The Impact of Positive Discrimination in Education in India: Evidence from a Natural Experiment[*]

Guilhem Cassan[†]

November 3, 2011

Abstract: *I use a natural experiment in order to assess the impact of positive discrimination in India on targeted groups' educational attainment. I take advantage of the harmonization of the Schedule Castes and Tribes lists within the Indian states taking place in 1976 to measure the increase of the educational attainment of the groups that suddenly became members of the "Scheduled Castes". This setting allows me to measure the impact of the policy on the castes that were added to the list from 1976 on, taking the castes already on the list as a control group. Using two different identification strategies, I show that this policy had no effect overall, while areas with better than average school supply and urban areas benefited from the policy.*

JEL Classification:
Keywords: scheduled caste; quota; positive discrimination.

Introduction

The question of the impact of positive discrimination policies has attracted a lot of attention in the past years. The literature mainly focuses on the US (Holzer and Neumark, 2006), but has also been interested in other contexts. As a matter of fact, a large literature already exists on the "reservations" for lower castes created at the Independence of India. This policy has three main dimensions: provision of reserved seats in the Assemblies (at the state and federal levels), quotas in public employment, and various advantages in the educational system (quotas in higher education as well as free secondary schooling). While the first two dimensions have been at the center of attention, positive discrimination in education has been less rigorously studied. Indeed, Pande (2003) studied the impact of reservations of seats in legislative assemblies for low castes and Chattopadhyay and Duflo (2004) as well as Beaman et al. (2009) studied the impact of reservations of seats for women in local assemblies for various outcomes. Howard and Prakash (2008) and Prakash (2009) chose to focus on the effect of reservation of public employment. On education however, most of the research has focused on the equity aspect of positive discrimination, with the case study of Bertrand et al. (2010) providing a thorough analysis of the question. However, the more simple question of evaluating the policy with respect to its goals has received only little attention (Chalam (1990), Chitnis (1972)). This is somehow surprising, as positive discrimination as a whole, and in education in particular is a heavily debated issue in India (Kumar, 1992), and the amounts spent to promote education for the Scheduled Castes and Scheduled Tribes is the largest budget of the compensatory discrimination policies (Galanter, 1984). Hence, I argue that a central question that remains to be answered is whether or not the positive discrimination policy managed to increase the educational attainment of its targeted population. This article ambitions to fill this gap, providing the first piece of systematic evidence on the causal impact of the positive discrimination policies on the educational attainment of the Scheduled Castes. Indeed, there is no clear evidence of the causal effect of the policy, as underlined by qualitative accounts: *"[...] it remains unclear just how much of the increase of the education for the backward classes represents 'special treatment' over and above what is provided for the whole population [...]"* (Galanter, 1984). An obvious reason for the lack of evaluation is that evaluating such a policy is a challenge by construction, as the population targeted is strongly selected. In order to avoid this problem, this paper takes advantage of a nationwide natural experiment on the beneficiaries of caste-based positive discrimination in order to measure its impact on educational attainment. Indeed, while the borders of the Indian states were redefined in

1956, the state wise lists of castes subject to positive discrimination remained defined as in the previous borders during 20 years. This created a situation where, depending on their location within a state, members of the same caste could be defined (or not) as having access to the positive discrimination system. Hence, the timing of access to the positive discrimination is exogenous, allowing the caste members that had access to the positive discrimination scheme from 1950 on to be a credible counterfactual for the members of the same castes that had access to the policy only from 1976 onwards. The main result of this paper is that such policy does not seem to have had an impact on the educational level of the Scheduled Caste. However, the effect of the policy are heterogenous, and the urban areas as well as the areas with a higher than average level of secondary education provision exhibit a significant effect on educational attainment. This points to an interpretation of the overall absence of effect of the policy as being due to the lack of provision of schooling facilities. The first part will detail the specific historical context of this natural experiment, the second one will detail the data and the empirical strategy, while the third part will present the results obtained.

1 Context

1.1 Positive discrimination in India

Positive discrimination (or "reservations") has a long history in India, dating from the colonial period. Indeed, the Government of India Act of 1919 (or "Montagu-Chelmsford Reforms") provided special representation to Non Brahmin (which include untouchables along many other "middle" or "low" castes), which lead to the implementation of quotas in education and public jobs in the Provinces of Madras and Bombay. However, it is not before the Independence and the Constitution of 1950 that a systematic positive discrimination policy was implemented across India. Reservations concerned 3 main items: legislative seats, education and public employment. This paper deals with the impact of the reservations in education, which consists in two main policies. The first one is free secondary education for members of the targeted groups (along with various schemes of scholarships), and the second one is seats reserved in higher education institutions. Positive discrimination might thus affect schooling through various channels. First of all, by reducing the cost of secondary education, it favors longer studies in the cost-benefit arbitrage of the household. Also, the quotas in higher education will allow, among those that had made the choice to pursue their studies up until this level, to effectively have access to it. Finally, the quotas in public employment also are an incitation to pursue

longer studies, as they decrease the probability of unemployment, and thus, increase the returns to education.

1.2 The definition of the Scheduled Castes

The Constitution defines 3 groups susceptible to be subject to positive discrimination: the "Scheduled Castes" (SC), the "Scheduled Tribes" (ST) and the "Other Backward Classes" (OBC). However, there is no precise definition of the criteria making a caste or a tribe eligible to the status, thus leaving the door open to some arbitrariness in the definition of the list. Indeed, the constitution of the list of caste subject to positive discrimination has been and still is the subject of debates, and has been subject to changes over time. This section, drawing from the work of Galanter (1984) provides a short history of the list of Scheduled Castes. The first all India classification of untouchables dates from 1936, even if earlier attempts had been made, with first estimates of the "Depressed classes" population, as they were then called, made in 1917. One of the main problems with the making of such classification is that the definition of untouchability has proven tricky: as untouchability varies in its meaning across the sub continent, it is hard to create a definition that would apply to the whole country. Indeed, if untouchable castes are relatively well identified in the South and West of India, it is not the case in the other parts of the country. Hence, the Constitution of 1950, which establishes reservations in legislative seats, higher education and public employment for Scheduled Castes and Tribes carefully avoids to define clearly the concept and only provides a procedure of designation: "*castes, races or tribes or parts of or groups within castes, races and tribes which shall for purposes of this Constitution be deemed to be Scheduled Castes in relation to that State*". It must also be noted that if caste is the main criterion for the inclusion in the SC list, other characteristics are also explicitly taken into account, and most notably religion [1] and territory. The list, specific to each state, is proposed by the President of India and can only be changed by an Act of Parliament. The list drawn in 1950 was essentially a reconduction of the classification of 1936. However, this classification has been criticized for its inconsistencies across States[2]. But the revision of the list is not a light process, and has thus not happened frequently. Indeed, it was only revised three times since the Independence, but with only one revision being of real

---

[1]In the 1936 list, only Hindus could claim the Scheduled Castes status, in 1950, 4 Sikh untouchable castes were added, and all of them in 1956, while the Buddhists could claim the SC status from 1990 (except in Maharashtra in which their SC status has been accepted since 1956).

[2]Bayly (1999) gives the example of the Khatik caste, considered as SC in Punjab, but classed as a "forward" caste in the neighboring state of Uttar Pradesh.

importance[3]. With around 2.4 million new SC over a population of roughly 80 million SC (Galanter, 1984), the Scheduled Castes and Scheduled Tribes (Amendment) Act of 1976 was the most dramatic change in the list of SC in India.

1.3 The Scheduled Castes and Scheduled Tribes (Amendment) Act of 1976

The history of this Act traces back to 1956, year which was to witness a vast reorganization of the borders of the Indian states along language lines. But, as the borders of the states were redefined, the state wise SC lists remained unaffected. This lead to a situation of large discrepancies between State borders and the list of SC, the latter one not being defined at the state level anymore, but by regions within each state, regions broadly corresponding to the pre 1956 borders (see Figures 1 and 2). It can be seen that the south of India has been the most affected by this change, the main change being the disappearance of the states of Bombay and Hyderabad.

[Figure 1 about here.]

The reason for the list not to be adjusted to the new borders is that the administration did not have enough time to prepare the new lists: *"It has been mentioned in the last report that the President has issued the SC and ST Lists (Modification) Order, 1956, specifying the SC and ST in the re-organized States. As these lists had to be issued urgently for the re-organized States, it was not possible to prepare comprehensive and consolidated lists and therefore, the SC and ST had to be specified in these list territory-wise within each re-organized State"* (Government of India, ed, 1958). But not only did the administration fail to change the lists on time, but it failed to do so for a period of twenty years. The yearly reports of the Commissioner on SC and ST are particularly telling in this aspect, as every year from 1957 until 1964 is the occasion to refer to the fact that *"[...]the question of preparation of comprehensive lists of SC and ST for the reorganized States [...] remained pending [...]"* (Government of India, ed, 1960). In 1965, a committee was appointed in order to revise the list and in 1967, a bill was proposed taking into consideration the recommendations of the Committee. However, the bill also proposed to exclude from Scheduled Tribes the converts to Christianity and Islam, which lead to such long debates that the Lok Sabha (Federal Legislative Assembly) was dissolved in 1970 before the law could be passed. Finally, it is only under the

---

[3]The change of 1956 mainly affected Rajasthan and Uttar Pradesh, and also allowed all Sikh untouchable castes to claim SC status, while the change of 1990 allowed the Buddhists to have access to the SC status in all the states.

emergency rule of Indhira Gandhi and the SC and ST (Amendment) Act of 1976 that the Scheduled Castes lists were harmonized within states. Hence, due to administrative reasons, a situation was created in which an individual from a caste could (or not) be considered as a member of a SC depending on her location in a state. From 1950 to 1956, this was due to the fact that this caste was split between two different states and arbitrarily put in a different category in the different states, but for the majority of the period, from 1956 to 1976, it was due to the incapacity of the administration to harmonize the Scheduled Castes lists within each newly created state. Hence, because the pre 1956 borders were not drawn according to the linguistic and cultural areas of India, certain castes were split in two across two states, facing different policies in terms of their SC status. With the reorganization of the state borders, they were facing different SC status, but in an identical political context. This situation thus creates a natural experiment setting in which an identical caste faced a different treatment with respect to positive discrimination due to an historical accident. According to Galanter (1984), the Area Restriction Removal Act of 1976 consisted in an almost complete disappearance of intra state area specifications, which dropped from 1,126 to 64. The removal of restrictions led to an increase of 2.4 million of the number of SC (3% of the 1971 SC total population), which was to be unevenly distributed across the Indian states. The states facing the largest increase being Maharashtra (+5%), Madhya Pradesh (+5.5%), Karnataka (+11%) and Kerala (+ 13%) (Government of India, ed, 1978), precisely the states that were the most affected by the 1956 change in borders. Due to the very small number of new members of SC in the other states of India, the analysis will focus on those four states. Figure 2 pictures their creation in 1956, the pre-1956 borders, as well as the SC area restrictions for the 1956-1976 period.

[Figure 2 about here.]

2 Data and Empirical Strategy

2.1 Data

The Indian Demographic and Health Survey of 1998-99 provides information on both the caste name and the district of residence. With a sample size of more than 90,000 households, this survey provides information on the education level of more than 500,000 individuals. Using the 1971 and 1981 Census lists of Scheduled Castes and Scheduled Tribes defined at the area of restriction level, I am able to identify the households that were added to the Scheduled Castes lists after 1976 (under the assumption that the

household did not migrate, as there is no information on the migrant status in the survey). However, the coding of each caste name in the two scheduled castes categories is not something obvious. Indeed, each caste can have various synonyms, varying locally. The SC list provides some synonyms, while the project People of India (Singh, ed, 1996) provided a state specific list of synonyms allowing me to code each caste into its SC group[4]. Overall, 9.9% of the population declaring a SC status is coded as having been added to the SC list in 1976[5]. This data is complemented by the District Information System for Education dataset[6] (round 2007). This Census of schools provides the year of construction of each school, allowing me to build a measure of the school supply in 1977, at the district level[7]. The information being disaggregated at a very fine level, I am able to distinguish, within a district, the number of primary and secondary schools, as well as their location (urban or rural areas). Combined with the Indian District Data (Vanneman and Barnes, 2000) which contains district level information on population taken from the Census, I can recompute the school supply in 1977 per children at school age in each urban and rural area of each district.

2.2 Data limitations

As the access to the positive discrimination benefits depends on the caste name, some caste identity manipulation might be at play[8], with the castes whose name was not on the list in 1950 trying to pass as members of a caste listed on the schedule. Moreover, it is likely that the households not willing to manipulate their caste identity but searching the protection of the SC status had a tendency to migrate to the areas in which their caste was listed as a SC. However, this selection into pre 1976 treatment would only bias the estimates towards zero, as the persons manipulating their caste identity or migrating are probably the ones that would have benefited the most from access to positive discrimination in 1976. An other serious concern in the tendency for certain members of SC to refuse to answer their precise caste name, but to prefer a generic name such as "Scheduled Caste" or "Harijan". This is of course an issue, as it is to be expected that this tendency is not random, leading me to run the regressions on a

---

[4]See Appendix A for the caste synonyms not listed in the SC list used to identify caste names.

[5]In Karnataka, 17.1% are coded as new SC, 11% in Kerala, 6.2% in Madhya Pradesh and 8.4% in Maharashtra.

[6]I thank Adrien Lorenceau for having provided me with the database.

[7]As the 2007 dataset only contains information on the schools still opened in 2007, the use of this dataset relies on the -weak- assumption that only a negligible number of the schools closed between 1977 and 2007, and if they did, that their closure was not correlated with the effectiveness of the policy.

[8]Cassan (2011) documents that caste identity manipulation was at play in colonial Punjab in response to a positive discrimination-type of policy.

selected subsample of the SC population. However, while this is clearly an issue for the external validity of the results, this selection into the declaration of the "real" caste name is unlikely to bias the estimates. Indeed, a bias in the results could take place only in the case where self declaration into "Harijan" would be different between the SC listed in 1950 and the SC listed in 1976 combined with a change in this difference taking place over time. There is no way to test if the tendency to self declare as an "Harijan" or a "Scheduled Castes" is different across castes. However, it is possible to test if this tendency has evolved over time. It can be seen in Appendix B that the tendency to declare oneself as a "Harijan" is not correlated with a trend, thus pointing to the fact that a change in the selection does not seem to be happening. Hence, the respondents declaring themselves as member of a SC, but either having a caste name not granting them access to the SC status in the state in which they are residing, or calling their caste "Harijan" or "Scheduled Caste" are excluded from the analysis.

2.3 Descriptive statistics

The difference in the timing of access to the SC status suggests that if the policy had an impact, the evolution of the educational status of the two groups should diverge during the 1950-1976 period, when their treatment status differs and converge after 1976, when both groups are treated. The precise identification of the individuals effectively affected by the change of status is not straightforward, as it is only the individuals that were young enough at the year of the change of status that could benefit from the positive discrimination in schooling. Table 1 presents the treatment status of the various cohorts.

[Table 1 about here.]

Figure 3 graphs the evolution of the number of years of education by cohorts of the two caste categories. Only the individuals aged 21 and above at the age of the survey are represented (and are used in the parts of the paper dealing with the number of years of education), as the younger cohorts might still be at school. As some individuals could still be studying above 21 years old, the number of years of schooling is caped at 15. The cohorts born before the 1940's seem to face a very similar trend, while a divergence takes place from the cohorts born after 1940, with no catching up happening in the 70's. This picture is broadly consistent with the access to the SC status being determinant for the cohorts young enough to benefit from it in 1950, while the reform of 1976 on the contrary did not seem to have an impact on the caste newly added to the scheme, as the divergence between the two caste group seems to continue even for the cohorts born

8

in the 1970's. According to the DISE 2007 database and the Census of India, in 1977, the number of primary schools was of 2.4 per thousand children[9] aged below 15 while the number of secondary schools was of 0.13 per thousand children aged below 20. The supply of schooling was thus very poor at the time, which might explain the apparent poor success of the policy.

[Figure 3 about here.]

As capping the number of years of schooling to 15 might be a concern if it was affecting a large share of the observations, Figure 4 presents an histogram of the distribution of the number of years of schooling of SC in the cohorts born between 1944 and 1979. It can be seen that less than 1.5% of this population has a number of years of education higher than 15[10].

[Figure 4 about here.]

However, this general evolution hides the fact that the trends are very different between rural and urban areas. Figure 5 pictures the evolution of the years of schooling of the two groups in urban and rural areas. The difference in the two pictures is striking. Indeed, in the urban setting, one can clearly see a difference between the groups for the cohorts born in the 1940's and 1950's, followed by a convergence for the cohorts born in the 60's. The picture in the rural areas is quite different, with the trends of the two groups remaining very similar until the cohorts born at the end of the 60's, for which, quite surprisingly, it seems that the divergence between the castes having access to the SC status in 1950 and those having access to it in 1976 begins during the 1970's. In sum, the divergence observed at the aggregated level seems to be coming from the combination of two different processes taking place in urban and rural areas: while the reservation program seems to have had an effect in the urban areas, such an evolution does not seem to be happening in rural areas. Indeed, in the rural areas, it is as if the effect of the program had been delayed by 30 years, with the caste groups added to the list in 1950 starting to diverge from the other only from the cohorts born in the 1970's, when a convergence was on the contrary to be expected.

[Figure 5 about here.]

---

[9]The 1977 population of school aged children is calculated as a linear interpolation between the 1971 and 1981 Census.

[10]The very high number of zeros in the data is also a concern that will be treated carefully in the econometric analysis.

However, as the individuals are surveyed in 1998-1999 only, the use of the older cohorts is problematic, as their survival rate might be correlated to their education level. Hence, the remainder of the paper will focus on the cohorts born from 1944 (i.e. aged 55 at most at the time of the survey, and aged 6 and under at the time of the first implementation of the positive discrimination policy, allowing them to fully benefit from it). Table 2 reports the main descriptive statistics for the age groups that were either too old (18 years old and above) to be treated, or young enough (6 years old and under) to benefit entirely from the change in the reservation status. The descriptive statistics are consistent with the polynomial smoothing, as the level of education seem to converge between the two caste groups in the urban areas as opposed to rural areas. Moreover, it is to be noted that the castes added to the Schedules in 1950 seem to be residing more in urban areas. This is in line with the case already discussed of the migration of individuals whose caste is not scheduled in their (potentially rural) area of residence to an urban area in which their caste is listed.

2.4 Identification Strategy

As the SC might have been the target of specific policies other than the Area Restriction Removal Act, around 1976, the evaluation of the access to the SC status is not straightforward, and can not rely on a comparison with non SC.

In order to account for this possibility I will use the SC already in the list in 1976 as a control group, allowing to distinguish the effect of a general improvement in access to education for those populations with the effect of having access to reservations. However, it is not obvious to determine who exactly was affected by the change in the SC status: while the cohorts too young to be at school by the time of the implementation of the policy were fully exposed (6 years or younger), those aged between 6 and 18 (i.e. supposedly at school or already having finished their education) would be less affected by the policy. Hence, the treatment variable will not be a simple dummy, but a variable taking value 0 for individuals aged 18 and above in 1977, 1 for those aged 6 and under, and linearly increasing from 0 to 1 for the cohorts aged 6 to 18 at the time of implementation[11]. The setting of the natural experiment suggests two complementary identification strategies. Indeed, within each state, the castes that are spread across areas of restrictions have members whose SC status might differ. Hence, the first identification strategy would be to compare members of the same caste, in a same state, facing a different SC status due to their location within the state, suggesting a caste and state fixed effect specification.

---

[11]The results of regressions taking into account only the fully treated cohorts are similar and are reported in Appendix D.

Table 1: Descriptive statistics by Scheduled Caste status.

	Urban and rural areas		Urban areas		Rural areas	
	1950 Scheduled Castes	1976 Scheduled Castes	1950 Scheduled Castes	1976 Scheduled Castes	1950 Scheduled Castes	1976 Scheduled Castes
1943< Birth Year <1979						
Years of education	5.2 (0.27)	3.8 (0.33)	7.0 (0.45)	5.6 (0.68)	4.0 (0.21)	3.1 (0.32)
Literacy rate	60% (0.02)	49% (0.04)	74% (0.03)	67% (0.06)	50% (0.02)	41% (0.04)
Urban resident	41% (0.05)	29% (0.06)				
N	4,053	564	1,751	183	2,302	381
1944< Birth Year <1960						
Years of education	3.9 (0.32)	2.61 (0.43)	5.8 (0.6)	3.2 (0.8)	2.6 (0.18)	2.4 (0.53)
Literacy rate	49% (0.02)	40% (0.05)	66% (0.04)	47% (0.09)	38% (0.02)	38% (0.06)
Urban resident	39% (0.05)	30% (0.07)				
N	1,239	156	504	52	735	104
1970< Birth Year <1979						
Years of education	6.8 (0.29)	5.0 (0.50)	8.6 (0.38)	7.7 (0.91)	5.5 (0.35)	3.8 (0.42)
Literacy rate	72% (0.02)	58% (0.05)	86% (0.02)	81% (0.06)	61% (0.03)	48% (0.05)
Urban resident	42% (0.05)	28% (0.07)				
N	1,355	199	603	67	752	132

Survey weights used. Standard errors in parenthesis.

However, while the fact that members of a same caste provide the counterfactual for one another is quite a satisfying feature of this identification strategy, it also relies on comparing potentially different areas within a state. This might be an issue if an area of restriction specific shock was to happen at the time of the implementation of the policies. Hence, a second strategy would consist of comparing within the same area of restriction castes which SC status has or not varied[12], suggesting an area of restriction fixed effect specification. This strategy, which avoids the comparison of different areas, however compares different caste groups, which might face different shocks. Overall, those two identification strategies, even if not devoid of bias, are exposed to very different type of bias (one might be biased by area specific shocks, while the other would be affected by caste specific shocks), and are thus complementary.

Hence, I will estimate regressions of the form:

$$Edu_{id} = constant + \beta sc76_{id} + \delta sc76_{id} * t76_i + \gamma t76_i$$
$$+ \lambda X_{id} + \epsilon_{id}$$

(1)

Where Edu_{id} is a measure of educational attainment of individual i residing in area of restriction d , $sc76_{id}$ a dummy indicating whether individual i residing in district d is member of a caste added to the SC list in 1976, $t76_i$ taking value 0 for individuals aged 18 and above in 1977, 1 for those aged 6 and under, and linearly increasing from 0 to 1 for the cohorts aged 6 to 18 at the time of implementation (year of the implementation of the law)and X_{id} a set of control variables[13]. Each specification then add either caste and state dummies or area of restriction dummies.

3 Results

3.1 Number of years of schooling

The first outcome to be looked at would be the number of years of schooling. As the data is by construction non negative and, as shown in Figure 4, contains a large share of zeros, OLS are not best suited. Several alternative specifications can be used and in particular Tobit regressions (accounting for the censored nature of the data) and Negative Binomial regressions (accounting for the count type of the data, in the presence of overdispersion).

---

[12]ie. comparing caste A, SC since 1950 in the region to caste B, becoming SC in the region with the Area Restriction Removal Act.

[13]Time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

This section details the results of the Negative binomial estimation, while Appendix E details the very similar results of the Tobit[14]. It can be seen in Table 3 that the point estimate on the coefficient of interest, sc76*t76 is small and non significant, pointing to an absence of effect of the positive discrimination policy on education levels. Regressions are reported with and without the sampling weights.

[Table 2 about here.]

3.2   Primary and secondary school completion

However, focusing on the number of years of education, I had so far implicitly assumed a smooth effect of the policy on the evolution of education. This is a strong assumption as it is more likely that the education choices target certain thresholds of education (completed primary education, completed secondary education...) more than a number of years of education. To account for this possibility, I use probit regressions on the outcomes "having completed primary education"[15] and "having completed secondary education"[16]. Table 4 and 5 respectively report the results of the probit regressions on primary and secondary schooling completion. It can be seen that in both cases, the interaction between sc76 and t76 is small and non significant, pointing once again to the absence of overall effect of the policy.

[Table 3 about here.]

[Table 4 about here.]

3.3   Urban vs rural areas

If the effect of the policy seems to have been overall negligible, Figure 5 had pointed to a potential heterogeneity of the effect across urban and ural areas. In order to explore further this aspect, I will allow the effect of the policy to be heterogenous:

$$
\begin{aligned}
Edu_{idl} = constant &+ \beta sc76_{id} + \delta sc76_{id} * t76_i \\
&+ sc76_{id} * t76_i * urban_l + t76_i * urban_l + sc76_{id} * urban_l \\
&+ \gamma t76_i + \lambda X_{idl} + \epsilon_{idl}
\end{aligned}
\tag{2}
$$

---

[14]OLS results are also reported in Annex F for reference.

[15]As the normal age to complete primary education is 14, I do not restrict the sample to the cohorts aged above 21 anymore, but to those aged above 14 for this outcome. Restricting to cohorts born aged over 21 does not alter the results.

[16]As the normal age to complete secondary education is 18, I do not restrict the sample to the cohorts born aged above 21 anymore, but to those born aged above 18 for this outcome. Restricting to cohorts aged above 21 does not alter the results.

With the same notation as earlier and $urban_l$ a dummy indicating whether the location in which individual i resides is urban or rural. The introduction of caste fixed and state fixed effects or of area of restriction fixed effect will depend on the identification strategy followed. Indeed, when allowing the effect of the policy to be heterogenous across urban and rural areas, the picture is somehow less pessimistic, as the point estimate on sc76*urban*t76 is relatively large and positive in all specifications, and significant in the all except the weighted within caste regression. A wald test of the significance of the sum of the coefficients on sc76*t76 and sc76*urban*t76 in the within area specifications confirms that the global effect of the treatment in urban areas is significantly different from 0[17], but not significantly different from 0 in the within caste specifications.

[Table 5 about here.]

Even more so, when looking at the primary school completion (reported in Table 7), the urban/rural divide this time appears significant in all specifications[18], and the overall effect of the policy in urban areas (sum of sc76*t76 and sc76*urban*t76) is positive and significant (at the 10% level) for the within area identification strategy.

[Table 6 about here.]

3.4 Supply of schooling

As the policy seems to have had some effect, if any, only in the urban areas, one potential explanation for this fact could come from the fact that urban areas had a better supply of schooling (both in terms of quality and quantity). In order to explore this possibility, using the data from the District Information System on Education (round 2007), I have built a district level (distinguishing urban and rural areas within districts) measure of the number of school per child of school age, both for primary schools (for children aged under 15) and secondary schools (for children aged under 20[19]). As underlined in the Descriptive Statistics section, the supply of schooling was very poor in the 1970's, which might explain the absence of effect of the policy: in the absence of school, a policy aiming at favoring the education of a sub population is not likely to be successful. In particular, the presence of a secondary school might provide a sufficient incentive to complete primary schooling. Finally, the effect of the policy are mainly affecting

---

[17]At the 5% level, in both weighted and unweighted regressions.

[18]Secondary education completion, not reported, is not significantly affected.

[19]Secondary school is in theory completed at 18. However, the age data from the Census does not allow to measure the population aged 18 or under.

secondary schooling (free for SC, and with various scholarship schemes aiming at favoring secondary education). I will thus run regressions of the type:

$$
\begin{aligned}
Edu_{idl} = constant &+ \beta sc76_{id} + \delta sc76_{id} * t76_i \\
&+ sc76_{id} * t76_i * secondary_l + t76_i * secondary_l + sc76_{id} * secondary_l \\
&+ \gamma t76_i + \lambda X_{idl} + \epsilon_{idl}
\end{aligned}
\tag{3}
$$

With the same notation as earlier, and $secondary_l$ the number of secondary schools per children aged 20 or under in 1977, for the location l, the urban or rural area of a district. This variable is centered and reduced in order to ease its interpretation: its coefficient will give the effect of an increase of the supply of schooling of a standard deviation above the mean. Table 8 gives the results of the probit regression of primary school completion on the treatment interacted with the supply of school[20]. It can be seen that in the places where the supply of secondary schools was average (coefficient on sc76*t76), the effect of the policy was insignificant (and the point estimates are small). In the places in which the supply of secondary schools was above the average, being added to the SC list lead to an increase in the probability to complete primary education. Hence, the results are clearly in line with an interpretation of the failure of the policy due to a failure to provide education facilities.

[Table 7 about here.]

4   Conclusion

This paper studies the impact of the positive discrimination policy in education conducted by the Indian Government since the Independence using a natural experiment. It shows, using two identification strategies, that the impact of reservations in education has been, at best, mixed. Indeed, only the SC residing in urban areas benefit from the policy, showing an increase in primary school completion. The reason of this absence of impact of the policy (and its differential impact between urban and rural areas) seems to come from the very poor supply of schooling which still continues to plague the Indian school system (Duflo et al., 2010): when allowing for a differential impact between the places with a higher supply of schooling in 1977 with others, the SC seem to benefit

---

[20]The results of Tobit and Negative Binomial regressions are not reported, but provide positive and non significant point estimates on the interaction term sc76*difsecondary*t76, pointing to the fact that the effect might be better estimated using levels of education attained than plain number of years of schooling.

from the policy in the places with a better school supply. In terms of policy making, to answer the questioning of Galanter (1984) cited in the introduction, it seems that most of the improvement of the education of the SC does not come from the reservations, but from the general improvement of the supply of schooling. Hence, in a context of poor supply of schooling, emphasize must be put first on improving the supply itself before trying to put the accent on the access to schooling of certain sub categories.

References

Bayly, Susan, *Caste, society and politics in India*, Cambridge University Press, 1999. Cited on page 4.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova, "Powerful Women: Does Exposure Reduce Bias?," *Quaterly Journal of Economics*, 2009, *124* (4), 1497–1540. Cited on page 2.

Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan, "Affirmative Action in Education: Evidence from college admissions in India," *Journal of Public Economics*, 2010, *94* (1-2), 16–29. Cited on page 2.

Cassan, Guilhem, "Law and identity manipulation: evidence from a natural experiment in colonial Punjab.," *Working Paper*, 2011. Cited on page 8.

Chalam, K.S., "Caste reservations and equality of opportunity in education," *Economic and Political Weekly*, 1990, *25* (41). Cited on page 2.

Chattopadhyay, Raghabendra and Esther Duflo, "Women as Policy Makers: Evidence from a India-Wide Randomized Policy Experiment," *Econometrica*, 2004, *72* (5), 1409– 1444. Cited on page 2.

Chitnis, Suma, "Education for equality: case of Scheduled Castes in Higher Education," *Economic and Political Weekly*, 1972, *7* (31-33). Cited on page 2.

Duflo, Esther, Rema Hanna, and Stephen Ryan, "Incentives Work: Getting Teachers to Come to School," *Working Paper*, 2010. Cited on page 21.

Galanter, Marc, *Competing equalities: law and the backward classes in India*, University of California Press, 1984. Cited on pages 2, 4, 5, 6, and 21.

Government of India, ed., *Report of the Commissioner for Scheduled Castes and Scheduled Tribes, 1957-1958.*, Manager Govt. Of India Press, 1958. Cited on page 6.

— **, ed.**, *Report of the Commissioner for Scheduled Castes and Scheduled Tribes, 1958-1959.*, Manager Govt. Of India Press, 1960. Cited on page 6.

— **, ed.**, *Report of the Commissioner for Scheduled Castes and Scheduled Tribes, 1975-1977.*, Manager Govt. Of India Press, 1978. Cited on page 7.

Holzer, Harry and David Neumark, "Affirmative Action: what do we know ?," *Journal of Policy Analysis and Management*, 2006, *25* (2), 463–490. Cited on page 2.

Howard, Larry and Nishit Prakash, "Does Employment Quota Explain Occupational Choice Among Disadvantaged Groups? A Natural Experiment from India," *Working Paper*, 2008. Cited on page 2.

Jaffrelot, Christophe, *Dr. Ambedkar. Leader Intouchable et père de la Constitution indienne*, Presses de Sciences Po, 2000. Cited on page 26.

Kumar, Dharma, "The Affirmative Action Debate in India," *Asian Survey*, 1992, *32* (3), 290–302. Cited on page 2.

Pande, Rohini, "Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence ? Theory and Evidence from India.," *American Economic Review*, 2003, *93*, 1132–1151. Cited on page 2.

Prakash, Nishit, "The Impact of Employment Quotas on the Economic Lives of Disadvantaged Minorities in India," *Working Paper*, 2009. Cited on page 2.

Singh, K.S., ed., *People of India: Communities, Segments, Synonyms, Surnames and Titles*, Anthropological Survey of India, 1996. Cited on pages 8 and 25.

Vanneman, Reeve and Douglas Barnes, "Indian District Data, 1961-1991: machine-readable data file and codebook," http://www.inform.umd.edu/ districts/index.html 2000. Cited on pages 8 and 29.

Appendices

A   Caste synonyms used

Each caste name can have various synonyms. While the list of SC provides certain synonyms for each caste, which are the caste names required to have access to the SC status. However, given that the survey was not an official administrative task, it is likely that the respondent have answered with the name of the caste they most commonly use more than with the synonym officially recognized. Indeed, a large number of respondents declare themselves as SC while their caste name is not on the SC list of the state. While those individuals might be migrants coming from states in which their caste is recognized as a SC (or are just wrong in declaring themselves as being SC), I believe that it is likely that in the context of the survey, they declare the caste name that they most commonly use, while confronted to the administration, they would declare their caste under the synonym name listed on the official list. Hence, I have coded each caste according to the official list, complemented with the list of synonyms provided by the volume on caste names synonyms of the People of India project (Singh, ed, 1996), if the respondent was declaring herself as being a member of a SC.Finally, in the cases in which the list of synonyms were different across areas of restrictions, then only the synonyms listed in the official list were used to code the timing of the treatment status[21]. Table 9 list by state all the choices made in terms of coding jatis as synonyms. If not otherwise specified, the source is Singh, ed (1996).

[Table 8 about here.]

---

[21]Say if castes A, A1 and A2 are synonyms, but only castes A and A1 are considered as SC in region 1 and only castes A and A2 are considered as SC in region 2, while after the Area Restriction Removal, they are listed as a synonyms, eligible everywhere, then an individual declaring her caste as A1 will be coded as having access to the reservations from 1950 in region 1 and from 1976 in region 2.

B Declaring oneself as a Harijan.

Since the identification strategy relies on the possibility to code each individual to a SC status through her caste name, the fact that a large share of individuals declare themselves as belonging either to the "Harijan" caste or the "Scheduled Castes" caste is a concern for our results. Table 10 details the share of persons individuals declaring their caste as Harijan or Scheduled Caste.

[Table 9 about here.]

Table 11, which shows the results of an OLS regression of the number of years of schooling on a "Harijan" dummie, clearly pictures that the selection into declaring oneself as a "Harijan" is correlated with the education level. Indeed, the individuals declaring themselves member of a SC and Harijan exhibit a lower level of education than the other SC. However, Table 12 shows the results of a probit regression of "declaring one's caste name as Harijan or SC" on the treatment variable (t77) and various controls. The coefficient on the treatment variable is both very small and non significant indicating that the tendency to declare oneself as a "Harijan" is not correlated with the timing of the treatment.

[Table 10 about here.]

[Table 11 about here.]

20

C   Descriptive statistics of the variable used.

The variables used as control come from the Indian District Database (Vanneman and Barnes, 2000). The districts have been made comparable over time between 1961 and 1991, and it is this definition of district that is used in the paper. Table 13 present the descriptive statistics of the district level variables used as controls in the regressions. The data from supply of schooling is from the DISE (2007) database. Table 14 gives the descriptive statistics of the secondary school supply variables, at the district*urban/rural level. The number of observation is lower than 200 because two districts are fully urban. Finally, Table 15 gives the descriptive statistics of the individual level variables taken from the DHS, for the population under study.

[Table 12 about here.]

[Table 13 about here.]

[Table 14 about here.]

D Regressions on cohorts fully treated.

Tables 16 and 17 report the results of regression described in Equations 1 and 2, but only on the cohorts fully treated (i.e. aged either 6 and under or 18 and above in 1977). The $t76$ treatment variable takes a value 1 for the cohorts aged 6 and under in 1977, 0 for the cohorts aged 18 and above in 1977. Tables 16 and 17 report the results of such a specification. It can be seen that they remain very similar to the results found on the full sample.

[Table 15 about here.]

[Table 16 about here.]

E Tobit regressions

The Tobit regressions on the number of years of schooling are reported in Table 18. They account for both the top coding of the number of years of schooling at 15 as well as of their positivity constraint. It can be seen that the point estimates on sc76*t76 is negative and non significant in all specifications (slightly significant in the unweighted fixed effect regression, and positive and non significant in the last specification). Table 19 reports the results of a Tobit regression allowing for an heterogeneous effect between urban and rural areas. It can be seen that the results obtained are robust to using a Tobit specification, with the coefficient on sc76*urban*t76 being positive and significantly estimated. It can also be noted that the coefficient on sc76*t76 is negative and significant in the within jati specifications. It is very hard to make sense of this coefficient, and it is reassuring to see that it loses its significance when a within area of restriction specification is used, putting forward the possibility that it was due to geographical determinants.

[Table 17 about here.]

[Table 18 about here.]

23

F OLS estimates

Table 20 reports the results of an OLS regression obtained for those two identification strategies.

[Table 19 about here.]

Indeed, it can be seen in Table 21 that the coefficient on the interaction of urban*sc76 with t76 is positive and significant in all specifications. However, the coefficient on sc76*t76 is negative in all specifications, and significant in most of them. This suggests that for rural areas a divergence between the two groups took place, precisely when both groups had access to the benefits of the status. It is very hard to make sense of this coefficient, and it is reassuring to see that it loses its significance when a within area of restriction specification is used, putting forward the possibility that it was due to geographical determinants.

[Table 20 about here.]

G   1956 and 1976 Scheduled lists

Table 2: Maharashtra's 1956 list of SC.

Throughout the state except the districts of Buldana, Akola, Amravati, Yeotmal, Wardha, Nagpur, Bhandara, Chanda, Aurangabad, Parbhani, Nanded, Bhir, Osmanabad and Rajura:	In the district of Akola, Amravati and Buldana:
Ager	Bedar
Bakad or Bant	
Bhambi, Bhambhi, Asadaru, Asodi, Chamadia, Chamar, Chambhar, Chamgar, Haralayya, Harali, Khalpa, Machigar, Mochigar, Madar, Madig, Mochi, Telegu Mochi, Kamati Mochi, Ranigar, Rohidas, Rohit or Samgar	**In the districts of Bhandara:**
Bhangi, Mehtar, Olgana, Rukhi, Malkana, Halalkhor, Lalbegi, Balmiki, Korar or Zadmalli	Chadar
Chalvadi or Channayya	Holiya
Chenna Dasar or Holaya Dasar	
Dhor, Kakkayya or Kankayya	**In the districts of Bhandara and Buldana:**
Garoda or Garo	Khangar, Kanera or Mirdha
Halleer	
Halsar, Haslar, Hulasvar or Halasvar	**In the districts of Amravati, Bhandara and Buldana:**
Holar of Valhar	Kori
Holaya or Holer	
Lingader	**In the districts of Aurangabad, Parbhani, Nanded, Rajura, Bhir and Osmanabad:**
Mahar, Taral of Dhegu Megu	Anamuk
Mahyavanshi, Dhed, Vankar or Maru Vankar	Araya (Mala)
Mang, Matang or Minimadig	Arwa Mala
Mang Garudi	Beda (Budga) Jangam
Meghval or Menghvar	Bindla
Mukri	Byagara
Nadia or Hadi	Chalvadi
Pasi	Chambhar
Shenva, Chenva, Sedma or Ravat	Dakkal (Dokkalwar)
Tirgar or Tirbanda	Dhor
Turi	Ellamalwar (Yellammalawandlu)
	Holeya

Maharashtra's list of SC of 1956 (continued)	
In the districts of Buldana, Akola, Amravati, Yeotmal, Wardha, Naghpur, Bhandara and Chanda:	Holeya Dasari
Bahna or Bahana	Kolupulvandlu
Balahi or Balai	Madiga
Basor, Burud, Bansor or Bansodi	Mahar
Chamar, Chamari, Mochi, Nona, Rohidas, Ramnami, Satnami, Surjabanshi or Surjyaramnami	Mala
Dom or Dumar	Mala Dasari
Dohor	Mala Hannai
Ganda or Gandi	Malajangam
Ghasi or Ghasia	Mala Masti
Kaikadi	Mala Sale (Netkani)
Katia or Patharia	Mala Sanyasi
Khatik, Chikwa or Chikvi	Mang
Madgi	Mang Garodi
Mahar or Mehra	Manne
Mang, Dankhni Mang, Mang Mahashi, Mang Garudi, Madari Garudi or Radhe Mang	Mashti
Mehtar or Bhangi	Mehtar
Sansi	Mitha Ayyalvar
	Mochi
	Samagara
	Sindhollu (Chindollu)

Table 3: Maharashtra's 1976 list of SC.

Throughout the state:	**Throughout the state:**
Ager	Kolupulvandlu
Anamuk	Kori
Aray Mala	Lingader
Arwa Mala	Madgi
Bahna, Bahana	Madiga
Bakad, Bant	Mahar, Mehra, Taral, Dhegu Megu
Balahi, Balai	Mahyavanshi, Dhed, Vankar, Maru Vankar
Basor, Burud, Bansor, Bansodi	Mala
Beda Jangam, Budga Jangam	Mala Dasari
Bedar	Mala Hannai
continued on next page...	

Maharashtra's list of SC of 1976 (continued)	
Bhambi, Bhambhi, Asadaru, Asdoi, Chamadia, Chamar, Chamari, Chambhar, Chamgar, Haralayya, Harali, Khalpa, Machigar, Mochigar, Madar, Madig, Mochi, Telegu Mochi, Kamati Mochi, Ranigar, Rohidas, Nona, Ramnami, Rohit, Samgar, Samagara, Satnami, Surjyabanshi, Surjyaramnami	Mala Jangam
Bhangi, Mehtar, Olgana, Rukhi, Malkana, Halalkhor, Lalbegi, Balmiki, Korar, Zadmalli	Mala Masti
Bindla	Mala Sale, Netkani
Byagara	Mala Sanyasi
Chalvadi, Channayya	Mang, Matang, Minimadig, Dankhni Mang, Mang Mahashi, Madari, Garudi, Radhe Mang
Chenna Dasar, Holaya Dsar, Holeya Dasari	Mang Garodi, Mang Garudi
Dakkal, Dokkalwar	Manne
Dhor, Kakkayya, Kankayya, Dohor	Mashti
Dom, Dumar	Meghval, Menghvar
Ellemalvar, Yellammalawandlu	Mitha Ayyalvar
Ganda, Gandi	Mukri
Garoda, Garo	Nadia, Hadi
Ghasi, Ghasia	Pasi
Halleer	Sansi
Halsar, Haslar, Hulasvar, Halasvar	Shenva, Chenva, Sedma, Ravat
Holar, Valhar	Sindhollu, Chindollu
Holaya, Holer, Holeya, Holiya	Tirgar, Tirbanda
Katia, Patharia	Turi
Khangar, Kanera, Mirdha	
Khatik, Chikwa, Chikvi	**In Akola, Amravati, Bhandara, Buldana, Nagpur, Wardha and Yavatmal districts and Chandrapur district other than Rajura Tahsil:**
	Kaikadi

Table 4: Karnataka's 1956 list of SC.

Throughout the State except Coorg, Belgaum, Bijapur, Dharwar, Kanara, South Kanara, Gulbarga, Raichur and Bidar districts and Kollegal taluk of Mysore district:	**In South Kanara district and Kollegal taluk of Mysore district:**

Karnataka's list of SC of 1956 (continued)	
Adi Andhra	Adi Andhra
Adi Dravida	Adi Dravida
Adi Karnataka	Adi Karnataka
Banjara or Lambani	Ajila
Bhovi	Arunthatiyar
Dakkaliga	Baira
Ganti Chores	Bakuda
Handi Jogis	Bandi
Kepmaris	Bellara
Koracha	Chakkiliyan
Korama	Chalavadi
Machala	Chamar or Muchi
Mochi	Chandala
Sillekyathas	Cheruman
Sudugadu Siddha	Devendrakulathan
	Dom, Dombara, Paidi or Pano
In the districts of Belgaum, Bijapur, Dharwar and Kanara:	Godagali
Ager	Godda
Bakad or Bant	Gosangi
Bhambi, Bhambhi, Asadaru, Asodi, Chamadia, Chamar, Chambhar, Chamgar, Haralayya, Harali, Khalpa, Machigar, Mochigar, Madar, Madig, Mochi, Telegu Mochi, Kamati Mochi, Ranigar, Rohidas, Rohit or Samgar	Holeya
Bhangi, Mehtar, Olgana, Rukhi, Malkana, Halalkhor, Lalbegi, Balmiki, Korar or Zadmalli	Jaggali
Chalvadi or Channayya	Jambuvulu
Chenna Dasar or Holaya Dasar	Kadaiyan
Dhor, Kakkayya or Kankayya	Kalladi
Garoda or Garo	Karimpalan
Halleer	Koosa
Halsar, Haslar, Hulasvar or Halasvar	Kudumban
Holar or Valhar	Kuruvan
Holaya or Holer	Madari
Lingader	Madiga
Mahar, Taral or Dhegu Megu	Maila
Mahyavanshi, Dhed, Vankar or Maru Vankar	Mala
Mang, Matang or Minimadig	Mavilan
Mang Garudi	Moger
Meghval or Menghvar	Mundala
Mukri	Nalakeyava

Nadia or Hadi	Nayadi
Pasi	Pagadai
Shenva, Chenva, Sedma or Ravat	Pallan
Tirgar or Tirbanda	Pambada
Turi	Panchama
	Panniandi
In Kanara district:	Paraiyan
Kotegar or Metri	Puthirai Vannan
	Raneyar
In the districts of Gulbarga, Bidar and Raichur:	Samagara
Anamuk	Samban
Aray (Mala)	Sapari
Arwa Mala	Semman
Beda (Budga) Jangam	Thoti
Bindla	Tiruvalluvar
Byagara	Valluvan
Chalvadi	
Chambhar	**In Kollegal Taluk of Mysore:**
Dakkal (Dokkalwar)	Pannadi
Dhor	Vathiriyan
Ellamalwar (Yellammalawandlu)	
Holeya	**In South Kanara district:**
Holeya Dasari	Bathada
Kolupulvandlu	Hasla
Madiga	Nalkadaya
Mahar	Paravan
Mala	
Mala Dasari	**In Coorg district:**
Mala Hannai	Adi Dravida
Malajangam	Adi Karnataka
Mala Masti	Adiya
Mala Sale (Netkani)	Balagai
Mala Sanyasi	Holeya
Mang	Madiga
Mang Gorodi	Muchi
Manne	Mundala
Mashti	Pale
Mehtar	Panchama
Mitha Ayyalvar	Paraya
Mochi	Samagara
Samagara	

Karnataka's list of SC of 1956 (continued)	
Sindhollu (Chindollu)	

Table 5: Karnataka's 1976 list of SC.

Throughout the state:	Throughout the state:
Adi Andhra	Kudumban
Adi Dravida	Kuravan
Adi Karnataka	Lingader
Ager	Machala
Ajila	Madari
Anamuk	Madiga
Aray Mala	Mahar, Taral, Dhegu, Megu
Arunthathiyar	Mahyavanshi, Dhed, Vankar, Maru Vankar
Arwa Mala	Maila
Baira	Mala
Bakad	Mala Dasari
Bakuda	Mala Hannai
Balagai	Mala Jangam
Bandi	Mala Masti
Banjara, Lambani	Mala Sale, Netkani
Bathada	Mala Sanyasi
Beda Jangam, Budga Jangam	Mang, Matang, Minimadig
Bellara	Mang Garudi, Mang Garodi
Bhangi, Mehtar, Olgana, Rukhi, Malkana, Halalkhor, Lalbegi, Balmiki, Korar, Zadmalli	Manne
Bhambi, Bhambhi, Asadaru, Asodi, Chamadia, Chamar, Chambhar, Chamgar, Haralayya, Harali, Khalpa, Machigar, Mochigar, Madar, Madig, Mochi, Telegu Mochi, Kamati Mochi, Ranigar, Rohidas, Rohit or Samgar	Masthi
Bhovi	Mavilan
Bindla	Meghval, Menghvar
Byagara	Moger
Chakkiliyan	Mukri
Chalavadi, Chalvadi, Channayya	Mundala
Chandala	Nadia, Hadi
Chenna Dasar, Holaya Dasar	Nalkadaya
Dakkal, Dokkalwar	Nalakeyava
Dakkaliga	Nayadi
	continued on next page...

Karnataka's list of SC of 1976 (continued)	
Dhor, Kakkayya, Kankayya	Pale
Dom, Dombara, Paidi, Pano	Palian
Ellamalwar, Yellammalawandlu	Pambada
Ganti Chores	Panchama
Garoda, Garo	Panniandi
Godda	Paraiyan, Paraya
Gosangi	Paravan
Halleer	Raneyar
Halsar, Haslar, Hulasvar, Halasvar	Samagara
Handi Jogis	Samban
Hasla	Sapari
Holar, Valhar	Sillekyathas
Holaya, Holer, Holeya	Sindhollu, Chindollu
Holeya Dasari	Sudugadu Siddha
Jaggali	Thoti
Jambuvulu	Tirgar, Tirbanda
Kadaiyan	Valluvan
Kalladi	
Kepmaris	**In Coorg district:**
Kolupulvandlu	Adiya
Koosa	
Koracha	**In Belgaum, Bijapur, Dharwar and North Kanara districts:**
Korama	Bant
Kotegar, Metri	

Table 6: Madhya Pradesh's 1956 list of SC.

In the districts of Bhind, Gird, Morena, Shivpuri, Guna, Rajgarh, Shajapur, Ujjain, Ratlam, Mandsaur, Indore, Dewas, Dhar, Jhabua and Nimar (MB):	**In the districts of Balaghat, Bilaspur, Durg, Raipur, Surguja, Bastar and Raigarh:**
Bagri or Bagdi	Ghasi or Ghasia
Balai	
Banchada	**In the districts of Balaghat, Betul, Bilaspur, Durg, Nimar, Raipur, Bastar, Surguja and Raigarh and in Hoshangabad and Seoni Malwa tahsils of Hoshangabad district; in Chhindwara district and in Sagar district:**

Madhya Pradesh's list of SC of 1956 (continued)	
Barahar or Basod	Katia or Patharia
Bargunda	
Bedia	**In Sagar and Damoh districts and in Hoshangabad and Seoni Malwa tahsils of Hoshangabad districts:**
Bhangi or Mehtar	Khangar, Kanera or Mirdha
Bhanumati	
Chamar, Bairwa, Bhambi, Jatav, Mochi or Regar	**In Chhindwara, Seoni, Betul, Jabalpur, Narsimhapur, Sagar, Mandla, Damoh, Nimar, Balaghat, Raipur, Durg, Bastar, Surguja and Raigarh districts; and in Hoshangabad district except Harda and Sohagpur tahsil thereof:**
Chidar	Kori
Dhanuk	
Dhed	**In Chhindwara, Seoni, Betul, Jabalpur, Narsimhapur, Sagar, Mandla, Damoh, Nimar, Balaghat, Raipur, Bilaspur, Durg, Bastar, Surguja and Raigarh districts; and in Hoshngabad district except Harda and Sohagpur tahsil thereof:**
Dom	Mahar or Mehra
Kanjar	
Khatik	**In Sohagpur tahsil of Hoshangabad district:**
Koli or Kori	Rujjhar
Kotwal	
Mahar	**In the districts of Datia, Tikamgarh, Chhatarpur, Panna, Satna, Rewa, Sidhi and Shahdol:**
Mang or Mang Garodi	Basor or Bansphor
Megwal	Beldar or Sunkar
Nat, Kalbelia or Sapersa	Chamar, Ahirwar, Chamar Mangan, Mochi or Raidas
Pardhi	Dharkar, Balmik or Lalbegi
Pasi	Dher
Sansi	Dom
Zamral	Domar or Doris
	Ghasia

In the districts of Chhindwara, Seoni, Betul, Jabalpur, Sagar, Damoh, Mandla, Hoshangabad, Narsimhapur, Nimar, Balaghat, Raipur, Bilaspur, Durg, Bastar, Surguja and Raigarh:	Kuchbandhia
Bahna or Bahana	Kumhar
Balahi or Balai	Mehtar, Bhangi or Dhanuk
Basor, Burud, Bansor or Bansodi	Moghia
Chamar, Chamari, Mochi, Nona, Rohidas, Ramnami, Satnami, Surjyabanshi or Surjyaramnami	Muskhan
Dom or Dumar	Pasi
Ganda or Gandi	Sansia or Bedia
Khatik, Chikwa or Chikvi	
Mang, Dankhni Mang, Mang Mahashi, Mang Garudi, Madari, Garudi or Radhe Mang	**In the districts of Raisen and Sehore:**
Mehtar or Bhangi	Balahi
Sansi	Bansphor or Basor
	Basar
In Bilaspur district:	Bedia
Audhelia	Beldar
	Chamar, Jatav or Mochi
In Sagar and Damoh districts:	Chitar
Chadar	Dhanuk
	Dhobi
In Damoh district:	Dome
Dahait, Dahayat or Dahat	Kanjar
	Khatik
In the districts of Bilaspur, Durg, Raipur, Bastar, Surguja and Raigarh:	Koli or Katia
Dewar	Mang
	Mehar
In Sagar district:	Mehtar or Bhangi
Dhanuk	Pasi
	Sansia
In Balaghat district:	Silawat
Dohor	
Holiya	
Madgi	

Table 7: Madhya Pradesh's 1976 list of SC.

Throughout the state:	Throughout the state:
Audhelia	Kanjar
Bagri, Bagdi	Katia, Patharia
Bahna, Bahana	Khatik
Balahi, Balai	Koli, Kori
Banchada	Khangar, Kanera, Mirdha
Barahar, Basod	Kuchbandhia
Bargunda	Mahar, Mehra, Mehar
Basor, Burud, Bansor, Bansodi, Bansphor, Basar	Mang, Mang Garodi, Mang Garudi, Dhanka Mang, Mang Mahasi, Madari, Garudi, Radhe Mang
Bedia	Meghwal
Beldar, Sunkar	Moghia
Bhangi, Mehtar, Balmik, Lalbegi, Dharkar	Muskhan
Bhanumati	Nat, Kalbelia, Sapera, Navdigar, Kubut
Chadar	Pasi
Chamar, Chamari, Bairwa, Bhambi, Jatav, Mochi, Regar, Nona, Rohidas, Ramnami, Satnami, Surjyabanshi, Surjyaramnami, Ahirwar, Chamar Mangan, Raidas	Rujjhar
Chidar	Sansi, Sansia
Chikwa, Chikvi	Silawat
Chitar	Zamral
Dahait, Dahayat, Dahat	
Dewar	**In Bhopal, Raisen and Sehore districts:**
Dhanuk	Dhobi
Dhed, Dher	**In Bhind, Dhar, Dewas, Guna, Gwalior, Indore, Jhabua, Khargone, Mandsaur, Morena, Rajgarh, Ratlam, Shajapur, Shivpuri, Ujjain and Vidisha districts:**
Dohor	Kotwal
Dom, Dumar, Dome, Domar, Doris	
Ganda, Gandi	**In Chhatarpur, Datia, Panua, Rewa, Satna, Shahdol, Sidhi and Tikamgarh districts:**
Ghasi, Ghasia	Kumhar
Holiya	
	In Bhind, Dhar, Dewas, Guna, Gwalior, Indore, Jhabua, Khargone, Mandsaur, Morena, Rajgarh, Ratlam, Shajapur, Shivpuri, Ujjain and Vidisha districts:
	Pardhi

Table 8: Kerala's 1956 list of SC.

Throughout the state:	In Malabar district:
Chakkiliyan	Adi Andhra
Kuravan, Sidhanar	Adi Dravida
Nayadi	Adi Karnataka
Pallan	Ajila
Paraiyan, Parayan (Sambavar)	Arunthathiyar
Valluvan	Baira
	Bakuda
Throughout the State except Kasaragod taluk of Malabar district:	Bandi
Kanakkan or Padanna	Bellara
Panan	Chamar or Muchi
	Chandala
Throughout the state except Malabar district (excluding Kasagarod taluk):	Cheruman
Paravan	Godagali
	Godda
Throughout the state except Malabar district:	Gosangi
Ayyanavar	Holeya
Bharatar	Kadaiyan
Boyan	Kalladi
Domban	Karimpalan
Kakkalan	Koosa
Kavara	Kudumban
Kootan (Koodan)	Maila
Mannan	Mavilan
Padannan	Moger
Palluvan	Mundala
Pathiyan	Nalakeyava
Perumannan	Pambada
Pulayan or Cheramar	Panchama
Thandan	Puthirai Vannan
Ulladan	Raneyar
Uraly	Samagara
Vallon	Samban
Vannan	Semman
Velan	Thoti
Vetan	

Kerala's list of SC of 1956 (continued)	
Vettuvan	**In Malabar district (excuding Kasagarod taluk):**
	Gavara
	Malayan
	Pulaya Vettuvan
	In Kasagarod taluk of Malabar district:
	Bathada
	Hasla
	Nalkadaya

Table 9: Kerala's 1976 list of SC.

Throughout the State:	Throughout the State:
Adi Andhra	Moger
Adi Dravida	Mundala
Adi Karnataka	Nalakeyava
Ajila	Nalkadaya
Arunthathiyar	Nayadi
Ayyanavar	Padannan
Baira	Pallan
Bakuda	Palluvan
Bandi	Pambada
Bathada	Panan
Bellara	Panchama
Bharatar	Paraiyan, Parayan, Sambavar
Chakkiliyan	Paravan
Chamar, Muchi	Pathiyan
Chandala	Perumannan
Cheruman	Pulayan, Cheramar
Domban	Pulaya Vettuvan
Gavara	Puthirai Vannan
Godagali	Raneyar
Godda	Samagara
Gosangi	Samban
Hasla	Semman
Holeya	Thandan
Kadaiyan	Thoti
Kakkalan	Vallon

Kerala's list of SC of 1976 (continued)	
Kalladi	Valluvan
Kanakkan, Padanna	Vannan
Karimpalan	Velan
Kavara	Vetan
Koosa	Vettuvan
Kootan, Koodan	
Kudumban	**Throughout the state excluding the areas comprising the Malabar district as specified by sub section (2) of section 5 of the States Reorganisation Act, 1956 (37 of 1956):**
Kuravan, Sidhanar	Boyan
Maila	
Mannan	**In the areas comprising the Malabar district as specified by sub section (2) of section 5 of the States Reorganisation Act, 1956 (37 of 1956):**
Mavilan	Malayan

Figure 1: Variation in States' borders in 1956.

Indian States in 1951.

Indian States in 1961.
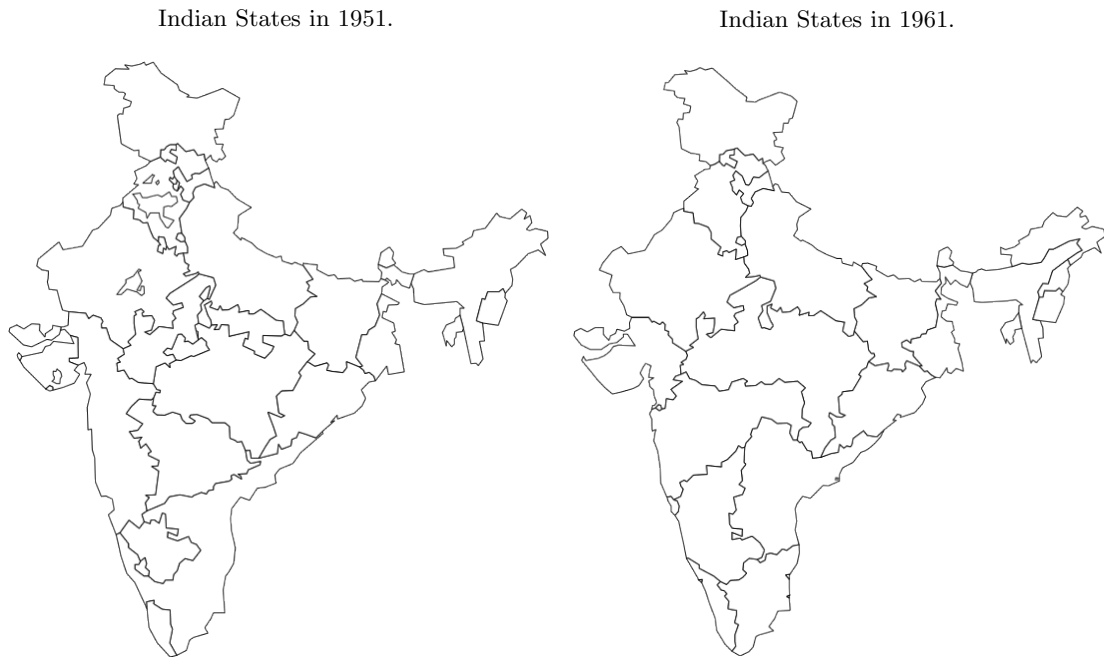
Figure 2: Area Restrictions and pre 1956 state borders

Indian States in 1951.

1956-1976 Area Restrictions.

Indian States in 1971.

Bhopal
Bombay
Coorg
Hyderabad
Madhya Bharat
Madhya Pradesh
Madras
Mysore
Travancore-Cochin
Vindhya Pradesh

Kerala
Madhya Pradesh
Maharashtra
Mysore
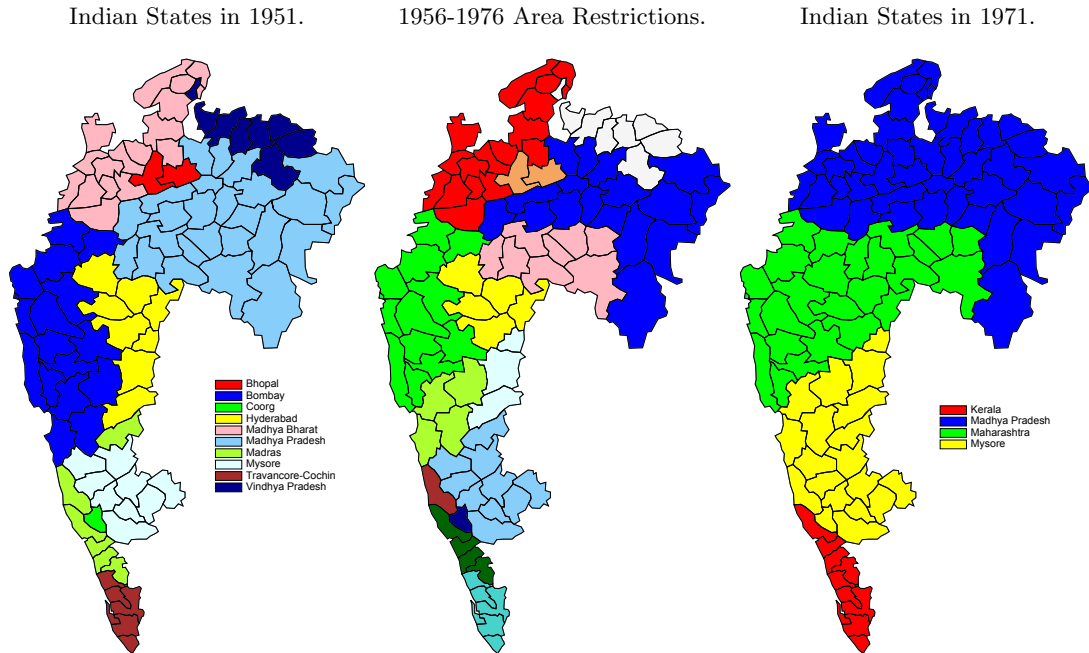
Figure 3: Evolution of years of schooling by SC status.
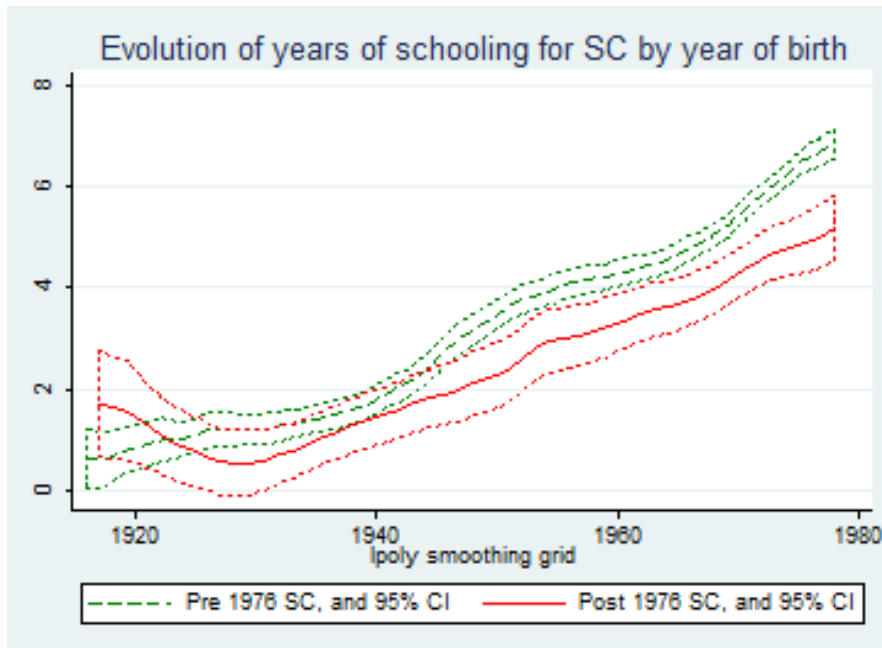
Polynomial smoothing of the years of schooling caped at 15 years.

Figure 4: Histogram of the years of education, for cohorts born between 1944 and 1979.
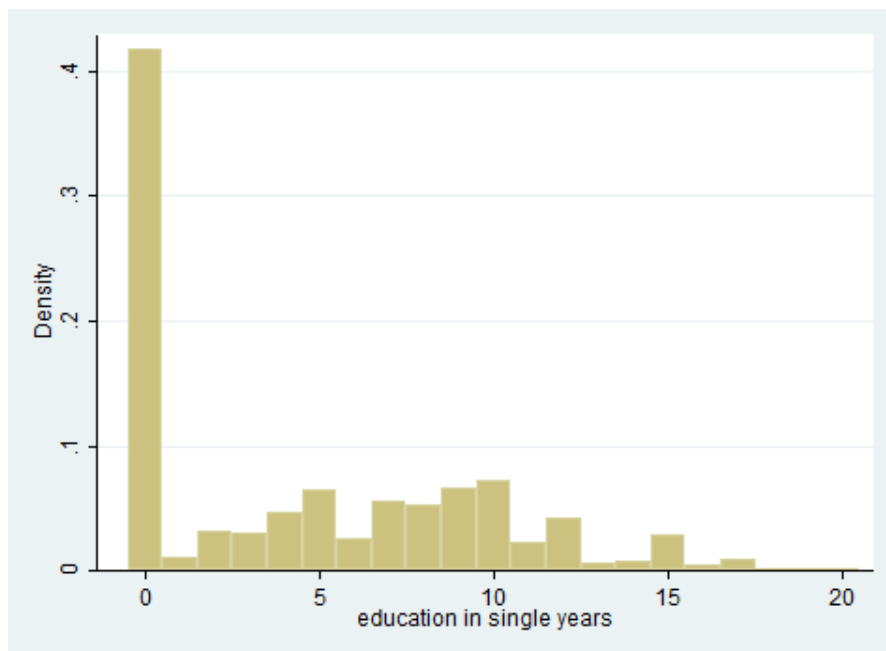
Figure 5: Evolution of years of schooling by SC status, by area of residence.

Urban areas only.
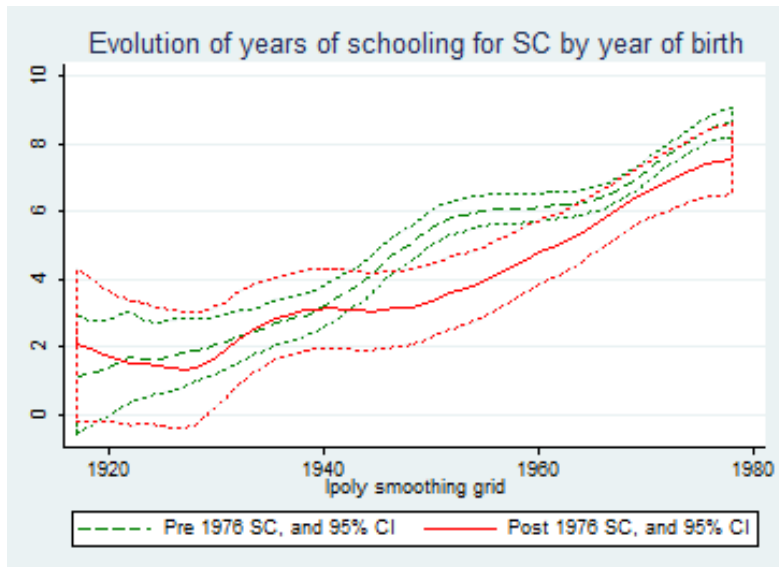
Polynomial smoothing of the years of schooling caped at 15 years.

Rural areas only.

Polynomial smoothing of the years of schooling caped at 15 years.

41
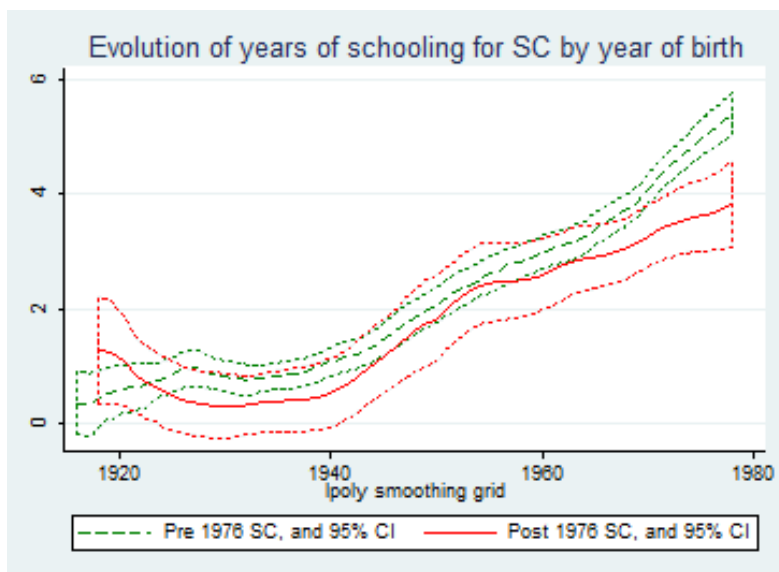
Table 10: Treatment by cohort and SC status.

Cohort	SC since 1950*	SC since 1976*
Birth Year < 1934	Not Treated	Not Treated
1933 < Birth Year < 1945	Ambiguous	Not Treated
1944 < Birth Year < 1960	Treated	Not Treated
1959 < Birth Year < 1971	Treated	Ambiguous
Birth Year > 1970	Treated	Treated

*If the laws were passed in 1950 and 1976, the SC status is considered to actually be implemented in the following year.

Table 11: Negative binomial regression of years of schooling.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**0.0103**	**0.0534**	**0.0779**	**0.109**	**0.140**
	(0.144)	(0.149)	(0.149)	(0.157)	(0.143)
sc76	-0.144	-0.102	-0.131	-0.117	-0.239
	(0.152)	(0.147)	(0.163)	(0.154)	(0.157)
t76	0.00369	-0.0297	0.000556	-0.0465	-0.0163
	(0.0999)	(0.0936)	(0.0971)	(0.118)	(0.115)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	4,617	4,592	4,617	4,592	4,617

Standard errors clustered at the district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 12: Probit regression on primary school completion.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-0.0422**	**0.00349**	**-0.0102**	**0.0334**	**0.0228**
	(0.0658)	(0.0754)	(0.0670)	(0.0791)	(0.0703)
sc76 (d)	-0.0188	-0.00722	-0.0233	0.00429	-0.0447
	(0.0591)	(0.0607)	(0.0646)	(0.0604)	(0.0684)
t76	0.00336	-0.00711	-0.00116	-0.0173	-0.00971
	(0.0475)	(0.0482)	(0.0484)	(0.0566)	(0.0555)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	6,150	6,095	6,150	6,095	6,150

Marginal effects, (d) for discrete change of dummy variable from 0 to 1. Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 13: Probit regression on secondary school completion.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**0.0187**	**0.0364**	**0.0383**	**0.0217**	**0.0378**
	(0.0485)	(0.0511)	(0.0487)	(0.0581)	(0.0514)
sc76 (d)	-0.0597*	-0.0665*	-0.0740**	-0.0620	-0.0993***
	(0.0334)	(0.0357)	(0.0321)	(0.0439)	(0.0283)
t76	0.0137	0.00349	0.00806	0.00553	0.00793
	(0.0347)	(0.0353)	(0.0347)	(0.0430)	(0.0404)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	5,120	4,960	5,120	4,960	5,120

Marginal effects, (d) for discrete change of dummy variable from 0 to 1. Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 14: Negative binomial regression of years of schooling, urban vs rural areas.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-0.258**	**-0.168**	**-0.188**	**-0.0706**	**-0.0773**
	(0.170)	(0.171)	(0.178)	(0.193)	(0.181)
sc76*urban*t76	**0.638****	**0.511***	**0.630****	**0.438**	**0.551****
	(0.269)	(0.288)	(0.255)	(0.311)	(0.255)
sc76	-0.00610	0.0327	0.0517	-0.0294	-0.129
	(0.208)	(0.169)	(0.229)	(0.173)	(0.214)
t76	0.158	0.128	0.156	0.0855	0.109
	(0.116)	(0.107)	(0.112)	(0.129)	(0.127)
urban*t76	-0.366***	-0.362***	-0.367***	-0.326***	-0.316***
	(0.0916)	(0.0954)	(0.0887)	(0.104)	(0.101)
urban	0.946***	0.882***	0.925***	0.813***	0.849***
	(0.116)	(0.107)	(0.113)	(0.113)	(0.117)
sc76*urban	-0.319	-0.303	-0.423	-0.192	-0.271
	(0.268)	(0.275)	(0.264)	(0.285)	(0.251)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	4,617	4,592	4,617	4,592	4,617

Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 15: Probit regression of primary school completion, urban vs rural areas.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-0.171****	**-0.107**	**-0.127***	**-0.0677**	**-0.0779**
	(0.0724)	(0.0863)	(0.0766)	(0.0904)	(0.0809)
sc76*urban*t76	**0.337*****	**0.286****	**0.299*****	**0.287****	**0.283****
	(0.116)	(0.124)	(0.115)	(0.143)	(0.126)
sc76 (d)	0.0638	0.0693	0.0658	0.0732	0.0220
	(0.0705)	(0.0726)	(0.0745)	(0.0711)	(0.0770)
t76	0.0569	0.0465	0.0515	0.0304	0.0346
	(0.0537)	(0.0547)	(0.0547)	(0.0628)	(0.0600)
urban*t76	-0.114**	-0.112**	-0.114**	-0.106	-0.101*
	(0.0459)	(0.0481)	(0.0442)	(0.0645)	(0.0581)
urban (d)	0.354***	0.342***	0.362***	0.322***	0.341***
	(0.0453)	(0.0422)	(0.0462)	(0.0473)	(0.0517)
sc76*urban (d)	-0.205**	-0.186*	-0.218***	-0.188*	-0.185**
	(0.0913)	(0.0983)	(0.0828)	(0.107)	(0.0901)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	6,150	6,095	6,150	6,095	6,150

Marginal effects. (d) for discrete change of dummy variable from 0 to 1. Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 16: Probit regression of primary school completion, secondary school supply.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**0.0563**	**0.117**	**0.0848**	**0.0941**	**0.0733**
	(0.0636)	(0.0790)	(0.0640)	(0.0832)	(0.0718)
sc76*secondary*t76	**0.394*****	**0.409****	**0.396*****	**0.359****	**0.346****
	(0.136)	(0.171)	(0.139)	(0.183)	(0.156)
sc76 (d)	-0.0490	-0.0175	-0.0698	0.00790	-0.0732
	(0.0601)	(0.0715)	(0.0610)	(0.0706)	(0.0691)
t76	0.00377	-0.00538	0.000666	-0.0162	-0.00697
	(0.0479)	(0.0479)	(0.0487)	(0.0564)	(0.0558)
secondary*t76	-0.00901	0.0176	-0.00404	-0.0157	-0.0215
	(0.0429)	(0.0711)	(0.0441)	(0.0794)	(0.0559)
secondary	-0.000572	-0.00358	0.0164	0.0361	0.0269
	(0.0360)	(0.0874)	(0.0313)	(0.0959)	(0.0543)
sc76*secondary	-0.131	-0.0337	-0.190	0.000477	-0.175
	(0.125)	(0.172)	(0.122)	(0.183)	(0.137)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	6,150	6,095	6,150	6,095	6,150

Marginal effects. (d) for discrete change of dummy variable from 0 to 1. Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 17: Caste synonyms

Caste	Synonym	Reason
Karnataka		
Bhovi	Vadda	is a synonym.
	Vodda	is a synonym.
	Wadda	is a synonym.
Kerala		
Cheruman	Cherumakkal	is a synonym.
Madhya Pradesh		
Bagri	Bagh	is a synonym of Bagdi, while for the SC list, Bagdi=Bagri.
Chadar	Athia	is a synonym.
Chamar	Bambi	is a subgroup of Chamar.
	Cobbler	is a traditional occupation of Chamar.
Dhed	Charia	is a synonym.
Dhobi	Rajak	is a synonym.
	Washerman	is a traditional occupation of Dhobi.
Kumhar	Prajapati	is a title of Kulhar.
Mahar	Chaudhary	is a title of Mahar.
	Jharia	is a synonym.
Sansi	Bagaria	is a got of Sansi.
Maharashtra		
Mahar	Budhist, Neo Buddhist	Due to the specific history of the Mahar caste in Maharashtra, most Buddhist of the state are Mahar converted to Buddhism (Jaffrelot, 2000).

Table 18: Share of Harijan in the SC population.

	All States	Kerala	Maharashtra	Madhya Pradesh	Karnataka
Share of Harijan	12.1%	2.5%	4.1%	14.6%	24.8%

Survey weights used. Cohorts born between 1943 and 1979.
SC population restricted to the castes listed as SC in the state.

Table 19: Number of years of education of the Harijans. Negative binomial regression.

	(1)
hari	-0.233***
	(0.0852)
Other Controls	YES
Observations	5,393

Standard errors clustered at the district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 20: Probit regression on declaring one's caste as Harijan.

	(1)	(2)	(3)
t77	0.00681	0.00716	0.0131
	(0.0101)	(0.00924)	(0.0144)
State FE	YES	YES	NO
District FE	NO	NO	YES
Other Controls	NO	YES	YES
N	5,393	5,393	3,772

Standard errors clustered at the district level in parentheses. Marginal effects. *** p<0.01, ** p<0.05, * p<0.1. Controls include: gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 21: District level variables descriptive statistics.

	Mean	Std. Deviation.	N
1977 share of migrants	33.3%	0.059	100
1977 share of literate	41.2%	0.169	100
1977 share of SC	11.8%	0.058	100
1977 share of ST	10.8%	0.167	100
1977 share of urban population	20.3%	0.145	100

Source: Indian District Database (Vanneman and Barnes, 2000).

Table 22: School supply variables descriptive statistics.

	Mean	Std. Deviation.	N
1977 Secondary schools per ¡20	0.000147	0.0002627	198
Difference with mean secondary school per ¡20	0.0000174	0.0002627	198

Source: DISE (2007).

Table 23: Individual level variables.

	Mean	Std. Deviation.	N
Urban	42%	0.49	4,617
Male	50%	0.50	4,617
Hindu	85%	0.36	4,617

Source: DHS (1998-1999). Statistics refer to the sub
sample used in the main regressions.

Table 24: Negative binomial regression of years of schooling, fully treated cohorts.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	-0.0242	0.00129	0.0166	0.0672	0.0870
	(0.156)	(0.157)	(0.165)	(0.162)	(0.156)
sc76	-0.153	-0.0457	-0.0836	-0.0355	-0.197
	(0.162)	(0.150)	(0.189)	(0.151)	(0.178)
t76	0.0542	0.00596	0.0398	0.0182	0.0562
	(0.109)	(0.105)	(0.108)	(0.130)	(0.131)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	2,939	2,923	2,939	2,923	2,939

Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$,
* $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban
area fixed effect, religion fixed effect, and district level controls: 1977 share of SC,
1977 share of ST, 1977 share of urban population, 1977 share of literate population
and 1977 share of migrants.

Table 25: Negative binomial regression of years of schooling, fully treated cohorts, urban/rural heterogeneity.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-0.316**[*]	**-0.255**	**-0.272**	**-0.142**	**-0.158**
	(0.192)	(0.185)	(0.203)	(0.199)	(0.197)
sc76*urban*t76	**0.706**[**]	**0.607**[*]	**0.698**[**]	**0.520**	**0.629**[**]
	(0.320)	(0.345)	(0.308)	(0.364)	(0.308)
sc76	-0.00232	0.133	0.137	0.0970	-0.0427
	(0.228)	(0.178)	(0.261)	(0.179)	(0.241)
t76	0.205[*]	0.164	0.200[*]	0.152	0.189
	(0.121)	(0.114)	(0.119)	(0.137)	(0.140)
urban*t76	-0.373[***]	-0.373[***]	-0.391[***]	-0.346[***]	-0.352[***]
	(0.0944)	(0.100)	(0.0925)	(0.108)	(0.102)
urban	0.958[***]	0.888[***]	0.936[***]	0.834[***]	0.868[***]
	(0.119)	(0.109)	(0.116)	(0.116)	(0.121)
sc76*urban	-0.365	-0.420	-0.519	-0.319	-0.390
	(0.338)	(0.373)	(0.336)	(0.384)	(0.327)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	2,939	2,923	2,939	2,923	2,939

Standard errors clustered at the district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 26: Tobit regression of years of schooling.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-0.330**	**-0.236**	**0.0279**	**-0.0500**	**0.259**
	(0.855)	(0.158)	(0.832)	(0.169)	(0.798)
sc76	-0.625	-0.200	-0.648	-0.254[*]	-1.209
	(0.851)	(0.129)	(0.888)	(0.138)	(0.886)
t76	0.377	0.223	0.350	0.175	0.355
	(0.630)	(0.152)	(0.621)	(0.217)	(0.728)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	4,617	4,617	4,617	4,617	4,617

Standard errors clustered at the district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 27: Tobit regression of years of schooling, urban/rural heterogeneity.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-1.926****	**-1.754*****	**-1.381**	**-1.249*****	**-0.797**
	(0.935)	(0.171)	(0.939)	(0.173)	(1.030)
sc76*urban*t76	**4.011*****	**3.854*****	**3.446****	**3.376*****	**2.830****
	(1.553)	(0.251)	(1.464)	(0.239)	(1.369)
sc76	0.104	0.709***	0.183	0.358**	-0.783
	(1.077)	(0.136)	(1.128)	(0.141)	(1.136)
t76	0.929	0.759***	0.880	0.603***	0.753
	(0.707)	(0.161)	(0.695)	(0.228)	(0.798)
urban*t76	-1.231**	-1.182***	-1.208**	-0.994***	-0.940
	(0.579)	(0.216)	(0.557)	(0.311)	(0.620)
urban	5.271***	4.894***	5.259***	4.540***	4.919***
	(0.780)	(0.188)	(0.739)	(0.270)	(0.766)
sc76*urban	-1.798	-2.201***	-2.043	-1.618***	-1.188
	(1.418)	(0.185)	(1.313)	(0.177)	(1.274)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	4,617	4,617	4,617	4,617	4,617

Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 28: OLS regression of years of schooling.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-0.359**	**-0.310**	**-0.185**	**-0.342**	**-0.163**
	(0.486)	(0.456)	(0.470)	(0.433)	(0.434)
sc76	-0.318	-0.0804	-0.368	0.0359	-0.579
	(0.394)	(0.382)	(0.434)	(0.408)	(0.437)
t76	0.430	0.334	0.409	0.304	0.406
	(0.357)	(0.343)	(0.353)	(0.417)	(0.410)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	4,617	4,617	4,617	4,617	4,617
Adjusted R^2	0.323	0.355	0.342	0.340	0.338

Standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.

Table 29: OLS regression of years of schooling, urban/rural heterogeneity.

	Unweighted regressions			Weighted regressions	
	(1)	(2)	(3)	(4)	(5)
sc76*t76	**-1.102****	**-1.087****	**-0.828**	**-1.017***	**-0.677**
	(0.509)	(0.489)	(0.505)	(0.537)	(0.543)
sc76*urban*t76	**2.217****	**2.409*****	**1.905****	**2.353*****	**1.713****
	(0.983)	(0.843)	(0.934)	(0.831)	(0.794)
sc76	0.0268	0.396	0.0262	0.434	-0.314
	(0.458)	(0.414)	(0.488)	(0.457)	(0.509)
t76	0.478	0.377	0.447	0.344	0.426
	(0.388)	(0.375)	(0.386)	(0.459)	(0.447)
urban*t76	-0.0792	-0.0669	-0.0675	-0.0523	-0.0129
	(0.347)	(0.347)	(0.337)	(0.409)	(0.388)
urban	2.728***	2.559***	2.768***	2.424***	2.670***
	(0.451)	(0.408)	(0.428)	(0.454)	(0.459)
sc76*urban	-1.036	-1.397**	-1.175*	-1.334**	-0.935
	(0.718)	(0.637)	(0.645)	(0.667)	(0.625)
Caste FE	NO	YES	NO	YES	NO
Area of restriction FE	NO	NO	YES	NO	YES
Other Controls	YES	YES	YES	YES	YES
Observations	4,617	4,617	4,617	4,617	4,617
Adjusted R^2	0.323	0.356	0.343	0.341	0.338

Standard errors clustered at the district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include: time trend, state fixed effect, gender fixed effect, urban area fixed effect, religion fixed effect, and district level controls: 1977 share of SC, 1977 share of ST, 1977 share of urban population, 1977 share of literate population and 1977 share of migrants.