Working paper

International Growth Centre

Convergence Across Castes

Viktoria Hnatkovska
Amartya Lahiri

July 2012

When citing this paper, please use the title and the following reference number:
S-35011-INC-1

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid
from the British people

Convergence Across Castes*

Viktoria Hnatkovska† and Amartya Lahiri†

July 2012

Abstract

The past 25 years have witnessed a remarkable economic catch-up by the historically discriminated against scheduled castes and tribes (SC/STs) in India towards non-SC/ST levels in the terms of their education attainment levels, their occupation choices as well as their wage and consumption levels. In this paper we document that these aggregate data patterns mask empirically significant sectoral hetereogeneity in the evolution of the caste gaps in wages: wage gaps have declined in services, widened in agriculture and stayed unchanged in manufacturing between 1983 and 2008. We develop a multi-sector model with two types of agents to show how aggregate TFP shocks along with a process of structural transformation can induce a convergence between the two groups without any other concurrent redistributive policy changes as long as there exists an *initial* affirmative action policy in education and/or jobs for the relatively disadvantaged group. We provide some indirect evidence in support of this channel by examining the convergence patterns of Muslims who were not covered by such affirmative action policies.

JEL Classification: J6, R2

Keywords: Castes, convergence, labor

1 Introduction

A perennial challenge of managing the development process is to balance the macroeconomic goals of growth and development with the microeconomic goals of equity and distributional fairness. These challenges often come to the fore during periods of rapid economic changes in growing economies.

*We would like to thank Satyajit Chatterjee and seminar participants at the IGC conference in Delhi, the SED meetings in Cyprus, Pennsylvania State University, FRB Philadelphia and the University of Maryland for helpful comments and discussions. Thanks to Arka Roy Chowdhury for providing research assistance. We also thank the India Growth Center network for providing funding to support this research.

†Department of Economics, University of British Columbia, 997 - 1873 East Mall, Vancouver, BC V6T 1Z1, Canada. E-mail addresses: viktoriya.hnatkovska@ubc.ca (Hnatkovska), amartya.lahiri@ubc.ca (Lahiri).

An example of this phenomenon is India over the past 25 years. This period has witnessed a rapid takeoff of the Indian economy with average annual growth rates doubling relative to the pre-reform phase. However, this takeoff has also been accompanied by a vigorous debate regarding the effects of economic reforms on poverty and economic inequality. A big part of the development challenge is to devise policy initiatives in order to manage these often clashing goals. This paper aims to contribute to this challenge. In particular, the paper aims to uncover the channels through which the large macroeconomic changes in India have affected the economic fortunes of different social groups, i.e., uncover the black-box of the linkages between the macro and micro developments.

We approach the issue by focusing on the experience of Scheduled Castes and Scheduled Tribes (SC/STs) – an historically underprivileged section of Indian society. In recent work, we have shown that these groups have experienced a rapid catch-up towards non-SC/ST levels in their education attainment levels, their occupation choices, as well as in their wages and consumption levels (see Hnatkovska, Lahiri, and Paul (2012a)). Accompanying this catch-up has been a sharp convergence in the intergenerational mobility rates in these three indicators as well (see Hnatkovska, Lahiri, and Paul (2012b)). These developments immediately raise the key question: what are the forces that have driven these convergent trends? Have they been sparked by the large aggregate changes in the Indian economy during this period which affected all sectors symmetrically? Or, is the convergence primarily due to changes in the caste gap in specific sectors of the economy? This paper is an attempt at decoupling these two forces of change and assessing the relative importance of the two.

Recognizing that the past 25 years have also been a period of rapid structural change in the Indian economy, we begin by breaking down the overall patterns by sectors. Specifically, we break down the evolution of the caste gaps in education, occupation and wages by the broad sectorial aggregates of agriculture, manufacturing and services. The key patterns that emerge from this *unconditional* examination of the data are that gaps between the castes in their education attainment levels have declined secularly in all three sectors. Moreover, the sectoral employment distributions of SC/STs and non-SC/STs have also become more aligned over time. However, we find that the caste wage gaps have not evolved symmetrically across the three sectors. While the wage gap between non-SC/STs and SC/STs has declined in the service sector, it has widened in the agricultural sector and stayed relatively unchanged in the manufacturing sector. We also compute sectoral total factor productivity and labor productivity and find that both these measures of productivity shared a common rising trend across sectors.

To examine whether the unconditional behavior of the sectoral wage patterns are statistically significant, we use the NSS data from five rounds (rounds 38 to 64) to study the determinants of wages and the time series behavior of the sectoral wage gaps. Of crucial interest to us is whether the asymmetric behavior of the sectoral wage gaps is statistically significant or not. Using a set of caste, sector and time dummies, we find that both the median and mean wage gaps between non-SC/STs and SC/STs declined significantly in services, rose significantly in agriculture while remaining relatively unchanged in manufacturing. Hence, the asymmetric behavior of the unconditional sectoral wage gaps are indeed statistically significant.

To rationalize the differential behavior of the wage gaps across sectors, we develop a simple model with endogenous skill formation and multiple sectors. The goal of the exercise is two-fold. First, we want to isolate the various margins that affect the skill and industry distribution. A particular focus here is on the effect of aggregate productivity increases as well as changes in the costs of skill acquisition in the context of an environment which generates a structural transformation of the economy in response to aggregate shocks. Second, we identify conditions under which the model can explain the overall wage convergence between the groups while simultaneously generating the the sectoral heterogeneity in changes in the wage gaps. Our results identify the existence of pre-existing education subsidies for SC/STs as a necessary condition for the model to match the observed dynamics of the aggregate and sectoral wage gaps. We interpret the reservations in education and jobs provided to SC/STs in India since 1951 as the policy counterparts to the conditions identified by the model. Moreover, we provide independent corroborating evidence in support of this channel by documenting the widening gaps experienced by Muslims in India – another relatively disadvantaged minority group who did not have access to similar affirmative action protection.

Our work is related to some recent and past work on gaps between castes by a number of different authors. In terms of the gaps between the castes, Banerjee and Knight (1985), Borooah (2005) and Madheswaran and Attewell (2007) have studied either wage or employment discrimination faced by SC/STs in the urban Indian labor market. Ito (2009) has examined wage and employment discrimination simultaneously in two Indian states – Bihar and Uttar Pradesh. Kijima (2006) has used NSS data to study consumption inequality of SC/ST households. In terms of conceptual mappings of caste based frictions to outcomes, Munshi and Rosenzweig (2006) show how caste-based networks affect education choices by gender.

There is also a parallel and older literature on the black-white gaps in the USA that is also related to our work. Human capital based explanations for inter-group differences can be found in Becker and Tomes (1986) and Loury (1981), while Becker (1957) formalizes the classic discrimination based differences across groups. In recent work, Bowles, Loury, and Sethi (2009) show how network learning effects can generate long run earnings disparities between groups despite there being no underlying differences in innate ability. Our explanation is distinct in that it focuses on the role of aggregate shocks and structural transformation as key components of the change along with affirmative action policies.

The next section lays out the broad empirical regularities in terms of overall economy as well as the caste gaps. Section 3 contains the wage regressions and the role of time and sectoral effects on the wage gaps. Section 4 formalizes the model and the main analytical results while the last section concludes.

2 Empirical regularities

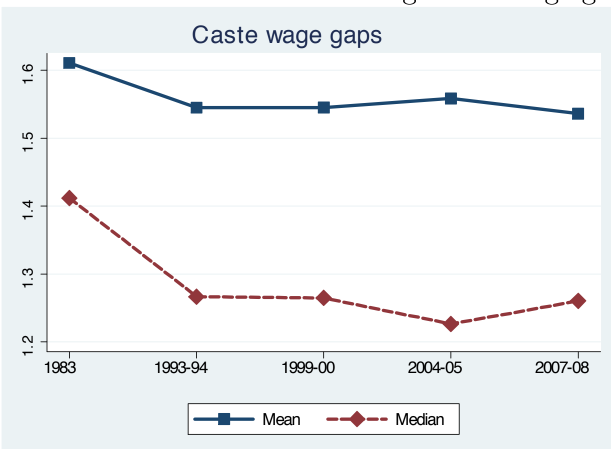
Our data comes from different sources. The primary data source is the National Sample Survey (NSS) rounds 38 (1983), 43 (1987-88), 50 (1993-94), 55 (1999-2000), 61 (2004-05) and 64 (2007-08). The NSS provides household-level data on approximately 600,000 individuals on education, employment, consumption, and wages as well as other social characteristics. We consider individuals between the ages 16-65 belonging to male-headed households who were not enrolled full time in any educational degree or diploma. The sample is restricted to those individuals who provided their 4-digit industry of employment code information as well as their education information.[1] Our focus is on full-time working individuals who are defined as those that worked at least 2.5 days per week, and who are not currently enrolled in any education institution. This selection leaves us with a working sample of around 165,000-182,000 individuals, depending on the survey round. The wage data is more limited. This is primarily due to the prevalence of self-employed individuals in rural India who do not report wage income. As a result, the sub-sample with wage data is limited to about 48,000 individuals on average across rounds. Details on the data are contained in the Data Appendix to this paper.

We start by reporting some aggregate facts regarding the education and wage gaps between SC/STs and non-SC/STs since 1983. These facts are mostly borrowed from the results reported in Hnatkovska, Lahiri, and Paul (2012b). Figure 1 reports the wage gaps between the castes. Panel (a) shows the mean and median wage gaps between the groups across the NSS rounds. The

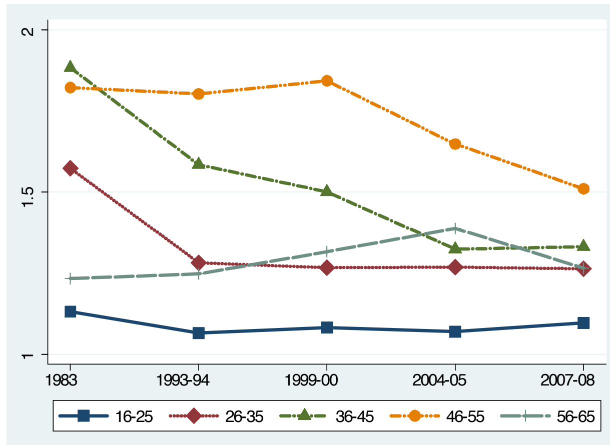
[1] We also consider a narrower sample in which we restrict the sample to only males and find that our results remain robust.

picture shows that the gap between the groups declined by both measures. Panel (b) breaks down the groups by age and plots the median wage gaps by age groups. For all except the oldest age-group, the median wage gaps declined secularly during this period. Hence, both plots reveal an unambiguous pattern of wage convergence between the two groups.

Figure 1: Wage gaps between castes



(a)



(b)

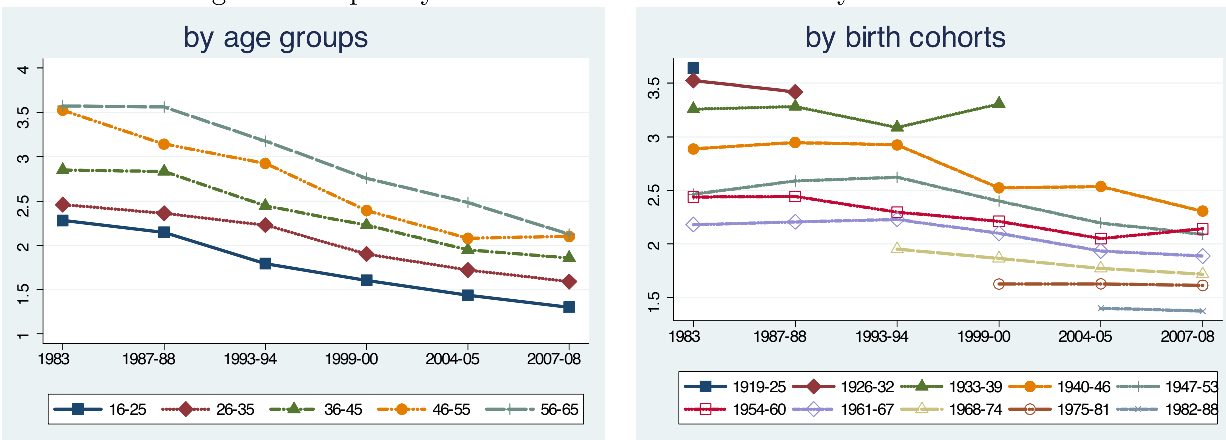
Notes: Panel (a) of this Figure presents the mean and median wage gaps between SC/STs and non-SC/STs (expressed as non-SCST/SCST) for various NSS rounds. Panel (b) shows the median wage gaps (non-SCST/SCST) by age groups for various NSS rounds.

Next we examine the education patterns of the two groups during this period. Figure 2 shows the relative gaps in the years of education between non-SC/STs and SC/STs. Panel (a) of the Figure shows the gaps for different age groups while panel (b) shows the corresponding gaps in the average years of schooling by birth cohorts. Again, both panels reveal the same pattern of convergence in education attainment rates between the two groups. In fact, the education convergence trends are even sharper than the trends in wage convergence.

Given the trends in Figures 1 and 2, the natural question to ask is how much of the wage convergence between the two groups is due to convergence in education attainment. In Hnatkovska, Lahiri, and Paul (2012b) we examine precisely this question and find that most of the wage convergence is, in fact, due to education convergence.

These trends, while interesting by themselves, raise the logical question about the deeper reasons behind the observed convergence between the groups during this period. While there may have been multiple factors operating simultaneously, in this paper we focus on the two biggest changes that occurred in the Indian economy during this period. As is well known, this period – 1983 to 2007-08 – has also been a period of major changes in economic policy accompanied by a sharp

Figure 2: Gaps in years of education: overall and by birth cohorts



(a) (b)

Notes: Panel (a) of this Figure shows the relative gap in average years of education (non-SCST/SCST) across the NSS rounds for different age groups while Panel (b) shows the gaps by birth cohorts.

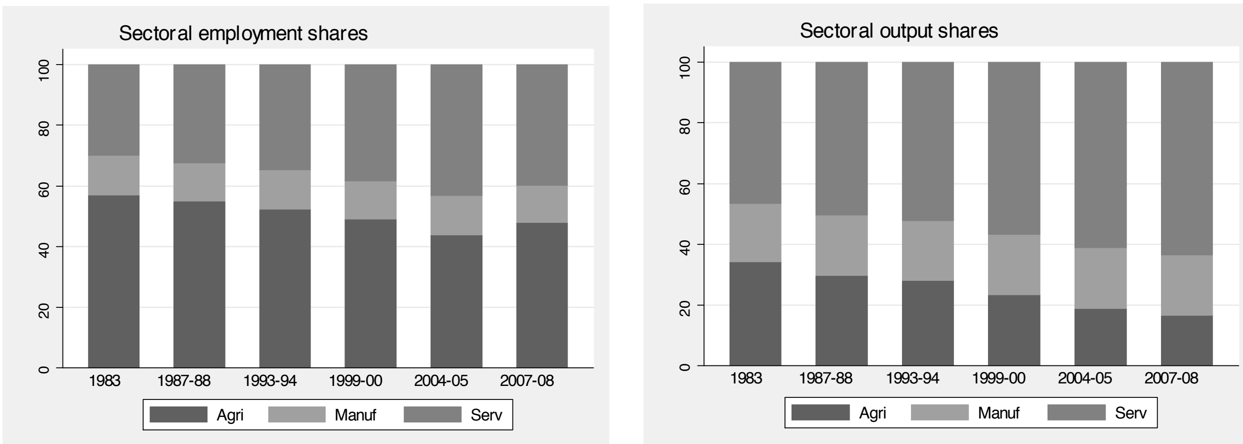
economic take-off in India. There were large scale trade and industrial reforms carried out in the mid-1980s and in the 1990s. Economic growth in India took off from an average of around 3 percent in the period between 1950 and 1985 to consistently being above 6 percent by the end of the 1990s. Second, this period was also marked by a very sharp structural transformation of the economy. The primary question we address is whether the aggregate productivity improvement could have induced a wage and education convergence across the castes through the structural transformation that it sparked?

Before proceeding, it is useful to document some of the key data facts related to the structural transformation of the economy since the early 1980s as well as a breakdown by caste of these structural changes. In order to present the structural transformation facts, we combine industry categories reported in the NSS by each individual into three broad industry categories: Agricultural, Manufacturing, and Services. Our grouping reflects the traditional industrial classification according to the United Nations classification system. See Appendix 8 for more details on the industry grouping.

As Figure 3 shows, the period was marked by a gradual contraction in the traditional agricultural sector while the service sector expanded both in terms of its share of output as well as employment (there was an expansion in the manufacturing sector too but much more tepid relative to that of the service sector).

This process of structural transformation coincided with rapid growth in both labor productivity and total factor productivity (TFP) at the aggregate and sectoral levels. Figure 4 reports labor

Figure 3: Industry distribution

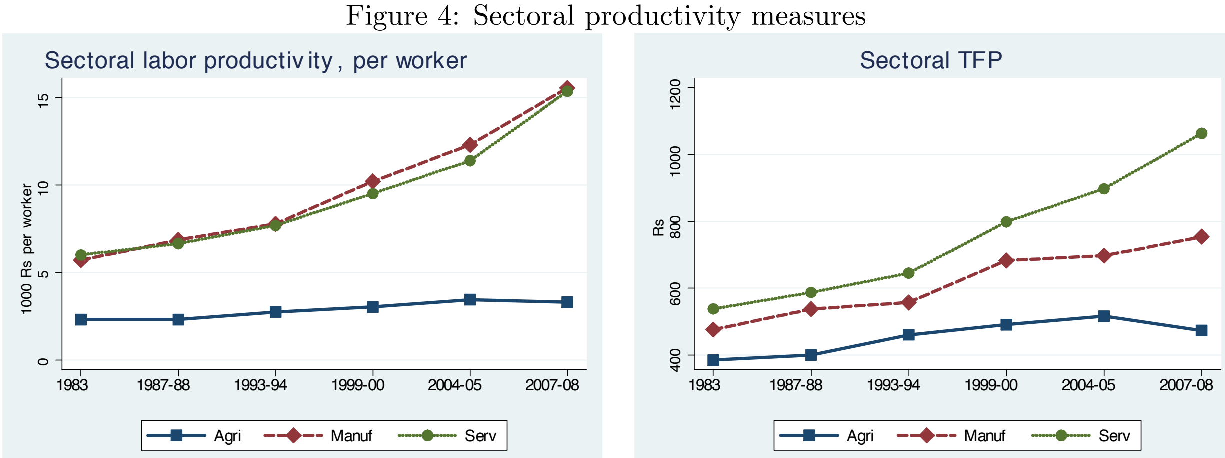


(a) (b)

Notes: Panel (a) of this Figure presents the distribution of workforce across three industry categories for different NSS rounds. Panel (b) presents distribution of output (measured in constant 1980-81 prices) across three industry categories.

productivity in each sector. Panel (a) is measured as output per worker, while panel (b) reports the sectoral total factor productivity numbers that we estimated assuming a Cobb-Douglas production function including capital, human capital and employment.[2] The figures show a common feature of productivity growth across the three sectors, especially in TFP.

Figure 4: Sectoral productivity measures



(a) (b)

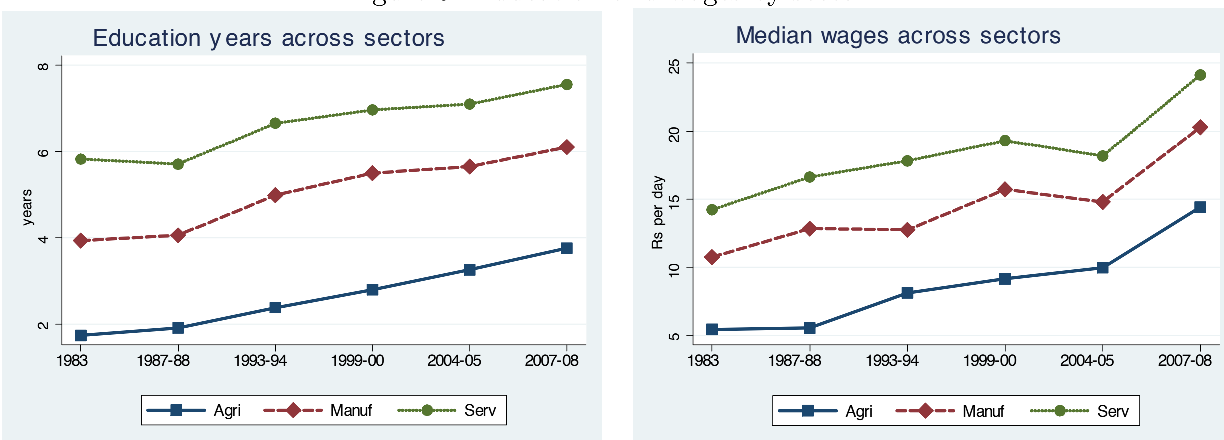
Notes: Panel (a) of this Figure presents labor productivity, measured as GDP (in constant 1980-81 prices) divided by number of workers in each sector. Panel (b) shows the sectoral total factor productivity by using a Cobb-Douglas production function for each sector using sectoral capital and labor.

Figure 5 reports mean years of education and median wages in the three sectors for various

[2]The sectoral human capital stocks were constructed using standard Mincer regressions by using the education attainment rates of the sectoral workforce.

survey rounds. The figures reveal a dramatic increase in both education attainments and median wages in India during 1983-2008 period.

Figure 5: Education and wages by sector



(a) (b)

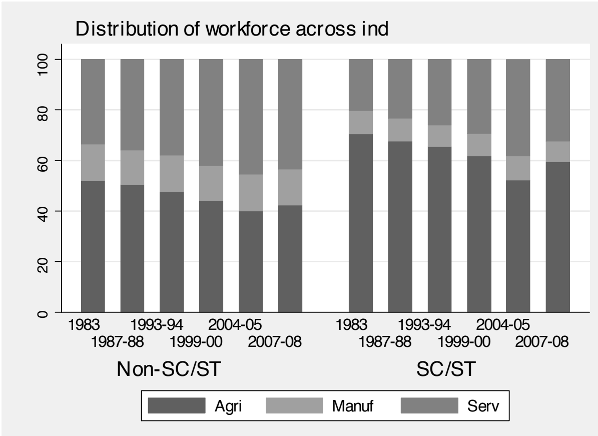
Notes: Panel (a) of this Figure presents average years of education of workers employed in each of the three sectors. Panel (b) reports median wages in the three sectors.

So, how did this overall transformation of the economy affect the two groups? Figure 6 reports the industry distribution of working individuals among SC/STs and non-SC/STs, and the relative gaps in this distribution. Clearly, SC/STs were and remain more likely to be employed in agriculture and other farming activities (agri) than non-SC/STs. However the gap has somewhat narrowed in the last ten years of our sample. The second largest industry of employment for both social group is services, whose share has also been rising steadily over time. Interestingly, services also exhibit the sharpest convergence pattern between non-SC/STs and SC/STs followed by agriculture. In particular, the relative gap between non-SC/STs and SC/STs in employment shares in services has shrunk from more than 50 percent in 1983 to below 25 percent in 2007-08. Manufacturing shows little changes in the employment shares of the two groups over time.

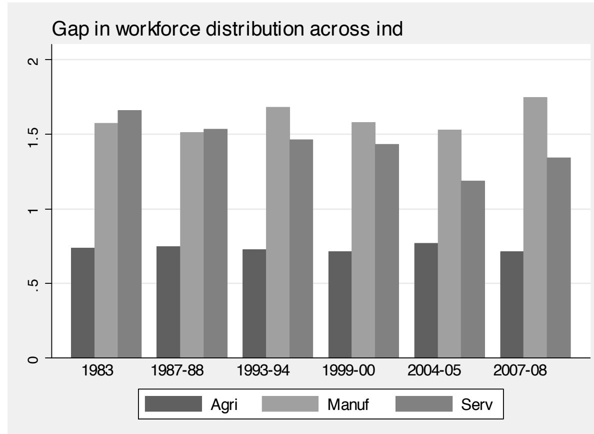
Figures 7 report the relative gaps in education attainments and median wages between non-SC/STs and SC/STs employed in each sector. The education gaps have narrowed significantly over time between the two caste groups. Median wage gaps on the other hand declined in Services, stayed unchanged in Manufacturing, but widened in Agriculture.

To summarize the data features documented above, the period 1983-2008 was characterized by high aggregate growth in the economy, rising output per worker in all three sectors and similar productivity growth across the sectors. Concurrently, there was a gradual transformation of the economy that was underway as well with services becoming a larger share of the economy both

Figure 6: Industry employment distribution across castes



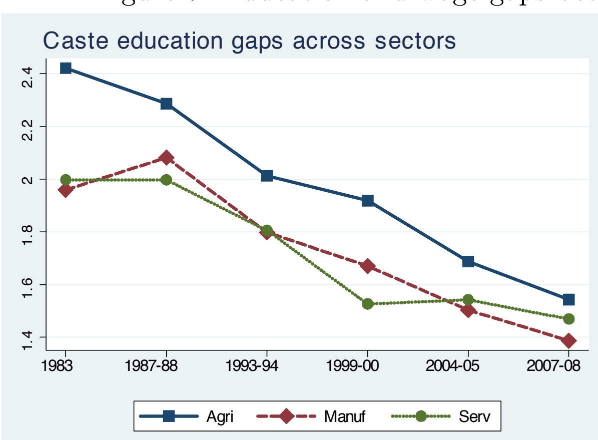
(a)



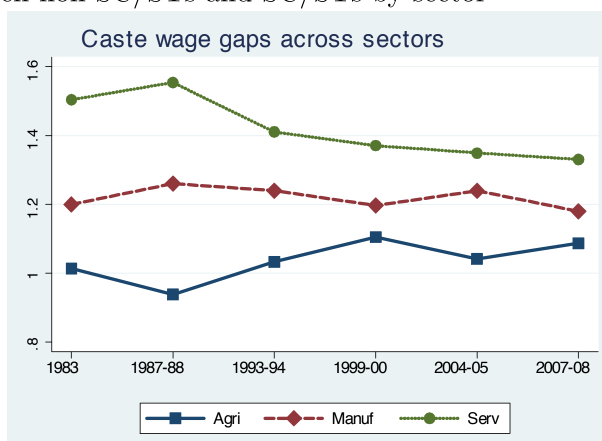
(b)

Notes: Panel (a) of this Figure presents the distribution of workforce across the 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 relative gaps in the distribution of non-SC/STs relative to SC/STs across three industry categories.

Figure 7: Education and wage gaps between non-SC/STs and SC/STs by sector



(a)



(b)

Notes: Panel (a) of this Figure presents relative gap in years of education between non-SC/STs and SC/STs. Panel (b) presents the ratio of non-SC/STs median wages to SC/STs median wages.

in terms of output and employment while the corresponding agriculture shares were shrinking. In terms of the caste distributions, both SC/STs and non-SC/STs appeared to be exiting from agriculture and moving into service sector employment during this period. The education gap between the castes declined in all three sectors. Moreover, while wages were converging *overall* between the castes, there were interesting contrasts in the patterns across the sectors. The wage convergence was strong in the service sector. The agricultural sector however saw a divergence in wages between the castes. Interestingly, the median wage gaps in the manufacturing sector stayed relatively unchanged over this period.

3 Empirics

The preceding results were all based on an unconditional examination of the data. The key question is whether these patterns are statistically significant. We now try to address this using some formal econometric tests. Two questions are of particular interest to us: what factors account for the time series behavior of the wage gap? Has the wage gap been narrowing symmetrically across sectors? To address these questions we introduce controls for the sector in which the individual worked. Our empirical method is structured to attempt to deduce the relative importance of aggregate versus sectoral factors underlying the observed wage patterns of the two groups. We use household survey data between 1983 and 2007-08 to construct a synthetic panel of households and individuals by combining the cross-sections from all NSS rounds. We then evaluate the changes in the relative contributions of the factors listed above for the wage trends.

We proceed sequentially. We start with a specification with only caste and round controls. We then introduce sectoral dummies, and interactions between the caste, sector and round dummies. The goal is to uncover the sectoral and aggregate features of the dynamics of the wage gap. We examine both the behavior of mean wages and median wages. For mean wages we use the standard OLS procedure while for median wages we use the Recentered Influence Function (RIF) regressions which were formalized in Firpo, Fortin, and Lemieux (2009).

Our initial specification is a regression of wages on a constant, a caste dummy, round dummies and interactions between the caste and round dummies for the different rounds of the survey. The base group in this regression are the non-SC/STs in 1983. Table 1 reports the results. Regressions (1) reports results for median wages while specification (3) pertains to mean wages. The SC/ST dummy is negative and significant for both specifications. However, the interaction term between caste and round always enters positively, both for median and mean wage gaps. Crucially, the sum of the coefficients on caste and the caste-round interaction are negative for all rounds. Hence, SC/STs had significantly lower median and mean wages relative to non-SC/STs in all the rounds. However, relative to the gap in 1983, the median and mean wage gaps declined significantly over the subsequent time period since the interactive terms between the caste and round dummies are all significantly positive and became larger over successive rounds. The decline in the SC/ST wage gap is more pronounced for the median than for the mean wages, suggesting important distributional changes within each social group taking place over time.

Are the changes in the wage gap over time symmetric across sectors? Given the large scale sectoral transformation that has been occurring in India during the sample period, controlling for sectoral developments may clearly be important. Accordingly, the specifications in regressions (2) (for median wages) and (4) (for mean wages) include sectoral dummies along with interactions of sector and caste, sector and round, and sector, round and caste. Given the dummies, all the coefficients in regressions (2) and (4) are interpreted relative to the base which is the wage in round 38 (i.e., in 1983) of a non-SC/ST individual employed in agriculture. Adding the SC/ST dummy and the interaction term between SC/ST and round gives the wage gap in agriculture in the concerned round. The wage gap in any other sector in that round is derived by adding the coefficients on the SC/ST-sector and SC/ST-sector-round interactions. Lastly, the change in the wage gap in any sector between 1983 and any other round is given by the sum of the coefficients on the SC/ST-round and the SC/ST-sector-round interactions.

Two features of the results are worth noting. First, the wage gaps were positive in all sectors in all rounds, i.e., non-SC/STs received a higher wage than SC/STs across the board. Second, across the survey rounds, the caste wage gap was the lowest in agriculture. Third, over the sample period, the wage gap in services declined by almost 35 percentage points. In agriculture however, the caste wage gap *widened* by almost 10 percentage points. The gap in manufacturing changed very little and it was not significant. Importantly, these basic patterns were similar for both median and mean wage gaps. The sectoral changes in the median and mean wage gaps between 1983 and 2008 are summarized in Table 2.

Overall, our results suggest that during the sample period, the average caste wage gap has declined significantly in services, stayed constant in manufacturing, and widened significantly in the agricultural sector. Clearly, the overall picture of declining wage gaps tends to mask important sectoral heterogeneity underneath the overall trends. Hence, any explanation of the changes in the overall caste wage gap during this period has to both take into account the sectoral developments in India as well as rationalize the sectoral heterogeneity in the behavior of the wage gap.

4 Explaining the Patterns

The preceding patterns in the data raise the natural question regarding the potential explanations for them. We now turn to this issue. In formalizing a structure to explain the facts, we take as a

Table 1: Wage Regressions: Time and sectoral effects

	median wage	median wage	mean wage	mean wage
	(1)	(2)	(3)	(4)
SC/ST	-0.373***	-0.023***	-0.366***	-0.042***
	(0.008)	(0.006)	(0.008)	(0.008)
SC/ST*round50	0.021*	-0.024**	0.025**	-0.036***
	(0.011)	(0.010)	(0.011)	(0.013)
SC/ST*round55	0.086***	-0.030**	0.044***	-0.021**
	(0.012)	(0.013)	(0.012)	(0.011)
SC/ST*round61	0.109***	-0.039**	0.063***	-0.009
	(0.014)	(0.017)	(0.012)	(0.011)
SC/ST*round64	0.217***	-0.090***	0.075***	-0.024**
	(0.012)	(0.016)	(0.011)	(0.011)
SC/ST*Manufacturing		-0.103***		-0.131***
		(0.027)		(0.023)
SC/ST*Services		-0.367***		-0.327***
		(0.019)		(0.018)
SC/ST*round50*Manufacturing		-0.132***		0.023
		(0.039)		(0.033)
SC/ST*round50*Services		0.194***		0.122***
		(0.027)		(0.026)
SC/ST*round55*Manufacturing		0.008		0.013
		(0.042)		(0.036)
SC/ST*round55*Services		0.297***		0.156***
		(0.028)		(0.027)
SC/ST*round61*Manufacturing		-0.093**		-0.101***
		(0.044)		(0.035)
SC/ST*round61*Services		0.278***		0.104***
		(0.030)		(0.025)
SC/ST*round64*Manufacturing		0.144***		-0.006
		(0.039)		(0.032)
SC/ST*round64*Services		0.436***		0.112***
		(0.028)		(0.024)
N	323590	323590	323590	323590
R-sq	0.091	0.238	0.114	0.278

Notes: The table reports regressions of log mean and median wages on covariates that vary across specifications. Regressions (1) and (2) report results for median wages while (3) and (4) report estimates for mean wages. Regressions (1) and (3) include a constant, caste and round dummies along with interactions of the caste and round dummies. Regressions (2) and (4) add in controls for sectors of employment, and interactions of sector and caste, sector and round as well as interactions between sector, round and caste. Standard errors are reported in parenthesis. * p-value≤0.10, ** p-value≤0.05, *** p-value≤0.01. We report a subset of the estimates. The complete results are available from the authors upon request.

starting point two data features from Section 2. The first is that there was sharp convergence in education attainment levels between the groups. the second is that there were common trends in productivity across the three sectors which we interpret as evidence for the presence of aggregate factors that impacted all sectors simultaneously during the period 1983-2008. Hence, we shall build a model which has human capital acquisition as a fundamental component. Moreover, the test of the structure that we shall formalize here is that it has to be consistent with the heterogeneity in the sectoral dynamics for the caste wage gaps that we have documented above.

Table 2: Changes in Sectoral Wage Gaps: 1983-2008

	Median	Mean
Agriculture	0.09***	0.024**
	(0.016)	(0.011)
Manufacturing	-0.054	0.030
	(0.042)	(0.034)
Services	-0.346***	-0.088***
	(0.032)	(0.026)

The table reports the change in the sectoral wage gaps between 1983 and 2008. The numbers are computed by adding the estimated coefficient on the SC/ST dummy to the coefficients on the interaction term between the SC/ST and 64th round dummies, and the interaction between the SC/ST, 64th round and sectoral dummies reported in Table 1. Standard errors are reported in parenthesis. * p-value≤0.10, ** p-value≤0.05, *** p-value≤0.01.

4.1 The model

We formalize a model with heterogenous ability of the workforce. Consider a one-period lived closed economy that is inhabited by a continuum of agents of measure L. A measure S of these agents belong to caste s (for scheduled castes and tribes or SC/STs) while a measure $N = L - S$ belong to caste n for non-SC/ST. Each agent i maximizes utility from

$$u\left(c_i\right) = \frac{c_i^{1-\rho}}{1-\rho}$$

where

$$c_i = \left(c_i^a - \bar{c}\right)^{\theta}\left(c_i^m\right)^{\eta}\left(c_i^h\right)^{1-\theta-\eta}$$

and c^{κ}, $\kappa = \{a, m, h\}$ denotes consumption of good κ. In the following, we shall refer to the a good as the agricultural good, the m good as the manufacturing good and the h good as the high skill good. $\bar{c}$ is the minimum level of consumption of the a good.

Each agent i is born with one unit of labor time that is supplied inelastically to the market and an endowment of ability e_i. The ability distribution is caste specific. So, for agents belonging to caste s the ability e_i is drawn from an i.i.d. process that follows the cumulative distribution function $G_s\left(e\right), \quad e \in [\underline{e}_s, \bar{e}^s]$. Similarly, for agents belonging to caste n the ability type is drawn from an i.i.d. process summarized by the distribution function $G_n\left(e\right), \quad e \in [\underline{e}_n, \bar{e}^n]$. In the following we shall retain the assumptions

Assumption 1: $\underline{e}_s \leq \underline{e}_n$

Assumption 2: $\bar{e}^s \leq \bar{e}^n$

Assumptions 1 and 2 imply that caste s members draw their ability types from a distribution with lower levels of both the lower and upper supports relative to caste n members. This is intended to be a stand-in for the fact that centuries of discrimination against the lower castes left them with little to no education which, in turn, was perpetuated through the generations. In as much as education affects ability and some of that can be passed across generations, this may have left current generation of SC/STs with lower inherited ability levels, or shifted the ability distribution to the left.[3]

Ability is a productive input in both production and in skill acquisition. An agent can work in either of the three sectors. Sector a does not require any training or special skills, hence agents who choose to work in this sector can supply their labor endowment to this sector as is. Working in sector m requires some special, sector-specific skill. Agent i can acquire this skill by spending f_{ji}^{m} units of the sector m good where $j = s, n$ denotes the caste to which agent i belongs. This specification allows the skill acquisition costs to be caste specific. Similarly, to work in sector h the worker i needs to acquire a different skill level which can be acquired by expending f_{ji}^{h} units of the m good. In the following we shall assume that the costs of acquiring skills are decreasing in the ability level of the individual:

Assumption 3: $f_{ji}^{k} = f_{j}^{k}\left(e_i\right), \quad f_{j}^{k\prime} \leq 0, \quad j = s, n, \quad k = m, h.$

The technologies for producing the three goods are all linear in the labor input. In particular, an unskilled worker with ability e_i supplying one unit of labor time to sector a produces

$$y_i^a = Ae_i.$$

An m-sector worker with ability e_i produces the manufacturing good m according to

$$y_i^m = Me_i.$$

Lastly, an $h-$sector worker with ability e_i produces the high skill good according to

$$y_i^h = He_i.$$

Here A, M and H denote productivity levels in agriculture, manufacturing and high-skill sectors,

[3] The evolution of these supports of the distributions is clearly a dynamic issue and endogenous to time and investment decisions regarding education that are made by families. We intend to address these issues more fully in future work.

respectively. Note that labor supply is inelastic and indivisible, so each worker supplies one unit of labor time to whichever sector he/she works in.

Let p_m denote the relative price of good m and p_h denote the relative price of good h. Throughout we shall use good a as the numeraire. The budget constraint of worker i is then given by

$$c_i^a + p_m c_i^m + p_h c_i^h = \hat{y}_i,$$

where

$$\hat{y}_i = \max\left\{y_i^a, p_m\left(y_i^m - f^m\left(e_i\right)\right), p_h y_i^h - p_m f^h\left(e_i\right)\right\}.$$

The subscript i refers to household i with ability e_i. Recall that each worker will be working in only one sector.

The optimality conditions governing consumption of the three goods for household i are

$$\begin{aligned} \frac{\eta}{\theta}\left(\frac{c_i^a - \bar{c}}{c_i^m}\right) &= p_m \\ \frac{(1-\eta-\theta)}{\theta}\left(\frac{c_i^a - \bar{c}}{c_i^h}\right) &= p_h \end{aligned}$$

Using these solutions along with the household's budget constraint gives

$$\begin{aligned} c_i^a &= \bar{c} + \theta\left(\hat{y}_i - \bar{c}\right) \\ p_m c_i^m &= \eta\left(\hat{y}_i - \bar{c}\right) \\ p_h c_i^h &= (1-\eta-\theta)\left(\hat{y}_i - \bar{c}\right) \end{aligned}$$

4.2 Occupation and Skill Choice

The decisions about which occupation to choose and what skill level to acquire are joint in this model since skills are matched uniquely to sectors. Thus, an agent of caste j with ability e_i will choose to remain unskilled and work in sector a if and only if

$$Ae_i \geq p_m\left(Me_i - f_j^m\left(e_i\right)\right)$$

and

$$Ae_i \geq p_h He_i - p_m f_j^h\left(e_i\right).$$

The two conditions can be rewritten as

$$\begin{aligned}
\frac{f_j^m(e_i)}{e_i} &\geq M - \frac{A}{p_m}, \\
\frac{f_j^h(e_i)}{e_i} &\geq \frac{p_h}{p_m}H - \frac{A}{p_m}.
\end{aligned}$$

The right hand sides of these two conditions are the relative gains from working in sector m or h while the left hand sides are the relative costs. Crucially, the right hand side variables are aggregate variables that private agents take as given. The left hand sides, on the other hand, are individual specific and are clearly decreasing functions of e_i. Hence, these two conditions define two cutoff thresholds:

$$z_j^m\left(\hat{e}_j^m\right) = M - \frac{A}{p_m}, \quad j = s, n \tag{4.1}$$

$$z_j^h\left(\hat{e}_j^h\right) = \frac{p_h}{p_m}H - \frac{A}{p_m}, \quad j = s, n, \tag{4.2}$$

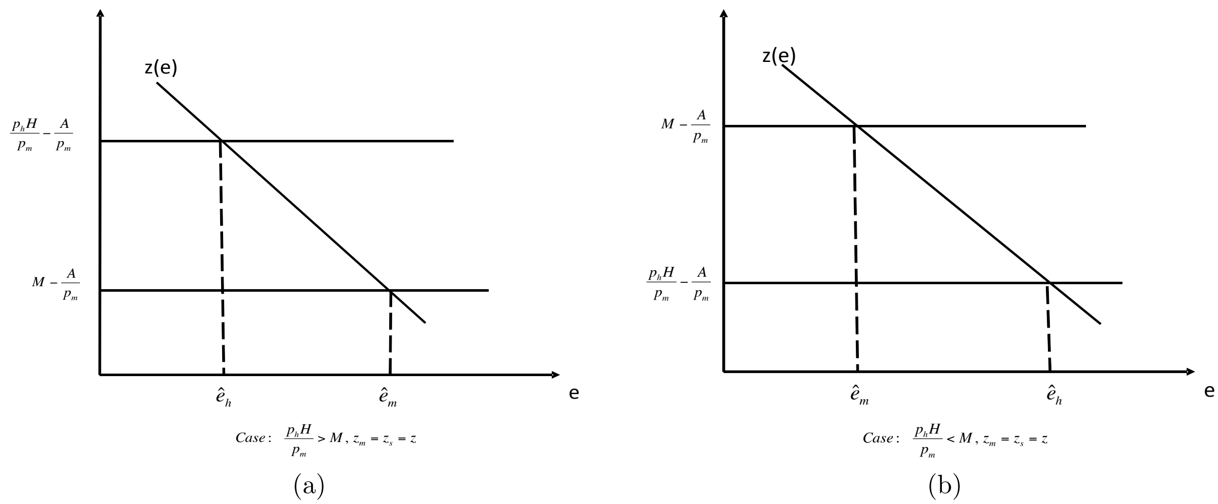
where we have used the definitions $z_j^m(e) \equiv \frac{f_j^m(e)}{e}$ and $z_j^h(e) \equiv \frac{f_j^h(e)}{e}$. Since z_j^k is decreasing in e for $j = m, h$ and $k = s, n$, all ability types above $\hat{e}_m^j$ will choose to acquire skill m and work in sector m while all types i with e_i above $\hat{e}_h^j$ will choose to work in sector h by acquiring skill h instead of staying unskilled and working in sector a. The rest will remain unskilled and work in sector a. Note that since the right hand sides of the threshold conditions given by equations (4.1) an d(4.2) are not caste specific (they are aggregate variables), the threshold conditions also imply that

$$\begin{aligned}
z_s^m\left(\hat{e}_s^m\right) &= z_n^m\left(\hat{e}_n^m\right) \\
z_s^h\left(\hat{e}_s^h\right) &= z_n^h\left(\hat{e}_n^h\right)
\end{aligned}$$

In order to characterize the distribution of the different ability types in the three sectors, it is important to note that there are four possible configurations of cases: (i) $\frac{p_h H}{p_m} \geq \frac{A}{p_m}$ and $z_j^h(e) \geq z_j^m(e)$; (ii) $\frac{p_h H}{p_m} \geq \frac{A}{p_m}$ and $z_j^h(e) < z_j^m(e)$; (iii) $\frac{p_h H}{p_m} < \frac{A}{p_m}$ and $z_j^h(e) \geq z_j^m(e)$; and (iv) $\frac{p_h H}{p_m} < \frac{A}{p_m}$ and $z_j^h(e) < z_j^m(e)$. Figure 8 depicts the ability thresholds to get skilled and work in a non-agricultural sector in the special case of $f_j^h = f_j^m = f_j$. This collapses cases (i) and (ii) into one category and (iii) and (iv) into another leaving us with two cases for each caste. The two cases are shown in the two panels of Figure 8.

Panel (a) of the Figure shows the case in which $\hat{e}_j^m > \hat{e}_j^h$. Intuitively, when the $h-$sector is more productive than the $m-$sector, relatively low ability types find it profitable to pay the higher cost of skill acquisition in order to work in the $h-$sector even though they do not find it profitable to get skilled to work in the $m-$sector where productivity is lower. As a result, the cutoff threshold for the $h-$sector is lower than the $m-$sector where only the very high ability types find it profitable to invest and work. Panel (b) shows the opposite configuration.

Figure 8: Ability Cutoffs for Sectoral Allocation



4.3 Market clearing and Equilibrium

Markets for each good must clear individually. Hence, we must have

$$c^a = y^a \tag{4.3}$$

$$c^m = y^m - F \tag{4.4}$$

$$c^h = y^h \tag{4.5}$$

where F denotes the total skill acquisition costs incurred by workers employed in sector m and sector h respectively. The market clearing condition for the m good recognizes that part of the use of the good is for acquiring skills. We shall derive the exact expression for F below.

Definition: *The Walrasian equilibrium for this economy is a vector of prices $\{p_m, p_h\}$ and quantities $\left\{c^a, c^m, c^h, y^a, y^m, y^h, F^m, F^h,\ \hat{e}_s^m, \hat{e}_s^h, \hat{e}_n^m, \hat{e}_n^h\right\}$ such that all worker-households satisfy their* *optimality conditions, budget constraints are satisfied and all markets clear.*

4.4 Aggregation

There are six variables to aggregate – aggregate consumption of the three types of goods as well as their aggregate productions. We start with the consumption side. Aggregate consumption of each good is the sum of the consumptions of the good by each type of household. Recall that there are three types of households in this economy and that each household opts into only one of the three available occupations. The solution for c_i^a derived above says that

$$c_i^a = (1-\theta)\,\bar{c} + \theta \hat{y}_i.$$

Using the net income of each type of worker, $\hat{y}_i$, aggregate consumption of sector a goods are given by

$$\frac{c_a}{1-\theta}=\frac{\theta}{1-\theta}\sum_{j=s,n} s_j\left[\int_{\underline{e}_j}^{\hat{e}_j^m} Ae_i dG_j\left(e\right)+p_m\int_{\hat{e}_j^m}^{\hat{e}_j^h}\left\{Me_i-f_j^m\left(e_i\right)\right\}dG\left(e\right)+\int_{\hat{e}_j^h}^{\bar{e}_j}\left\{p_hHe_i-p_mf_j^h\left(e_i\right)\right\}dG_j\left(e\right)\right]+L\bar{c} \tag{4.6}$$

where $s_s = S$, and $s_n = N$ represent the population sizes of the two two castes. Similarly, we have $p_m c_{mi} = \eta\left(\hat{y}_i - \bar{c}\right)$ which when aggregated across all households gives

$$\frac{p_m c_m}{\eta}=\sum_{j=s,n} s_j\left[\int_{\underline{e}_j}^{\hat{e}_j^m} Ae_i dG_j\left(e\right)+p_m\int_{\hat{e}_j^m}^{\hat{e}_j^h}\left\{Me_i-f_j^m\left(e_i\right)\right\}dG\left(e\right)+\int_{\hat{e}_j^h}^{\bar{e}_j}\left\{p_hHe_i-p_mf_j^h\left(e_i\right)\right\}dG_j\left(e\right)\right]-L\bar{c} \tag{4.7}$$

Lastly, since $p_h c_{hi} = (1-\eta-\theta)\left(\hat{y}_i - \bar{c}\right)$, aggregate consumption of the h good is

$$\begin{aligned}\frac{p_h c_h}{(1-\eta-\theta)} &= \sum_{j=s,n} s_j\left[\int_{\underline{e}_j}^{\hat{e}_j^m} Ae_i dG_j\left(e\right)+p_m\int_{\hat{e}_j^m}^{\hat{e}_j^h}\left\{Me_i-f_j^m\left(e_i\right)\right\}dG\left(e\right)+\int_{\hat{e}_j^h}^{\bar{e}_j}\left\{p_hHe_i-p_mf_j^h\left(e_i\right)\right\}dG_j\left(e\right)\right] \\ &\quad -L\bar{c}\end{aligned} \tag{4.8}$$

Next, the expected output of a worker i of caste j in each sector is given by

$$\bar{w}_j^a = Ey_{ji}^a = \int_{\underline{e_j}}^{\hat{e}_m^j} Ae_i \frac{dG_j\left(e\right)}{G_j\left(\hat{e}_j^m\right)} \tag{4.9}$$

$$\bar{w}_j^m = Ey_{ji}^m = \int_{\hat{e}_m^j}^{\hat{e}_h^j} Me_i \frac{dG_j(e)}{G_j\left(\hat{e}_j^h\right) - G_j\left(\hat{e}_j^m\right)} \tag{4.10}$$

$$\bar{w}_j^h = Ey_{ji}^h = \int_{\hat{e}_h^j}^{\bar{e}_j} He_i \frac{dG_j(e)}{1 - G_j\left(\hat{e}_j^h\right)} \tag{4.11}$$

Given the linearity of the production technologies in ability, these are just the conditional means of the relevant distributions of ability in each sector.

Using these expected outputs, the aggregate output of each sector is

$$y^\kappa = \sum_{j=s,n} y_j^\kappa \quad , \quad \kappa = a, m, h$$

where κ indexes the sector. Clearly, $y_j^a = s_j G_j\left(\hat{e}_j^m\right)\bar{w}_j^a$, $y_j^m = s_j\left[G_j\left(\hat{e}_j^h\right) - G_j\left(\hat{e}_j^m\right)\right]\bar{w}_j^m$ and $y_j^h = s_j\left[1 - G_j\left(\hat{e}_j^h\right)\right]\bar{w}_j^h$, where $j = s, n$ and $s_j = S, N$. Substituting in the relevant expressions gives

$$y^a = S\int_{\underline{e}_s}^{\hat{e}_s^m} Ae_i dG_s(e) + N\int_{\underline{e}_n}^{\hat{e}_n^m} Ae_i dG_n(e) \tag{4.12}$$

$$y^m = S\int_{\hat{e}_s^m}^{\hat{e}_s^h} Me_i dG_s(e) + N\int_{\hat{e}_n^m}^{\hat{e}_n^h} Me_i dG_n(e) \tag{4.13}$$

$$y^h = S\int_{\hat{e}_s^h}^{\bar{e}_s} He_i dG_s(e) + N\int_{\hat{e}_n^h}^{\bar{e}_n} He_i dG_n(e) \tag{4.14}$$

Note that the aggregation above represents case (b) of Figure 8 above where $\hat{e}_j^h > \hat{e}_j^m$. The limits of integration would have to be altered appropriately for the opposite configuration.

To derive the skill acquisition costs, note that the average costs of skill acquisition by caste j conditional on the sector of employment is

$$\begin{aligned} F_j^m &= \int_{\hat{e}_j^m}^{\hat{e}_j^h} f_j^m(e_i)\frac{dG_j(e)}{G_j\left(\hat{e}_j^h\right) - G_j\left(\hat{e}_j^m\right)} \\ F_j^h &= \int_{\hat{e}_j^h}^{\bar{e}_s} f_j^h(e_i)\frac{dG_j(e)}{1 - G_j\left(\hat{e}_j^h\right)} \end{aligned}$$

Hence, total skill acquisition cost of caste j is

$$F_j = s_j\left[\left\{G_j\left(\hat{e}_j^h\right) - G_j\left(\hat{e}_j^m\right)\right\}F_j^m + \left\{1 - G_j\left(\hat{e}_j^h\right)\right\}F_j^h\right], \quad j = s, n$$

where $s_s = S$ and $s_n = N$. Summing the costs across the castes then gives the total cost of acquiring skills by the different groups as

$$F = S\left[\int_{\hat{e}_s^m}^{\hat{e}_s^h} f_s^m(e_i)\,dG_s(e) + \int_{\hat{e}_s^h}^{\bar{e}_s} f_s^h(e_i)\,dG_s(e)\right] + N\left[\int_{\hat{e}_n^m}^{\hat{e}_n^h} f_n(e_i)\,dG_n(e) + \int_{\hat{e}_n^h}^{\bar{e}_n} f_n^h(e_i)\,dG_s(e)\right] \tag{4.15}$$

These relationships can then be used in equations (4.3-4.5) to derive the specific market clearing conditions for the a, m, and h goods. Note that only two of the three market clearing conditions are free - if two markets clear then the third must clear as well.

The equilibrium for the economy can be reduced to a system of four equations in four unknowns, $\hat{e}_s^m, \hat{e}_s^h, \hat{e}_n^m, \hat{e}_n^h$. The four equations that jointly determine these variables are two out of the three market clearing conditions along with the conditions $z_s^m(\hat{e}_s^m) = z_n^m(\hat{e}_n^m)$ and $z_s^h\left(\hat{e}_s^h\right) = z_n^h\left(\hat{e}_n^h\right)$. Note that the prices p_m and p_h can be eliminated from this system of equations by using any two of the four threshold conditions $z_j^m\left(\hat{e}_j^m\right) = M - \frac{A}{p_m}$ and $z_j^h\left(\hat{e}_j^h\right) = \frac{p_h}{p_m}H - \frac{A}{p_m}, \quad j = s, n.$

4.5 Thresholds and Wage Gaps

The key endogenous variables in this model are the four threshold ability levels $\hat{e}_j^m$ and $\hat{e}_j^h$ for $j = s, n$. The sectoral and overall wage gaps are all functions of these four thresholds. We now illustrate the relationship between these thresholds and the sectoral wage gaps under a special case for the skill cost function and the ability distribution. Specifically, while we retain Assumptions 1 and 2 so that $\underline{e}_s \leq \underline{e}_n$ and $\bar{e}^s \leq \bar{e}^n$, we impose the additional assumptions:

Assumption 4: The skill acquisition cost is given by $f_j^k(e) = \phi\left(\gamma_j^k - \alpha e\right)$ for $j = s, n$ and $k = m, h$ with $\gamma_j^k > \alpha \bar{e}^j$.
Assumption 5: $\frac{\gamma_j^h}{\gamma_j^m} = \beta$ for $j = s, n, \;\; \beta > 0$
Assumption 6: $G_j(e)$ is uniform on the support $\left[\underline{e}_j, \bar{e}^j\right]$ for $j = s, n$.

Assumption 4 imposes linearity on the skill cost function with the marginal effect of ability on the cost assumed to be identical for both castes and sectors. Crucially though, the specification allows the intercept term on the cost function to vary by sector and caste. The condition $\gamma_j > \alpha \bar{e}^j$ ensures that getting skilled involves a positive cost for even the highest ability type. Assumption 5 says that the proportional difference in the fixed costs of training between the m and h sectors are identical for the two castes. Assumption 6 incorporates the uniform distribution for ability which just makes the analytics simple.

Under this formulation, it is easy to check that

$$
\begin{aligned}
\frac{\hat{e}_n^m}{\hat{e}_s^m} &= \frac{\gamma_n^m}{\gamma_s^m} \\
\frac{\hat{e}_n^h}{\hat{e}_s^h} &= \frac{\gamma_n^h}{\gamma_s^h}
\end{aligned}
$$

In other words, the relative sectoral ability thresholds of the two groups are proportional to their relative fixed costs of acquiring skills to work in that sector. Crucially, the conditions say that the ability cutoff of caste n for working in sector $\kappa = m, h$ will be greater than the corresponding cutoff for caste s if and only if their fixed skill costs exceed the corresponding cost for caste s.

Using this proportional relationship, we can determine the relationship between the thresholds and the sectoral wage gaps. We define the sectoral wage gaps as $\Delta w^a = \frac{\bar{w}_n^a}{\bar{w}_s^a}, \Delta w^m = \frac{\bar{w}_n^m}{\bar{w}_s^m}$ and $\Delta w^h = \frac{\bar{w}_n^h}{\bar{w}_s^h}$. Further, we assume that parameters are such $\hat{e}_j^m < \hat{e}_j^h$ for $j = s, n$. Hence, we are assuming that the $h-$sector cutoff is always greater for both groups. This reflects the fact (as we showed in the empirical section earlier) that average education levels in the service sector are always higher than in the manufacturing sector for both groups and in all the survey rounds.

A key point of interest for us is the relative wage gap between SC/STs and non-SC/STs. Using equations (4.9) and (4.10) and evaluating them under the assumption of our model gives the sectoral wage gaps:

$$
\begin{aligned}
\frac{\bar{w}_n^a}{\bar{w}_s^a} &= \Delta w^a = \frac{\hat{e}_n^m + \underline{e}_n}{\hat{e}_s^m + \underline{e}_s} = \frac{\frac{\gamma_n^m}{\gamma_s^m}\hat{e}_s^m + \underline{e}_n}{\hat{e}_s^m + \underline{e}_s} \\
\frac{\bar{w}_n^m}{\bar{w}_s^m} &= \Delta w^m = \frac{\hat{e}_n^h + \hat{e}_n^m}{\hat{e}_s^h + \hat{e}_s^m} = \frac{\frac{\gamma_n^h}{\gamma_s^h}\hat{e}_s^h + \frac{\gamma_n^m}{\gamma_s^m}\hat{e}_s^m}{\hat{e}_s^h + \hat{e}_s^m} \\
\frac{\bar{w}_n^h}{\bar{w}_s^h} &= \Delta w^h = \frac{\bar{e}_n + \hat{e}_n^h}{\bar{e}_s + \hat{e}_s^h} = \frac{\bar{e}_n + \frac{\gamma_n^h}{\gamma_s^h}\hat{e}_s^h}{\bar{e}_s + \hat{e}_s^h}
\end{aligned}
$$

Proposition 4.1 *Under Assumptions 4, 5 and 6,* Δw^a *is increasing (decreasing) in* $\hat{e}_s^m$ *as* $\frac{\gamma_n^m}{\gamma_s^m} \geq (<) \frac{\underline{e}_n}{\underline{e}_s}$*;* Δw^h *is increasing (decreasing) in* $\hat{e}_s^h$ *as* $\frac{\gamma_n^h}{\gamma_s^h} \geq (<) \frac{\bar{e}_n}{\bar{e}_s}$*; and* Δw^m *is independent of* $\hat{e}_s^m$ *and* $\hat{e}_s^h$*.*

The proof follows from the facts that

$$
\frac{\partial \Delta w^a}{\partial \hat{e}_s^m} \underset{<}{\geq} 0 \text{ as } \frac{\gamma_n^m}{\gamma_s^m} \underset{<}{\geq} \frac{\underline{e}_n}{\underline{e}_s}
$$

$$\frac{\partial \Delta w^m}{\partial \hat{e}_s^m} = \frac{\partial \Delta w^m}{\partial \hat{e}_s^h} = 0 \text{ as } \frac{\gamma_n^h}{\gamma_s^h} = \frac{\gamma_n^m}{\gamma_s^m}$$

$$\frac{\partial \Delta w^h}{\partial \hat{e}_s^m} \gtreqless 0 \text{ as } \frac{\gamma_n^h}{\gamma_s^h} \gtrless \frac{\bar{e}_n}{\bar{e}_s}$$

The Proposition illustrates that the sectoral wage gaps in the a and h sectors depend on the trade-offs between the relative fixed costs of skill acquisition and the relevant relative ability mark-ups. Clearly, if $\frac{\gamma_n^h}{\gamma_s^h} > \frac{\bar{e}_n}{\bar{e}_s}$ and $\frac{\gamma_n^m}{\gamma_s^m} > \frac{\underline{e}_n}{\underline{e}_s}$ then the wage gap in both sectors a and h will co-move positively with the ability threshold $\hat{e}_s^m$ and $\hat{e}_s^h$. Thus, consider aggregate shocks that reduce the ability thresholds $\hat{e}_j^m$ and $\hat{e}_j^h$ and thereby induce more individuals to get skilled. This will reduce the wage gaps in the a and h sectors only if the relative fixed cost of getting skilled for the s caste is sufficiently low so as to overcome their initial skill gaps at both the low and top ends of the ability distributions. In the $m-$sector, changes in the ability thresholds have no effect on the wage gap because the cost mark-up of training to move up from sector m to sector h is the same for both castes. Hence, the sectoral wages of both castes react symmetrically.

5 A two-sector illustration

The analysis above has described how changes in the ability thresholds of the two groups affect the sectoral wage gaps between the groups. The key aspect that we have been silent on thus far is the effect of exogenous productivity shocks on the ability thresholds. This is key since a goal of the model is to assess the role of aggregate productivity shocks in accounting for the process of wage and education convergence between the castes that we have seen over the past three decades.

In order to gain some analytical insights on this issue, we now specialize the three-sector model developed above to the two-sector case. In particular, we assume that there are only two sectors – a and m, i.e., we eliminate the h sector. We should note that in this two-sector example there is only one cut-off ability threshold for each group. Hence, in the following we drop the sectoral superscripts from the notation for the thresholds and use $\hat{e}_j$ to denote the threshold ability of caste $j = s, n$ so that all individuals with ability levels greater than $\hat{e}_j$ will get skilled in order to work in the $m-$sector.

Under these functional form assumptions, the equilibrium of the economy is determined by the system of equations:

$$p_m = \frac{\left(\frac{1-\theta}{\theta}\right)\left[y^A - \bar{c}L\right]}{y^M - F}$$

$$\begin{aligned} p_m &= \frac{A\hat{e}_s}{M\hat{e}_s - \phi\left(\gamma_s - a\hat{e}_s\right)} \\ \frac{\hat{e}_n}{\hat{e}_s} &= \frac{\gamma_n}{\gamma_s} \end{aligned}$$

where y^A, y^M and F are given by equations (4.12), (4.13) and (4.15) above (without the h-sector terms). The first equation above is the market clearing condition for the agricultural good while the second equation gives the ability threshold for group s members to get skilled in order to work in sector m. The third equation comes from the relation $z_s\left(\hat{e}_s\right) = z_n\left(\hat{e}_n\right)$. Under the assumed linear skill cost technology, this condition implies that the relative ability thresholds of the two castes are proportional to the relative fixed costs of getting skilled.

Combining these three relations gives the key equilibrium condition

$$\frac{A\hat{e}_s}{M\hat{e}_s - \phi\left(\gamma_s - a\hat{e}_s\right)} = \frac{\left(\frac{1-\theta}{\theta}\right)\left[y^A - \bar{c}L\right]}{y^M - F} \tag{5.16}$$

which involves only one unknown $\hat{e}_n$ since $\hat{e}_s = \frac{\gamma_s}{\gamma_s}\hat{e}_n$. Note that in this case we have

$$\begin{aligned} y^A - \bar{c}L &= \frac{A}{2}\left[S\frac{\left(\frac{\gamma_s}{\gamma_s}\hat{e}_n\right)^2 - \underline{e}_s^2}{\bar{e}_s - \underline{e}_s} + N\frac{\hat{e}_n^2 - \underline{e}_n^2}{\bar{e}_n - \underline{e}_n} - \frac{2\bar{c}L}{A}\right] \\ y^M &= \frac{M}{2}\left[S\frac{\bar{e}_s^2 - \left(\frac{\gamma_s}{\gamma_s}\hat{e}_n\right)^2}{\bar{e}_s - \underline{e}_s} + N\frac{\bar{e}_n^2 - \hat{e}_n^2}{\bar{e}_n - \underline{e}_n}\right] \\ F &= \phi S\left(\frac{\bar{e}_s - \hat{e}_s}{\bar{e}_s - \underline{e}_s}\right)\left[\gamma_s - \frac{a}{2}\left(\bar{e}_s + \frac{\gamma_s}{\gamma_s}\hat{e}_n\right)\right] + \phi N\left(\frac{\bar{e}_n - \hat{e}_n}{\bar{e}_n - \underline{e}_n}\right)\left[\gamma_n - \frac{a}{2}\left(\bar{e}_n + \hat{e}_n\right)\right] \end{aligned}$$

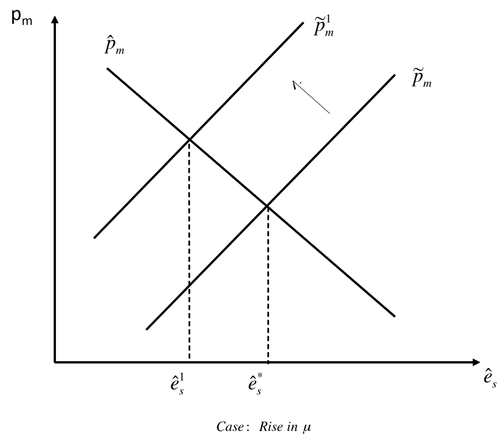
A key question of interest to us is whether the aggregate changes in the Indian economy over the past two decades can explain, at least partly and qualitatively, the wage and education convergence across the castes. Toward that end, we define

$$A = \mu\bar{A} \tag{5.17}$$

$$M = \mu\bar{M} \tag{5.18}$$

This specification nests sectoral and aggregate productivity changes with changes in μ being aggregate shocks while changes in $\bar{A}$ and $\bar{M}$ are sector-specific productivity shocks. Moreover, to avoid

Figure 9: An aggregate productivity shock



scale effects we also set

$$\phi = \frac{\mu}{\bar{\phi}} \tag{5.19}$$

If the skill acquisition cost were not indexed to the aggregate productivity parameter μ, the cost of acquiring skills would become progressively smaller as a share of total output of the economy simply in response to aggregate productivity growth. The specification avoids this scale effect of growth.

In order to determine the effect of aggregate productivity growth on the economy, it is useful to define the following:

$$\hat{p}_m = \frac{A\hat{e}_s}{M\hat{e}_s - \phi\left(\gamma_s - a\hat{e}_s\right)} \tag{5.20}$$

$$\tilde{p}_m = \frac{\left(\frac{1-\theta}{\theta}\right)\left[y^A - \bar{c}L\right]}{y^M - F} \tag{5.21}$$

Figure 9 plots the two equations. Clearly, $\hat{p}_m$ is a downward sloping function of $\hat{e}_s$ while $\tilde{p}_m$ is an upward sloping function. The equilibrium threshold ability is $\hat{e}_s^*$. All ability types above this critical level get skilled and work in the $m-$sector while the rest work in the $a-$sector. Recall that $\hat{e}_n = \frac{\gamma_n}{\gamma_s}\hat{e}_s$, so this solves for the ability threshold for type-n individuals as well.

Our primary interest is in determining the effect of an aggregate productivity shock on this economy. The following Proposition summarize the effects of a TFP shock on this economy:

Proposition 5.2 *An increase in aggregate labor productivity μ decreases the ability threshold $\hat{e}_s$. This (i) reduces the caste wage gap in sector a if and only $\frac{\gamma_n}{\gamma_s} > \frac{\underline{e}_n}{\underline{e}_s}$; and (ii) reduces the caste wage gap in sector m if and only if $\frac{\gamma_n}{\gamma_s} > \frac{\bar{e}_n}{\bar{e}_s}$.*

The logic behind Proposition 5.2 is easiest to describe using Figure 9 which shows the effect of an aggregate TFP shock on the equilibrium system of equations 5.20-5.21. A rise μ leaves the $\hat{p}_m$ locus unchanged but shifts the $\tilde{p}_m$ locus up and to the left. As a result the equilibrium threshold ability level $\hat{e}_s$ declines. Intuitively, the aggregate TFP shock leaves unchanged the relative gains and losses from getting skilled since the productivity of all sectors (including the education sector) are affected symmetrically. On the hand, a higher μ raises the aggregate supply of the agricultural good *net* of the subsistence amount $\bar{c}L$ while leaving the aggregate supply of the manufacturing good net of the training cost unchanged. The resultant excess supply of the agricultural good induces a terms of trade worsening of the agricultural good. As agents increase their demand for good m its relative price p_m rises. All else equal, this increases the attractiveness of working in the $m-$sector. Consequently, the threshold ability falls and agents with lower ability now begin to get trained. Part (b) os the Proposition follows directly from Proposition 4.1.

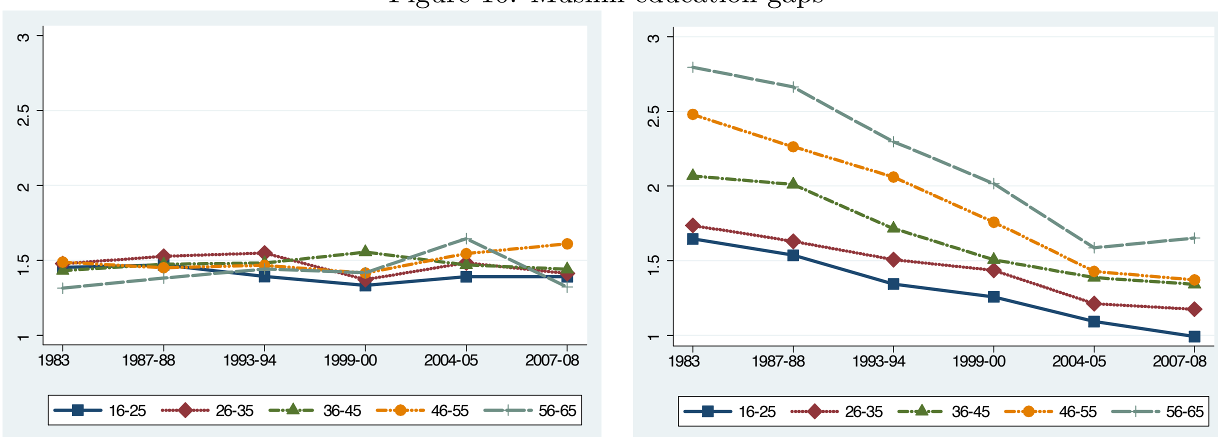
Before closing this section it is worth also examining the effect of productivity changes that are biased against the agricultural sector, a feature that characterizes India during this period. Specifically, suppose A remains unchanged while M rises. The direct effect of this shock would be to shift the $\hat{p}_m$ schedule down and to the left while simultaneously shifting the $\tilde{p}_m$ schedule down and to the right. The effect of these changes would be an unambiguous decline in the relative price of the manufacturing good, i.e., p_m falls. The effect on $\hat{e}_s$ is ambiguous and depends on the net strength of the two shifts. The key feature to note though is that the structure is perfectly consistent with both an improvement in the agricultural terms of trade as well as an increase in the share of each group getting skilled.

6 Some Independent Evidence: Muslims

The model that we outlined above predicts that a key necessary feature for there to be convergence between the groups is the presence of a pre-existing subsidy to acquiring skills for SC/STs. The direct connection of such subsidies to actual policies in India are the reservations in jobs and education that were provided to SC/STs in the Indian constitution when it came into effect in 1951. Alongside SC/STs, another group that was worse off relative to the mainstream in terms of education and income were the Muslims. However, the Constitution of India does not provide any reservations for Muslims. If the margin identified by the model is correct then the aggregate productivity rise in India since the mid 1980s should have led to a strong catch-up of the education and wage levels of SC/STs to the Muslim levels..

Figure 10 shows the education attainment levels by age groups of non-SC/STs relative to Muslims (Panel (a)), and Muslims relative to SC/STs (Panel (b)). Panel (a) of the The Figure clearly shows that Muslims had lesser education relative to non-SC/STs to start with. Moreover, over the sample period, their education gaps relative to non-SC/STs marginally worsened while the gaps with SC/STs declined sharply.

Figure 10: Muslim education gaps



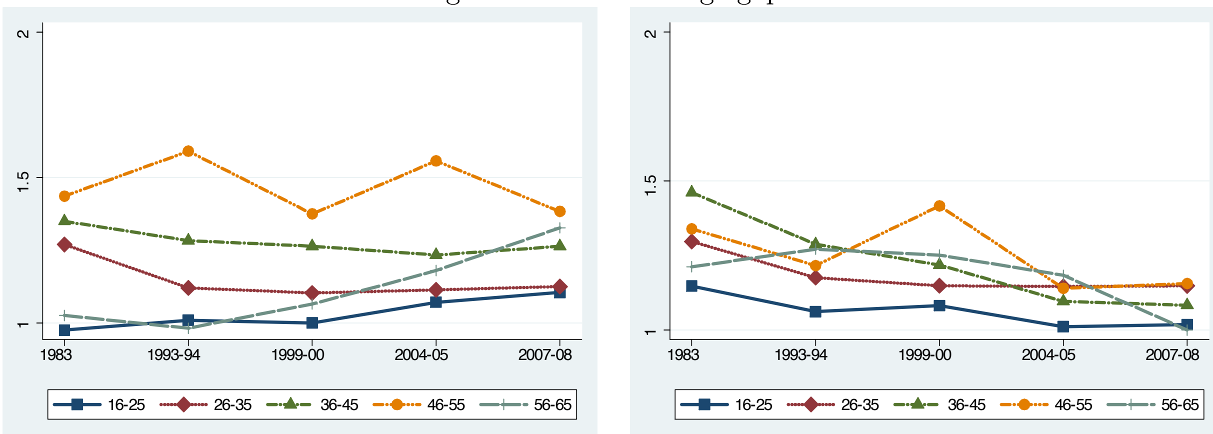
(a) Non-SCSTs / Muslims (b) Muslims / SCSTs

Notes: Panel (a) of this Figure presents the gap in years of education between non-SC/STs and Muslims for different age groups. Panel (b) presents the gap in years of education between Muslims and SC/STs by age-groups.

Figure 11 shows the wage gaps between Muslims and the other two groups. The Muslim wage dynamics are similar to the education dynamics. They were worse off relative to non-SC/STs at the beginning of the sample period and better off relative to SC/STs. Over the period 1983-2008, their wage gaps with non-SC/STs marginally widened while simultaneously declining relative to SC/STs.

We consider these patterns to be independent evidence in support of the role of a pre-existing affirmative action policy in accounting for the wage and education convergence of SC/STs that was suggested by our model.

Figure 11: Muslim wage gaps



(a) Non-SCSTs / Muslims (b) Muslims / SCSTs

Notes: Panel (a) of this Figure presents the wage gap by age-group between non-SC/STs and Muslims. Panel (b) presents the wage gap by age-group between Muslims and SC/STs.

7 Conclusion

The past three decades have seen a significant convergence in the education attainments, occupation choices and wages of scheduled castes and tribes (SC/STs) in India to toward the corresponding levels of non-SC/STs. In this paper we have shown that underlying this overall convergent patterns, there has been significant sectoral heterogeneity. In particular, we have found that while the wage gaps between SC/STs and non-SC/STs in the service sector declined significantly, the wage gap in agriculture actually widened. Manufacturing, on the other hand, displayed no change in the wage gaps. At the same time, we also found evidence that there were common trends in the sectoral productivity patterns suggesting the presence of important aggregate developments that were common to all sectors even though these dynamics were occuring in the backdrop of a significant structural transformation of the economy which was gradually shifting out of agriculture into non-agricultural sectors, particularly services.

Motivated by these patterns, we developed a model that is able to generate the overall convergence in wages as well as the structural transformation of the economy away from agriculture. Importantly, we show that aggregate productivity shocks in the model induce the sectoral heterogeneity in movements in the wage gaps, making it consistent with the data. The central feature of the model is education and training of the labor force and the incentives for workers to invest in this training. Our key finding from the model is that a multi-sector, heterogenous agent economy, when subjected to aggregate productivity shocks, can induce convergence across the two groups if there are some pre-existing institutions in place to lower the costs of acquiring education/skills for the SC/STs relative to non-SC/STs. Given that reservations in education for SC/STs were incorporated into the Indian constitution in 1950 in order to offset their historical disadvantage, this condition appears to have some support in the facts. We provide some indirect evidence on other minorities (Muslims) that were not covered by these affirmative action programs that is supportive of the channel formalized here.

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.

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

Bowles, S., G. C. Loury, and R. Sethi (2009): "Group Inequality," Economics Working Papers 0088, Institute for Advanced Study, School of Social Science.

Firpo, S., N. M. Fortin, and T. Lemieux (2009): "Unconditional Quantile Regressions," *Econometrica*, 77(3), 953–973.

Hnatkovska, V., A. Lahiri, and S. Paul (2012a): "Breaking the Caste Barrier: Intergenerational Mobility in India," *Journal of Human Resources (forthcoming)*.

Hnatkovska, V., A. Lahiri, and S. Paul (2012b): "Castes and Labor Mobility," *American Economic Journal: Applied Economics*, 4(2), 274–307.

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.

Kijima, Y. (2006): "Caste and Tribe Inequality: Evidence from India, 1983-1999," *Economic Development and Cultural Change*, 54(2), 369–404.

Loury, G. (1981): "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49(4), 843–867.

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.

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.

8 Data Appendix

To obtain our sectoral groupings we use the 4-digit industry code that individuals report in the NSS data and convert it into a one-digit code. This gives us seventeen categories. We then group these seventeen categories into the three broader industry categories: Agriculture, Manufacturing, and Services. Table 3 summarizes one-digit industry codes in our dataset and their grouping into the three broad industry categories.

Table 3: Industry categories

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

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about our work on our website
www.theigc.org

For media or communications enquiries, please contact
mail@theigc.org

Subscribe to our newsletter and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International Growth Centre

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid from the British people

Designed by soapbox.co.uk