Affirmative Action, Education and Gender: Evidence from India.*

Guilhem Cassan†

March 3, 2016

Abstract: *This paper studies the impact of affirmative action on education in India. Taking advantage of a change in the borders of the Indian states which led to within caste variation in exposure to these policies, I can exploit within caste variation and within state variation. The impact of those policies widely differs across genders, with males gaining about 2 years of education while females lose around 1.5 years. I interpret those findings as a reallocation of education expenses from girls to boys in a context of low perceived returns to education of girls combined with credit constraints.*

JEL Classification: I24; O15; H41
Keywords: identity, scheduled caste; quota; affirmative action; gender; India; education.

*I am grateful to Gani Aldashev, Denis Cogneau, Christelle Dumas, Esther Duflo, Hemanshu Kumar, Eliana La Ferrara, Sylvie Lambert, Andreas Madestam, Marco Manacorda, Annamaria Milazzo, Barbara Petrongolo, Rohini Somanathan, Vincenzo Verardi and Yanos Zylberberg, as well as participants in the seminar and conferences in which this paper has been presented. Marion Leturcq's and Lore Vandewalle's comments on an earlier draft vastly improved this paper. I am indebted to Julien Grenet for his help in the data collection, and to Ashwini Natraj, with whom I started this project. This paper is produced as part of the project "Actors, Markets, and Institutions in Developing Countries: A micro-empirical approach" (AMID), a Marie Curie Initial Training Network (ITN) funded by the European Commission. Funding from the CEPREMAP is gratefully acknowledged. This research used resources of the "Plateforme Technologique de Calcul Intensif (PTCI)" (http://www.ptci.unamur.be) located at the University of Namur, Belgium, which is supported by the F.R.S.-FNRS under the convention No. 2.4520.11. The PTCI is member of the "Consortium des Equipements de Calcul Intensif (CECI)" (http://www.ceci-hpc.be). All remaining errors are mine.

†University of Namur, CRED. Email: guilhem.cassan@unamur.be.

Introduction

Affirmative action policies in education are an heavily debated issue in all the countries in which they are implemented. In India, in particular, the country with the largest affirmative action program in the world, those policies have been controversial since they were systematically introduced after the Independence. However, due to the very nature of those programs, it is very difficult to assess their impact on the average beneficiaries, and as a result, we lack evidence on that dimension (as underlined by Chalam (1990), Chitnis (1972) and Galanter (1984)). Indeed, most studies have relied on regression discontinuity design estimates, hence measuring the impact of those programs on the marginal beneficiary (Bertrand et al. (2010), Krishna and Frisancho Robles (2012)). With the exception of Khanna (2013), this paper is, to my knowledge, the first to propose an estimate of the causal impact of affirmative action policies on education for the average beneficiary. In order to do so, I use a natural experiment on the access to the "Scheduled Caste" (SC) status[1] to study its impact on educational attainment. This paper relies on a unique natural experiment on the access to the "SC" status. As a matter of fact, the list of the castes considered as SC were drawn by each Indian state at the Independence[2]. In 1956, the borders of the Indian states have been redrawn, while the lists of SC remained unchanged. This increased greatly the within State discrepancies where the "Scheduled Caste" status of the members of a same caste could vary across space[3]. This situation lasted until 1976, when the list of castes considered as SC were harmonized within each state, allowing 2.4 million individuals (Government of India, ed, 1978) to have access to the SC status. This historical setting allows me to assess the fate of the individuals who had access to the SC status in 1976.

As the treatment status varies across castes, and within castes across time and space, this paper innovates by coding the treatment status of individuals based on their precise caste name instead of their declared beneficiary status[4]. By using the precise caste name I can use demanding specifications, which identify the effects of the policy within a single caste, something that, to the best of my knowledge, has never been done in the literature. As a matter of fact, the main specifications rely on caste*cohorts and

[1] This affirmative action program grants lower castes several benefits including reduced education costs and increased returns to education, as will be discussed further in the paper

[2] With some within states discrepancies in the lists

[3] In practice, each state had several lists of SC, by regions usually grouping several districts. Those regions are listed in Appendix A.

[4] Indeed, most of the micro level literature on affirmative action in India uses the the household's declaration of its SC status as a basis for identifying the "treatment" group (Khanna (2013), Hnatkovska et al. (2012), Hnatkovska et al. (2013), Prakash (2009))

region*cohorts fixed effects, which allow to identify the treatment effect using only the within caste-cohort and within region-cohort variation. In a robustness check, I show that adding a specific trend for the treatment group does not alter the results. Hence, the results presented in this paper can not be driven by any kind of pre-existing trends, and can safely be attributed to the acquisition of the Scheduled Caste status.

I can then show that the overall effect of affirmative action policies on educational attainment is close to non existent. However, very large disparities exist across genders: indeed, once the effect is allowed to vary across gender, it can be seen that males gain about 2 years of education while females lose around 1.5 years[5]. This result is to my knowledge new in the literature, which had so far not focused much on the gender dimension of affirmative action policies, or, when it did (Bertrand et al., 2010), documented a null effect on female, but not a negative effect. A plausible interpretation of those results would be that in a context of lower perceived returns of education of girls and imperfect financial markets, an increase in the returns to schooling would lead to an increase in boy's education whose opportunity cost would have to be compensated by a decrease in girls' education.

Hence, this paper relates to the literature on gender discrimination (Jensen, 2012) and in particular to the study of Foster and Rosenzweig (1999) who show that the differential mortality rate of girls vs boys can increase in response to common local increases of returns to human capital. It is also related to the literature on asymmetric effect of social policies in developing countries (Foster and Rosenzweig (2003), Rosenzweig and Schultz (1982)) and in particular to Ashraf et al. (2015) who find a differential effect by gender of school construction program, which they relate to marriage customs. It is also close to the paper of Qian (2008) on the response of gender mortality to variations in gendered returns to working hours. The paper serves as a cautionary tale to policy makers underlining that certain pro-poor policies can have detrimental effects on part of the targeted population. In the first section of this paper, I will examine the context and the natural experiment exploited in this paper. I will then describe the data as well as the empirical strategy, which will open the way to the presentation of the results, their discussion and various robustness checks (varying treatment across cohorts, differential trends and migration).

[5]Obviously, this does not mean that the level of education of females is decreasing in absolute terms: in a context of increasing access to education, females benefit less to this increase than they would have otherwise.

1 Context

1.1 Affirmative action in India

While the first affirmative action policies for the what was then called the "untouchable" castes were implemented under the British rule, it is not before the Independence that a systematic positive discrimination policy was implemented across India. Also called "reservations", they have 3 main dimensions: legislative seats, education and public employment. Within education, affirmative action consists in various policies. There are quotas in higher education institutions, but secondary schooling was also made free, while each state also has various policies (specific scholarships - for girls in particular, schools and hostels, free mid day meals, etc). In addition, in the mid-1990's quotas were also implemented in panchayat elections for SC among other discriminated groups. Positive discrimination might thus affect schooling through various channels. By reducing the cost of education, it favors longer studies in the cost-benefit arbitrage of the household, while quotas in higher education will help the pursuit of studies after secondary schooling. Moreover, the quotas in public employment increases the returns to education. Hence, this paper does not evaluate the effect of affirmative action *in* the educational sector but the effect of the package of affirmative action policies *on* educational attainment.

1.2 The definition of the Scheduled Castes

This section, drawing from the work of Galanter (1984) provides a short history of the list of Scheduled Castes.

One of the main problems with the making of such classification is that the definition of "untouchability", the criteria to be considered a SC, is not straightforward: as "untouchability" varies in its meaning across the subcontinent, it is hard to create a definition that would apply to the whole country. Indeed, while 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 avoids to define a clear concept and only provides a procedure of designation[6] that each State is to follow. This allowed for the possibility of inconsistencies across States[7] as well as within States, as certain States decided to give the SC status to certain castes only in certain areas. Despite those in-

[6] "*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.*"

[7] Bayly (1999) gives the example of the Khatik caste, considered as SC in Punjab, but classed as a "forward" caste in Uttar Pradesh, a neighboring State at the time of the establishment of the lists.

consistencies, the lists were revised only four times since the Independence[8]. With an increase of 2.4 million SC over an original population of 80 million SC, 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 Orders (Amendment) Act of 1976

In 1956, India reorganized the borders of its States along linguistic lines[9]. But, as the borders of the States were redefined , the State-wise SC lists remained unchanged. This led to a large increase of within state discrepancies in the SC lists, since there were now systematically several SC list per state. Hence, from 1956, the number of castes considered as SC in one part of a State and non SC in an other part of the same State vastly increased.It is only in 1976 that the SC lists were harmonized within states by the SC and ST (Amendment) Act of 1976, also called the "Area Restriction Removal Act"[10]. This Act removed almost all intra-State restrictions[11]. This removal of restrictions led to an increase of 2.4 million of the SC population[12].

Hence, individuals from the same caste could be considered as SC in one part of a State but not in an other until 1976, when anyone whose caste was on the SC list

[8]The first change, in 1951 was a matter of correcting small anomalies. The second change, in 1956 mainly affected Rajasthan and Uttar Pradesh, and also allowed all Sikh untouchable castes to claim SC status, while the fourth change of 1990 allowed the Buddhists to have access to the SC status in all the States. The most important change took place in 1976 and is described in more details in the paper.

[9]And in 1960, the states of Maharashtra and Gujarat were created from the former state of Bombay.

[10]The reason for the list not to be adjusted to the new borders was the slowness of the administration: *"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, 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 many of its yearly occurrences 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).*

[11]According the Galanter (1984) their number dropped from 1,126 to 64.

[12]The states affected by the reform were Andhra Pradesh, Bihar, Gujarat, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Rajasthan, Uttar Pradesh, Tamil Nadu, West Bengal and Himachal Pradesh. The analysis excludes Himachal Pradesh, as when this Union Territory was granted the State status in 1966, large portions of the state of Punjab were also transferred to it: hence, between 1956 and 1966, the population living in the contemporary borders of Himachal Pradesh were exposed to different State policies, preventing me from identifying the sole effect of the access to the SC status. Results are unaffected by the inclusion of Himachal Pradesh. Note that if the changes in border mainly affected the South of India, the Area Restriction Removal Act also affected northern states such as Bihar and Uttar Pradesh, who had within states variations in their SC list before 1956.

somewhere in a state would be considered SC in that state.

This situation thus creates a natural experiment setting in which the members of a same caste faced different access to affirmative action status within the same state.

This unique historical event creates a plausible exogenous variation in the SC status across individuals, allowing me to assess the causal impact of the SC status on educational attainment.

2 Data and Empirical Strategy

2.1 Data and descriptive statistics

The National Family and Health Survey of 1998-99 is, to my knowledge, the only dataset offering both the precise caste name (the "jati") of respondents, their district of residence and a sufficient sample size to perform the type of analysis done in this paper[13]. Using the 1971 and 1981 Census lists of Scheduled Castes and the district of residence of households, I am able to identify the households that were granted the SC status in 1976[14]. This methodology differs importantly from most studies of caste using nationally representative surveys. Indeed, while most study rely on the declarations of the respondent on their Scheduled Caste status, I use the jati name of the respondent to attribute the SC status, which allows to have within jati estimates of the access to the SC status[15].

Tables 1 and 2 provide the summary statistics for the variables used throughout the paper for males and females[16]. As expected, one can see that the educational attainment in the sample is lower for females than for males, and that overall, "Old SC" have a higher level of education than "New SC". It can also be noted that the declared religions are somewhat different across groups. All the results shown are robust to restricting the sample to only Hindus[17].

[13]For example, the 2004-5 round of the NFHS does not contain information on the district of residence of the respondents. The first round of the NFHS, conducted at the beginning of the 1990s also offers jati identifiers. However, the documentation does not allow to map the jati identifiers to jati names. Both Indian Human Development Survey round have district and jati identifiers, but the sample size is unfortunately too small for the type of exercise conducted here.

[14]In order to do so, I simply match the 1999 districts of the NFHS data to their 1971 counterparts using the Indian Administrative Atlas of 2001, which follows district changes over time.

[15]Most specifications will control for declaration of the SC status.

[16]I restrict my sample to individuals born between 1951 (below school age in 1956, year of the redefinition of borders) and aged 18 and above at the time of the survey, to allow them to have completed secondary education).

[17]Results available on request.

Table 1: Descriptive Statistics: Male population.

	New SC		Old SC	
	Mean	Std. Dev.	Mean	Std. Dev.
Education variables				
Years of Schooling up to Secondary	4.59	4.11	5.44	4.08
Literacy	0.61	0.49	0.68	0.47
Schooling	0.64	0.48	0.72	0.45
Primary Completion	0.51	0.50	0.61	0.49
Incomplete Secondary	0.45	0.50	0.53	0.50
Secondary Completion	0.20	0.40	0.26	0.44
Control variables				
Urban	0.24	0.43	0.30	0.46
Hindu	0.99	0.11	0.91	0.29
Christian	0.01	0.09	0.03	0.17
Sikh	0.00	0.00	0.00	0.05
Buddhist/Neo Buddhist	0.00	0.00	0.06	0.23
No Religion	0.00	0.04	0.00	0.01
Declares to be SC	0.68	0.47	0.82	0.38
N	680		6640	

Table 2: Descriptive Statistics: Female population.

	New SC		Old SC	
	Mean	Std. Dev.	Mean	Std. Dev.
Education variables				
Years of Schooling up to Secondary	1.91	3.45	2.67	3.79
Literacy	0.26	0.44	0.35	0.48
Schooling	0.28	0.45	0.39	0.49
Primary Completion	0.21	0.41	0.30	0.46
Incomplete Secondary	0.19	0.39	0.25	0.43
Secondary Completion	0.08	0.28	0.11	0.31
Individual and household variables				
Urban	0.20	0.40	0.29	0.45
Hindu	0.98	0.12	0.90	0.29
Christian	0.01	0.11	0.03	0.17
Sikh	0.00	0.00	0.00	0.05
Buddhist/Neo Buddhist	0.00	0.00	0.06	0.24
No Religion	0.00	0.04	0.00	0.00
Declares to be SC	0.67	0.47	0.82	0.38
N	653		6636	

2.2 Identification Strategy

As the SC are a different population from the general population, much poorer in particular, comparing the "new" SC to the general population does not allow to identify the specific effect of the access to the SC status on educational attainment from other social policies.

Hence, the natural counterfactual are the SC already on the lists in 1976. I will call them the "Old SC". This comparison between "Old" and "New" SC is new in the literature and allows to identify the effect of the policy within caste. As the number of treated individuals is relatively small, it was impossible to use an identification strategy Ã la Duflo (2001) which interacts the treatment status and cohort fixed effects: the population in each cell would have been too small. However, I present graphically the point estimates of such regression in Figure 1 with cohorts of 5 years and states fixed effects (and no other controls or fixed effects). The first cohort to be treated is not a priori clearly defined since treated individuals would be the ones still at school at the time of the implementation of the reform. This obviously varies across individuals. This point will be discussed more thoroughly in the results sections, but let us assume for now that the first treated cohorts should be somewhere between 1966 (aged 11 in 1977) and 1971 (aged 6 in 1977). Two striking points emerge. First, no differential trend pre treatment years exists, granting confidence in our difference in differences identification strategy. Second, in the full sample, the cohort exposed to the treatment do not seem to be affected. However, in the gender specific samples, one can see that while males and females were on the same trend as their respective control groups pre-treatment, this widely changed for the treated cohorts, with males' level of education increasing at the expense of females'.

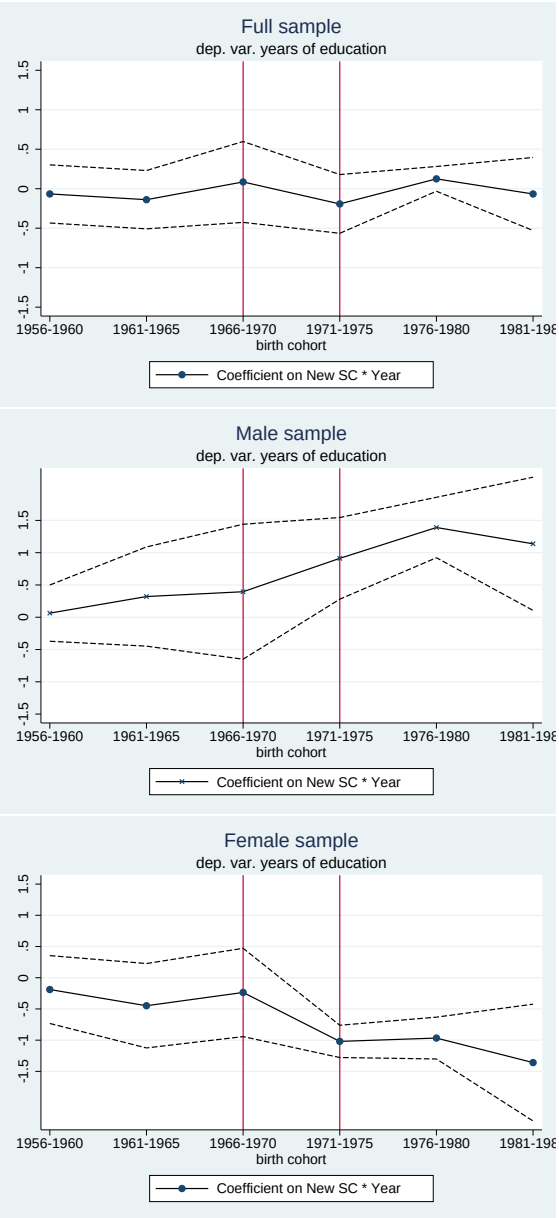
To further this analysis, the identification strategy of this paper will rely on difference in differences. The first difference will compare cohorts too old to benefit from the access to the SC status in 1977[18] from the point of view of their educational attainment to those that were young enough. Throughout the paper, I will consider as "young enough" the cohorts aged either 6 or 11 years old or below in 1977: .

I will thus run regressions of the type:

$$Edu_{idt} = constant + \beta NewSC_{id} + \delta NewSC_{id} * posttreatment_t + \gamma posttreatment_t + \lambda X_{idt} + \epsilon_{idt} \quad (1)$$

[18] Year of implementation of the 1976 change.

Figure 1: Cohort specific effects. Coefficients of the interaction terms "New SC"* years cohorts dummies (controlling for state FE). Population born after 1950 and aged 18 and above at time of survey.



Where Edu_{idt} is a measure of educational attainment of individual i born in year t and residing in district d, $NewSC_{id}$ a dummy indicating whether individual i residing in district d is member of a caste added to the SC list in 1976, $posttreatment_i$ a dummy taking value 1 if individual i is in a treated cohort and X_{idt} a set of control variables.

In my preferred specifications, the control variables will include: region fixed effects interacted with 5 years cohort dummies, as well as caste fixed effects interacted with 5 years cohort dummies, along with dummies for religions, urban residency and a dummie for declaring to be SC.

3 Results

3.1 Main specifications

The difference in differences estimates presented in Figure 1, while informative, could be biased by many correlates. I now turn to the full specification, and main identification strategy of this paper: a within jati variation in the exposure to the SC status. The full specification used in this section relies on within jati-5 years cohort FE and region-5 years cohort FE, along with other standard controls. That is, the estimation is now solely identified on within region-cohort and within jati-cohort variation. As discussed in the previous section, the first cohort to be exposed to the treatment is not well defined a priori. Indeed, treated individuals are the ones that would still be at school at the time of the reform. Since I do not have that information (I observe only the schooling choices after the reform has been implemented, which is endogenous) I resort to a rule of thumb. I assume that the first treatment cohort would be the cohort that reaches the median age of schooling in 1977, that is cohort 1966 for males[19] and cohort 1971 for females[20]. I then repeat the regression twice, first with 1966 as the cut off years and second with 1971.

Tables 3 presents the results of those specifications. It can be seen that the difference in differences results are confirmed and that the SC status leads to an increase of 1 to 2 years of schooling for males, spread across pre secondary schooling levels. For women, access to the SC status leads to a decrease of 1 to 1.4 years of schooling, driven by a decrease in schooling and primary school completion. Note that this is a relative decline: the context is one of overall increase of schooling, so the treated group does not see its absolute level of education decrease. Note also that the treatment effect seems larger for males when the treatment year is 1966 instead of 1971, which points to the fact that male cohorts were treated earlier than 1971, while the contrary is true for females. We will test that more formally in the next subsection.

Hence, it is now clear that the overall null effect of the access to the SC status on

[19] 11 years old in 1977 for a median of 5 years of schooling for the cohorts born up to 1971.
[20] 6 years old in 1977, for a median of 0 years of schooling for the cohorts born up to 1971.

Table 3: Educational attainment. OLS regressions.

	Years of Secondary	**Schooling**	**Primary**	**Literate**	**Some Secondary**	**Secondary**
			Full Sample			
New SC*post 1966	0.517	-0.008	0.069	0.043	0.110***	0.048
	[0.395]	[0.058]	[0.056]	[0.053]	[0.041]	[0.046]
New SC*post 1971	-0.143	-0.041	0.002	-0.009	0.015	-0.017
	[0.452]	[0.064]	[0.054]	[0.063]	[0.038]	[0.061]
N	14609	14609	14609	14609	14609	14609
			Male Sample			
New SC*post 1966	1.918***	0.133*	0.271***	0.217**	0.240***	0.09
	[0.499]	[0.081]	[0.072]	[0.087]	[0.049]	[0.071]
New SC*post 1971	0.982**	0.082	0.159***	0.173***	0.135***	-0.002
	[0.399]	[0.051]	[0.055]	[0.066]	[0.050]	[0.097]
N	7320	7320	7320	7320	7320	7320
			Female Sample			
New SC*post 1966	-0.944***	-0.128*	-0.128**	-0.092*	-0.052*	-0.026
	[0.339]	[0.072]	[0.050]	[0.049]	[0.031]	[0.038]
New SC*post 1971	-1.382***	-0.142*	-0.172***	-0.138**	-0.149***	-0.067
	[0.504]	[0.073]	[0.062]	[0.066]	[0.056]	[0.054]
N	7021	7021	7021	7021	7021	7021

Horizontal lines mark different regressions. *Significant at the 10%, **significant at the 5%, ***significant at the 1%. Standard errors in brackets are two way clustered at the jati and region level. Controls include region * 5 years cohorts FE, jati * 5 years cohorts FE, religion of household head FE, urban FE and declaration to be SC FE (all partialled out). Population aged 18 and above at time of the survey and born after 1950.

educational attainment in fact hides strong asymmetrical effects by gender which cancel each other out in the aggregate. Indeed, while males do benefit from the SC status, it seems to be at the expense of females. It is in particular interesting to see that while the levels of education at which each gender is affected is somehow different, the impact on the length of education is remarkably symmetric, as if the girls were compensating one for one the additional time spent by boys at school, lending credibility to an interpretation in terms of reallocation of resources within the household. Indeed, one could imagine that girls would have to compensate for the opportunity cost of the time spent by their brothers at school by staying at home to help their parents. Indeed, recall from Table 6 that for both gender, an important reason for the lack of schooling was that they were required for household work or to work on the family business. As a consequence, if a child saw its duration of schooling increase, an other child from the household would have to compensate under a given budget constraint and lack of financial market. Since girls' returns to education are often perceived as null, or smaller than that of the boys, from their family perspective, they would be the ones that would naturally compensate. Ideally, one would test this interpretation by comparing households by gender composition of the siblings. Unfortunately, the data does not allow for such a test: as I am using all the individuals of the household roster, I do not have any information of the gender composition of the household head's and his wife's siblings.

3.2 Varying treatment year

We have seen that female's treatment effect seemed to be higher when the treated cohort was assumed to be 1971 instead of 1966, and that the contrary was true for male. This has a logical interpretation: given the education differential between boys and girls, much less women were still at school at age 11 in 1977 compared to males. As a result, the year of treatment should vary across gender. To test more formally this intuition, I will allow the treatment effect to vary both for cohorts born in 1966 and after and for cohorts born in 1971 and after. This will allow me to test if indeed males are treated earlier than women. Table 4 presents the results of such a specification. It can be seen that indeed females are affected by the reform from cohort 1971 onwards, while males are affected from cohort 1966 onward. This reinforces our conclusion that the results are indeed driven by the access to the SC status: if the access to the SC status is to have an impact on school choices, it can only affect the cohorts still at school at the time of its implementation. As a result, cohorts of females born between 1966 and 1971 should be much less affected than their male counterpart, since most of them would already have

left school in 1977, while the contrary is true for males.

Table 4: Educational attainment. OLS regressions, alternative treatment cohorts.

	Years of Secondary	Schooling	Primary	Literate	Some Secondary	Secondary
			Full Sample			
New SC*post 1966	1.172***	0.041	0.129*	0.094	0.186***	0.114**
	[0.334]	[0.070]	[0.066]	[0.088]	[0.056]	[0.052]
New SC*post 1971	-0.890*	-0.067	-0.08	-0.068	-0.104*	-0.09
	[0.502]	[0.080]	[0.066]	[0.098]	[0.056]	[0.081]
N	14609	14609	14609	14609	14609	14609
			Male Sample			
New SC*post 1966	2.110***	0.131	0.276**	0.171	0.250***	0.152**
	[0.637]	[0.132]	[0.111]	[0.168]	[0.063]	[0.077]
New SC*post 1971	-0.289	0.003	-0.007	0.07	-0.015	-0.094
	[0.574]	[0.101]	[0.092]	[0.147]	[0.068]	[0.123]
N	7320	7320	7320	7320	7320	7320
			Female Sample			
New SC*post 1966	0.298	-0.034	0.014	0.035	0.137*	0.056
	[0.393]	[0.076]	[0.062]	[0.080]	[0.077]	[0.066]
New SC*post 1971	-1.581**	-0.12	-0.181**	-0.162	-0.240**	-0.104
	[0.653]	[0.081]	[0.081]	[0.100]	[0.099]	[0.085]
N	7021	7021	7021	7021	7021	7021

Horizontal lines mark different regressions. *Significant at the 10%, **significant at the 5%, ***significant at the 1%. Standard errors two way clustered at the jati and region level in brackets. Controls include region * 5 years cohorts FE, jati * 5 years cohorts FE, religion of household head FE, urban FE and declaration to be SC FE (all partialled out). Population aged 18 and above at time of the survey and born after 1950.

3.3 Gender and schooling

Before turning to robustness checks, I will first propose a hypothesis so as to why such differences are observed across gender. Women in India are generally at a disadvantage in terms of schooling. For example, in the sample, women have on average 2.8 years of schooling as opposed to 5.8 for their males counterpart. One of the main reasons usually put forward to explain this difference are the lower returns to schooling (Kingdon, 1998), or lower perceived returns to schooling (Dreze and Sen, 2002) for females. Indeed, as is evident in Table 5, women's participation in the labour force is much smaller than males', with 63% of women "not classified by occupations"[21] as opposed to 14% of males. As a

[21] Which essentially means being a housewife.

result, women's returns to education are often perceived as quasi nonexistent. Moreover, as marriage practices are generally patrilocal, the perception of the investment in the daughter's skills is often seen as a waste from the point of view of the parents. Dreze and Sen (2002) write for example: "[...] the gender division of labour [...] tends to reduce the perceived benefits of female education. [...] It is in the light of these social expectations about the adult life of women that female education appears to many parents to be of somewhat uncertain value, if not quite 'pointless' ."[22]. Note that this tendency is in no way specific to the lower castes in India. Actually, lower castes tend to be less gender biased than higher castes (Srinivas (1966), Chakravarti (1993), Mencher (1988), Kapadia (1997), Field et al. (2010), Cassan and Van de Walle (2015)).

Table 5: Occupation by gender among SC, in percent.

Occupation	**Female**	**Male**
Professional, technical and related work	1.22	2.65
Administrative, executive & managerial	0.03	0.63
Clerical and releated workers	0.56	2.91
Sales workers	1	5.14
Service workers	2.5	4.32
Farmers, fishermen, hunters, loggers &	22.09	32.72
Production & related workers, transport	9.53	37.66
Workers not classified by occupations	63.23	13.97
Total	100	100
N	7,280	7,316

The NFHS2 survey contains a question on the reason why a person did not attend school. Table 6 presents the responses to this question by gender. The main reasons for males are first and foremost the cost, and then that the child is required for household work or work on farm or family business. For females, the picture is strikingly different: the main reasons for which they do not attend school is because it is not considered necessary, closely followed by the fact that they are required for household work, and that schooling costs too much. Hence both cost and opportunity costs appear to be a major reason for which children are not sent to school, with the additional issue that often, education is not considered necessary for girls.

In this context of lower perceived returns from education for girls from the point of

[22] Dreze and Sen (2002) also underline a second important reason for the low enrolment of women: the low quality of schooling infrastructures. Indeed while, say, the absence of a nearby school could be thought to be gender neutral, in practice, girls are more affected, as parents tend to be more reluctant to send their daughters to far away schools than their sons. Finally, an other reason put forward by Dreze and Sen (2002) for the lower education of girls is that even if education can be an asset on the marriage market, it can only be so if the girl's education remains lower than that of her potential husband.

Table 6: Reason not going to school by gender.

main reason never went to school	**Female**	**Male**
school too far away	3.38	2.44
transport not available	0.91	0.76
education not considered necessary	23.69	11.94
required for household work	20	12.76
required for work on farm/family business	3.71	12.47
required for outside work for payment	2.6	8.03
cost too much	19.84	25.56
no proper school facilities for girls	3.45	
required for care of siblings	2.49	0.72
not interested in studies	6.85	12.47
other	5	4.3
don't know	7.65	8.55
Total	100	100
N	4 496	2 093

view of the parental household, an increase in overall returns to education and decrease in costs of education might have asymmetrical effects by gender. Indeed, education of the children is costly, both in terms of direct cost and opportunity costs (since children are often required to work on household chores or family business). Hence, a gender-symmetric increase in the returns of schooling might lead to a reallocation of resources within the household in order to allow the boys to stay longer at school. A perception by the parents that returns might be gender-asymmetric either because they do not believe that returns to schooling exist for girls, or because they benefit less from the returns to schooling of girls than that of boys is sufficient a mechanism to drive those type of behaviour. Under imperfect access to credit and budgetary constraints, a household would have to chose to sacrifice the education of a child for which returns to schooling are perceived as lower in order to favor the education of an other one. See for example Banerjee and Duflo (2011) for a discussion of this type of reallocation of schooling investment across children. Hence, a plausible interpretation of the asymmetric results by gender of the access to the affirmative action program are gender norms.

4 Robustness check

4.1 Differential trend

A concern with any difference in difference estimation would be that different trends across treatment and control groups would lead to biased results. We have seen in Figure 1 that it did not seem to be the case, while our within caste-cohort and within region-

cohort estimations already wash out almost all differential trends concerns. However, this subsection deals with concern more formally by adding a "new" SC specific trend to the specification. As can be seen in Table 7, the results are essentially not affected, even if slightly less significant, thus underlining that differential trends are not behind the results.

Table 7: Educational attainment. OLS regressions, "New SC" specific trend.

	Years of Secondary	**Schooling**	**Primary**	**Literate**	**Some Secondary**	**Secondary**
			Male Sample			
New SC*post 1966	2.802**	0.353**	0.306*	0.310*	0.23	0.143*
	[1.175]	[0.172]	[0.185]	[0.175]	[0.149]	[0.087]
New SC*post 1971	0.148	0.171	0.005	0.174	-0.043	-0.111
	[0.773]	[0.108]	[0.097]	[0.175]	[0.099]	[0.148]
New SC*post 1966	2.815**	0.362**	0.307*	0.319*	0.228	0.138*
	[1.196]	[0.175]	[0.185]	[0.174]	[0.152]	[0.083]
New SC*post 1971	0.256	0.185*	0.017	0.186	-0.034	-0.106
	[0.793]	[0.096]	[0.094]	[0.168]	[0.102]	[0.151]
N	7320	7320	7320	7320	7320	7320
			Female Sample			
New SC*post 1966	0.339	0.008	-0.011	-0.017	0.127*	0.007
	[0.568]	[0.101]	[0.093]	[0.109]	[0.065]	[0.062]
New SC*post 1971	-1.233*	-0.068	-0.173**	-0.179	-0.187*	-0.123
	[0.694]	[0.078]	[0.087]	[0.110]	[0.103]	[0.077]
New SC*post 1966	0.608	0.022	0.025	0.02	0.170**	0.033
	[0.590]	[0.103]	[0.094]	[0.113]	[0.075]	[0.073]
New SC*post 1971	-1.338*	-0.072	-0.177**	-0.182	-0.216**	-0.129
	[0.685]	[0.077]	[0.086]	[0.112]	[0.109]	[0.084]
N	7289	7289	7289	7289	7289	7289

Horizontal lines mark different regressions. *Significant at the 10%, **significant at the 5%, ***significant at the 1%. Standard errors two way clustered at the jati and region level in brackets. Controls include a linear trend and its interaction with the "New SC" dummy, region * 5 years cohorts FE, jati * 5 years cohorts FE, religion of household head FE, urban FE and declaration to be SC FE (all partialled out). Population aged 18 and above at time of the survey and born after 1950.

4.2 Selective migration

One concern with the estimation is selective migration. Indeed, I attribute the "New SC" status based on the district of residence of respondents at the time of the survey. It is possible that before 1976, certain households would have migrated in order to benefit from the SC status. As a consequence, the results of the estimations might be biased as

those migrants would wrongly be coded as "Old" SC. However, migration is relatively low in India (Munshi and Rosenzweig, 2009) and particularly so in the 1970's and before. Indeed, from the Migration Volumes of the 1981 Census, one can estimate that only 6.1% of the males and 11.1% of females had changed district before 1977[23]. In addition, the households choosing to migrate would probably be the ones that would have benefited the most from the access to the SC status had they not migrated. This means that, if anything, this selective migration is likely to bias the estimates for males downwards and for females upwards, as I would consider those households as "Old SC". Finally, given the difference in differences specification and the differential effect by gender, only a very specific and unlikely form of migration could bias the results. Indeed, one would need the migrants to have at least one of the two following characteristics: 1. that its male members born before 1966-1971 would have a higher education than the "Old SC" of the same age; 2. that its male members born after 1966-1971 would have a lower education than the "Old SC" of the same age. And that these characteristics should be the opposite for the females. Moreover, note that the first cohorts to be treated are born before 1976, that is, before the law is passed. This means that for selective migration to bias the results, the selection into migration should change at the time when the first cohorts are treated, that is up to ten years before the law is passed. Thinking about a mechanism that would lead to such change in behaviour prior to the enactment of the law requires more imagination that I unfortunately have at my disposition.

As the NFHS data does not offer useful migration information, it is not possible to directly verify if the migrating household indeed are different in any form from the other households. Despite the weak probability that migration could be an issue for my estimation strategy, I will show in this subsection that the results hold even with very strong assumptions on the profiles of migrants.

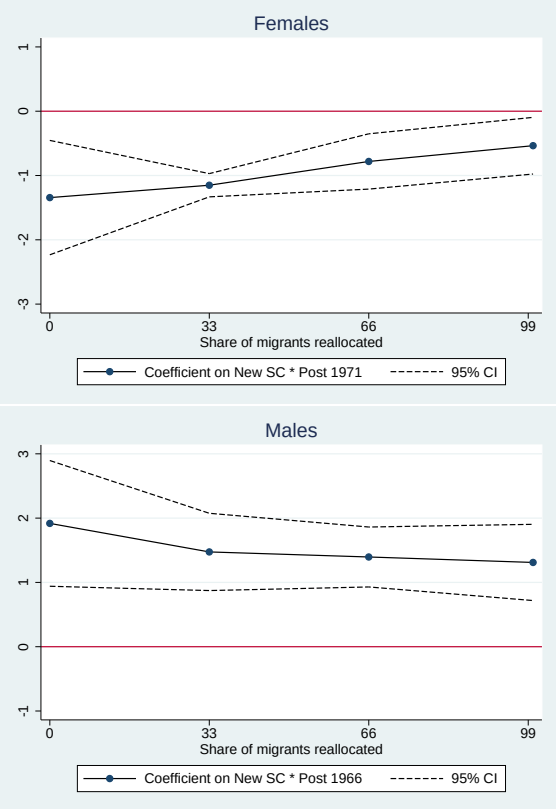
The robustness check proposed in this section proposes to reallocate the individuals who contribute the most to the results from control to treatment. I identify the individuals that are the most likely to contribute to bias the results if they had been migrants. That is, for male, I identify "old SC" born prior the first treatment cohort with the highest level of education, and the born after the first treatment cohort with the lowest level of education (and symmetrically for female). Among those identified individuals, I then randomly draw a number corresponding to 33%, 66% and 100% of the total migrating population[24], reallocate them to the "New SC" status and bootstrap 1000 times the main

[23]Note that this is an overestimation of the type of migration that concerns us, since the bias might emerge if those migrant moved in a way that would change their SC status, which is often not the case when crossing a district border.

[24]That is migrants that would account for 6.1% of the total male "New SC" and 11.1% of the total

estimations at each of those steps. Figure 2 plots the coefficients of those estimations. It can be seen that while moving towards zero, even if all the potential migrants, identified as those whose level of education would contribute the most to a potential bias, were reallocated, the results would not be qualitatively affected. As a result, migration can not drive the results of this paper.

Figure 2: Reallocation of potential migrants and robustness of the coefficients.



5 Conclusion

This paper studies the impact of the access to affirmative action policies in India. It shows that the access to the Scheduled Caste status has an ambiguous effect on schooling attainment. Indeed, the overall null effect actually hides large differences across genders, with male benefiting largely from those policies while female to the contrary seem to suffer. This suggest that affirmative action policies might need to be complemented by

female "New SC".

other policies in order to prevent the weakest population of the targeted group from being excluded or to lose from the program. One reason that can be put forward to explain somehow counterintuitive results would be the difference in perceived returns to schooling by gender combined with imperfect credit markets: as returns to schooling are perceived to increase more for boys than for girls, girls might be prevented from going to school in order to compensate for the opportunity cost of the boy's increased education.

References

Ashraf, Nava, Nathalie Bau, Nathan Nunn, and Alessandra Voena, "Bride Price and Female Education," *Working Paper*, 2015.

Banerjee, Abhijit and Esther Duflo, *Poor Economics: A Radical Rethinking of the Way to Flight Global Poverty*, Public Affairs, 2011.

Bayly, Susan, *Caste, Society and Politics in India*, Cambridge University Press, 1999.

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.

Cassan, Guilhem and Lore Van de Walle, "Overlapping Identities and Identity Based Policies: Evidence from Gender Quotas in India," *Working Paper*, 2015.

Chakravarti, Uma, "Conceptualising Brahmanical Patriarchy in Early India: Gender, Caste, Class and State," *Economic and Political Weekly*, 1993, *28* (14).

Chalam, K.S., "Caste Reservations and Equality of Opportunity in Education," *Economic and Political Weekly*, 1990, *25* (41).

Chitnis, Suma, "Education for Equality: Case of Scheduled Castes in Higher Education," *Economic and Political Weekly*, 1972, *7* (31-33).

Dreze, Jean and Amartya Sen, *India. Development and Participation.*, 2 ed., Oxford University Press, 2002.

Duflo, Esther, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 2001, *4* (91), 795–813.

Field, Erica, Rohini Pande, and Seema Jayachandran, "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India," *American Economic Review Papers and Proceedings*, 2010.

Foster, Andrew and Mark Rosenzweig, "Missing Women, the Marriage Market and Economic Growth," *Working Paper*, 1999.

_ **and** _ , "Technological change and the distribution of schooling: evidence from green-revolution India," *Journal of Development Economics*, 2003, *74* (1), 87–111.

Galanter, Marc, *Competing Equalities: Law and the Backward Classes in India*, University of California Press, 1984.

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

_ , ed., *Report of the Commissioner for Scheduled Castes and Scheduled Tribes, 1958-1959.*, Manager Govt. Of India Press, 1960.

_ , ed., *Report of the Commissioner for Scheduled Castes and Scheduled Tribes, 1975-1977.*, Manager Govt. Of India Press, 1978.

Hnatkovska, Viktoria, Amartya Lahiri, and Paul Sourabh B., "Castes and Labor Mobility," *American Economic Journal: Applied Economics*, 2012, *4* (2).

_ , _ , and _ , "Breaking the Caste Barrier: Intergenerational Mobility in India," *Journal of Human Resources*, 2013, *48* (2).

Jensen, Robert, "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *Quarterly Journal of Economics*, 2012, *2* (127), 153–192.

Kapadia, Karin, "Mediating the Meaning of Market Opportunities: Gender, Caste and Class in Rural South India," *Economic and Political Weekly*, 1997, *35.*

Khanna, Gaurav, "That's Affirmative: Incentivizing Standards or Standardizing Incentives? Affirmative Action in India," *Working Paper*, 2013.

Kingdon, Geetah, "Does the Labour Market Explain Lower Female Schooling in India ?," *Journal of Development Studies*, 1998, *35* (1).

Krishna, Kala and Veronica Frisancho Robles, "Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch," *Working Paper*, 2012.

Mencher, Joan P, "Women's Work and Poverty: Women's Contribution to Household Maintenance in South India," in Daisy Dwyer and Judith Bruce, eds., *Home Divided*, Stanford University Press, 1988.

Munshi, Kaivan and Mark Rosenzweig, "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," *Working Paper*, 2009.

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

Qian, Nancy, "Missing Women and the Price of Tea in China: the Effect of Sex-Specific Earnings on Sex Imbalance," *Quarterly Journal of Economics*, 2008, *123* (3).

Rosenzweig, Mark and T. Paul Schultz, "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India," *American Economic Review*, 1982, *72* (4), 803–815.

Srinivas, *Caste in Modern India and Other Essays*, University of California Press, 1966.

Appendices

A Areas of restriction

The following list presents state by state the different "regions" for which separate SC list were drawn until 1976. This information has been extracted from the 1971 Census' Social and Cultural Tables, publicly available, which also contain the whole SC lists by region. It can be seen that certain "regions" are included in larger ones. In that case, in the empirical exercises of the paper, I took the most conservative approach, that is: each region is given a specific fixed effect, while for the computation of the clustered standard errors, those regions are included in larger ones. In addition, in certain cases, only taluks are listed, not entire districts. In that case, I took the most conservative coding approach, which is to code the whole district as part of the region.

Andhra Pradesh

Region A: throughout the state; Region B: districts of Srikakulam, Visakhapathnam, East Godavari, West Godavari, Krishna, Gunter, Ongole, Nellore, Chittoor, Cuddapah, Anantapur and Kurnool; Region C: districts of Mahbubnagar, Hyderabad, Medak, Nizamabad, Adilabad, Karimnagar, Warangal, Khamman and Nalgonda.

Assam

No regional caste list. This state is not used in the paper.

Bihar

Region A: throughout the state; Region B: Patna and Tirhut divisions and Monghyr, Bhagalpur, Palamay and Purnea districts; Region C: districts of Patna, Shahabad, Gaya and Palamau.

Gujarat

Region A: throughout the state except Rajkot division and Kutch district; Region B: Dangs districts and Umbergaon taluka of Surat district; Region C: Rajkot division; Region D: Kutch district.

Haryana

Region A: throughout the state; Region B: throughout the state except in Mahendragarh and Jind districts; Region C: districts of Mahendragarh and Jind.

Himachal Pradesh

As Himachal Pradesh is carved out of Punjab in 1966, it is not used in the paper. Region A: districts of Chamba (excluding the towns of Dalhousie MC, Dalhousie CB and Bakloh CB), Mandi, Bilaspur, Mahasu, Sirmaur and Kinnaur; Region B: districts of Kangra, Kulu, Lahul & Spiti, Simla and the towns of Dalhousie MC, Dalhousie CB and Bakloh CB in Chamba district.

Jammu & Kashmir

No regional caste list. This state is not used in the paper.

Kerala

Region A: Throughout the state; Region B: throughout the state except in Kasaragod taluka of Malabar district; Region C: throughout the state except Malabar district (excluding Kasaragod taluka); Region D: throughout the state except Malabar district; Region E: Malabar district; Region F: Malabar district (excluding Kasaragod taluka); Region G: Kasaragod taluka of Malabar district.

Madhya Pradesh

Region A: districts of Bind, Gird, Morena, Shivpuri, Goona, Rajgarh, Shajapur, Ujjain, Ratlam, Mandsaur, Bhilsa, Indore, Dewas, Dhar, Jhabua and Nimar; Region B: districts of Chhindwara, Seoni, Betul, Jabalpur, Sagar, Damoh, Mandla, Hoshangabad, Narsimharpur, Nimar, Balaghat, Raipur, Bilaspur, Durg, Rastar, Surguja and Raigarh; Region C: Bilaspur district; Region D: Sagar and Damoh districts; Region E: Damoh district; Region F: districts of Bilaspur, Durg, Raipur, Bastar, Surguja and Raigarh; Region G: Sagar district; Region H: Balaghat district; Region I: districts of Balaghat, Bilaspur, Durg, Raipur, Surguja, Bastar and Raigarh; Region J: districts of Balaghat, Betul, Bilaspur, Durg, Nimar, Raipur, Bastar, Surguja, Chhindwara, Sagar and Raigarh, Hoshangabad and Seoni Malwa thesils of Hoshangabad district; Region K: districts of Sagar and Damoh, Hoshangabad and Seoni Malwa thesils of Hoshangabad district; Region L: districts of Chhindwara, Seoni, Betul, Jabalpur, Narsimhapur, Sagar, Damoh, Mandla, Nimar, Balaghat, Raipur, Durg, Bastar, Surguja, Raigarh and Hoshangabad (except Harda and Sohagpur thesils); Region M: districts of Chhindwara, Seoni, Betul, Jabalpur, Narsimhapur, Sagar, Damoh, Mandla, Nimar, Balaghat, Raipur, Bilaspur,

Durg, Bastar, Surguja, Raigarh and Hoshangabad (except Harda and Sohagpur thesils); Region N: Sohagpur thesil of Hoshangabad district; Region O: districts of Datia, Tikamgarh, Chhatarpur, Panna, Stana, Rewa, Sidhi and Shahdol; Region P: districts of Raisen and Sehore.

Maharashtra

Region A: throughout the state except Buldana, Akola, Amravati, Yeotmal, Wardha, Nagpur, Bhandara, Chanda, Aurangabad, Parbhani, Nanded, Bhir, Osmanabad and Rajura districts; Region B: districts of Greater Bombay, Dhulia, Jalgaon, Nasik, Ahmednagar, Poona, Satara, Sangli, Kolhapur, Sholapur, Thana, Kolaba and Ratnagiri; Region C: districts of Buldana, Akola, Amravati, Yeotmal, Wardha, Nagpur, Bhandara and Chanda; Region D: districts of Akola, Amravati and Buldana; Region E: Bhandara district; Region F: districts of Bhandara and Buldana; Region G: districts of Amravati, Bhandara and Buldana; Region H: districts of Aurangabad, Parbhani, Nanded, Rajura, Bhir and Osmanabad.

Manipur

No regional caste list. This state is not used in the paper.

Meghalaya

No regional caste list. This state is not used in the paper.

Karnataka

Region A: throughout the state except Coorg, Belgaum, Bijapur, Dharwar, Kanara, South Kanara, Gulbarga, Raichur and Bidar districts, and Kollegal talukla of Mysore district; Region B: districts of Belgaum, Bijapur, Dharwar and Kanara; Region D: Kanara district; Region E: districts of Gulbarga, Bidar and Raichar; Region F: district of South Kanara and Kollegal taluka of Mysore district; Region G: Kollegal taluka of Mysore district; Region H: district of South Kanara; Region I: Coorg district.

Orissa

Region A: throughout the state; Region B: Sambalpur district.

Punjab

Even though there were sub regions, those were not removed in 1976, so Punjab is not used in this paper. Region A: throughout the state; Region B: throughout the state except Patiala, Bhatinda, Kapurthala and Sangrur districts; Region C: districts of Patiala, Bhatinda, Kapurthala and Sangrur districts.

Rajasthan

Region A: throughout the state except Ajmer district, Abu Road taluka of Sirohi district and Sunel Tappa taluka of Jhalawar district; Region B: Ajmer district; Region C: Abu Road taluka of Sirohi district; Region D: Sunel Tappa taluka of Jhalawar district.

Tamil Nadu

Region A: throughout the state; Region B: throughout the state except Kanyakumari district and Shencottah taluka of Tirunelveli district; Region C: Nilgiri district; Region D: districts of Coimbatore and Salem; Region E: Kanyakumari district and Shencottah taluka of Tirunelveli district; Region F: Tanjore district.

Tripura

No regional caste list. This state is not used in the paper.

Uttar Pradesh

Region A: throughout the state; Region B: throughout the state excluding Agra, Meerut and Rohilkhand divisions; Region C: Bundelkhand division and the portion of Mirzarpur district south of Kaimur range.

West Bengal

Region A: throughout the state;

Region B: throughout the state except in Purulia district and the territories transferred from Purnea district of Bihar; Region C: in Purulia district and the territories transferred from Purnea district of Bihar; Region D: the territories transferred from Purnea district of Bihar.

Arunachal Pradesh

No regional caste list. This state is not used in the paper.

Chandigarh

No regional caste list. This state is not used in the paper.

Dadra and Nagar Haveli

No regional caste list. This state is not used in the paper.

Delhi

No regional caste list. This state is not used in the paper.

Goa, Daman and Diu

No regional caste list. This state is not used in the paper.

Mizoram

No regional caste list. This state is not used in the paper.

Pondicherry

No regional caste list. This state is not used in the paper.