sample year specific factors. 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 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 plot the OLS estimates in panel (a) of Figure 14 below, while panel (b) of Figure 14 presents our estimates from an instrumental variable (IV) regression.[21] There are three features 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 0.4 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 14: 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

[21] As with the other regressions, the complete estimation results are available in supplemental tables from http://faculty.arts.ubc.ca/vhnatkovska/research.htm.

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 0.9 to 0.4 and, by the end of our sample period, was below that for non-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. However, there are significant differences in the times series behavior of these mobility indicators for SC/STs relative to non-SC/STs. While intergenerational educational and wage mobility of the two groups have tended to converge, occupational and industry mobility rates have not converged similarly.

5 Conclusion

In this paper we have studied the evolution of occupation and industry choices, education attainment rates and wages in India between 1983 and 2004-05 with a special focus on the fortunes of scheduled castes and scheduled tribes (SC/STs). We have found that the 22-year period under study has been a period of dramatic changes for these historically disadvantaged groups. SC/STs have systematically reduced the gap with non-SC/STs in education attainment levels and have been changing occupation and industry of employment at increasingly faster rates. Moreover, the wage gap between SC/STs and non-SC/STs has narrowed sharply during this period. We have also found that the majority of the wage gap is accounted for by differences in education whose contribution has been rising over time. The caste effect on wages appears to have almost disappeared. Crucially, we find that these trends are the sharpest amongst the younger cohorts and in urban areas. The last two features are especially uplifting since they are potentially indicative of the types of changes one may expect in the future since India has been becoming increasingly urbanized and younger over time.

It is worth reiterating that SC/ST wages have been converging toward non-SC/ST levels across cohorts, education and occupation categories. Moreover, the speed of this convergence is impressive not just at an absolute level but also when compared to the wage convergence experienced by historically disadvantaged minority groups elsewhere such as Blacks and Hispanics in the USA. We find this evidence particularly reassuring in terms of the future prospects of SC/STs in India.

What explains these significant changes in the Indian social landscape? We believe that the rapid structural changes in the Indian economy over the past 25 years are at the heart of this progress. The liberalization of the previously restricted economy has opened up new opportunities for the private sector. While the increase in potential opportunities is common to all segments of the population, the more rapid response of SC/STs probably reflects a confluence of factors. One factor may be the competitive pressures that were unleashed on markets by the economic liberalization. As argued by Becker (1957), increasing competition raises the losses to businesses from pursuing discriminatory labor market practises. This reduces the degree of wage discrimination. The resultant decline in the wage gap could then also induce these disadvantaged groups to increase their education attainment rates since the returns to education rise. A second factor may be that the rapidly changing socioeconomic environment in India has presented SC/STs with a historic opportunity to break out of a centuries-old cycle of illiteracy and poverty, and they have been acting proactively to take advantage of it. A strengthening of community based networks of SC/STs along the lines suggested in Munshi (2010) may have also been at play in accelerating this process. The third possibility is that the reservations policy in place since 1950 for public sector jobs and higher education seats may have played a role in the declining wage gaps. The first two possibilities imply that caste may be becoming a less important factor in economic allocations in India while the third factor would put caste-based policies at the center of the explanation. We intend to examine these potential explanations in future work.

References

Banerjee, B., and J. Knight (1985): "Caste Discrimination in the Indian Urban Labour Market," *Journal of Development Economics*, 17(3), 277–307.

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

Becker, G., and N. Tomes (1986): "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4(3), 1–39.

Behrman, J., and P. Taubman (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.

Borooah, V. K. (2005): "Caste, Inequality, and Poverty in India," *Review of Development Economics*, 9(3), 399–414.

Fortin, N. M. (2006): "Greed, Altruism, and the Gender Wage Gap," Working papers, University of British Columbia.

Haider, S., and G. Solon (2006): "Life-cycle Variation in the Association Between Current and Lifetime Earnings," *American Economic Review*, 96(4), 1308–1320.

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

Jalan, J., and R. Murgai (2009): "Intergenerational Mobility in Education in India," Working papers, Indian Statistical Institute, Delhi.

Lee, C., and G. Solon (2009): "Trends in Intergenerational Income Mobility," *The Review of Economics and Statistics*, 91(4), 766–772.

Madheswaran, S., and P. Attewell (2007): "Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey," *Economic and Political Weekly*, 42(41), 4146.

Maitra, P., and A. Sharma (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 M. Rosenzweig (2006): "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy," *American Economic Review*, 96(4), 1225–1252.

——— (2009): "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," Working papers, Brown University, Department of Economics.

Prakash, N. (2009): "The Impact of Employment Quotas on the Economic Lives of Disadvantaged Minorities in India," Working papers, Dartmouth College, Department of Economics.

Solon, G. (1992): "Intergenerational Income Mobility in the United States," *The American Economic Review*, 82(3), 393–408.

——— (2002): "Cross-country Differences in Intergenerational Earnings Mobility," *The Journal of Economic Perspectives*, 16(3), 59–66.

A Appendix

A.1 Data Appendix

A.1.1 National Sample Survey (NSS)

The National Sample Survey Organisation (NSSO), set up by the Government of India, conducts rounds of sample surveys to collect socio-economic 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 minimised. 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 systematically 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 the time.

A.1.2 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.

We do not consider individuals whose occupation is not classified in the National Classification of Occupation – 1968. The NSS reports weekly (one week reference) as well as the usual occupation (one year reference) of household members. Household members can undertake more than one activity in the reference week. We consider one activity per member. If an individual undertook multiple activities over the reference period then we consider the activity in which she spent maximum days in a week. If there are more than one activities with equal days, we consider the one with paid employment (wage is not zero or missing). Wages of individual household members are reported by daily activity in the last one week. We calculate the daily wage by dividing total wage paid in last one week by days spent in that activity. 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). One of the main objectives of this paper is to document intergenerational mobility. Therefore, we include households only if the household head and at least one member from younger generations have information on occupation, education and at least sixteen years old. Lastly, we only include full time workers in our analysis. 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. We drop observations if total number of days worked in the reference week is more than seven.

We have two operational sub-samples – the overall sample and the wage sample:

1) The overall 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 different generation (son, grandson etc.) also in the 16-65 age group, where neither is enrolled in an educational institution, both have education and occupation information. Within included households, we only consider the head of the household and his direct male descendants.

2) The wage sample includes only those members in the overall sample who have non-missing and non-zero wage data.

Both the overall sample and the wage sample are further subdivided into two 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 grand children).

A.1.3 Occupation 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.

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 2
H	Hotels and Restaurants	Ind 2
I	Transport, Storage and Communications	Ind 2
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

A.2 Intergenerational education mobility

Table A3 presents average conditional 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 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 A3: 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	0.5013	0.3676	0.0915	0.0949	0.0810
	(0.0007)	(0.0007)	(0.0010)	(0.0002)	(0.0003)	(0.0004)
1993-94	0.5235	0.5450	0.4619	0.0828	0.0885	0.0663
	(0.0006)	(0.0006)	(0.0009)	(0.0002)	(0.0002)	(0.0003)
1999-00	0.5363	0.5486	0.5043	0.0903	0.0955	0.0770
	(0.0007)	(0.0007)	(0.0012)	(0.0003)	(0.0003)	(0.0004)
2004-05	0.5792	0.5764	0.5868	0.0887	0.0923	0.0790
	(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 (4.1), except we used a binary variable denoting education improvements or education reductions as the left-hand-side variable. Standard errors are in parenthesis.