Journal of Development Economics 136 (2019) 51–70

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Affirmative action, education and gender: Evidence from India☆

Guilhem Cassan

University of Namur, CRED, Belgium

ARTICLE INFO

JEL classification:
I24
O15
H41

Keywords:
Identity
Scheduled caste
Quota
Affirmative action
Gender
India
Education
Intersectionality

ABSTRACT

This paper studies the impact of India's affirmative action policies for Scheduled Castes on educational attainment. Using a plausibly exogenous variation, I show that affirmative action increases educational attainment. The main improvements are in literacy and secondary schooling and there is only small evidence of increases in higher education. The benefits are not distributed evenly across genders: only males show an increase in education (in literacy, primary and secondary completion). Individuals at the intersection of discriminated groups (low caste and female) may not be benefiting from these policies.

1. Introduction

The very nature of affirmative action policies makes their evaluation difficult, so, while these policies are often widely debated, the debate is rarely well informed. India is a particularly interesting case study. It has implemented the largest affirmative action program in the world, targeting the low castes and, in particular, the "Scheduled Castes" (SC).[1] Those policies have been controversial since their introduction at Independence. In the early 90s, with the expansion of affirmative action policies to a new group of castes, namely the "Other Backward Classes" (OBC), the debate has become particularly intense. It has focused on quotas in higher education, in which competition with the high castes was the strongest. Following this trend, most of the literature has dealt with the consequences of quotas in universities (Bertrand et al., 2010; Krishna and Frisancho Robles, 2012) or of affirmative action policies for OBC (Khanna, 2013). Some studies have looked into the evolution of education of the SC as a group (Desai and Kulkarni, 2008; Hnatkovska et al., 2012), but without an identification strategy that can associate trends in educational attainment with affirmative action policies. As a consequence there is still little knowledge about what is arguably the first order question for the evaluation of affirmative action policies' impact on education: that of their average effect on the SC population, a deficit already diagnosed several decades ago (Chalam, 1990; Chitnis, 1972; Galanter, 1984). According to the 2011 Census, only 56% of the SC population aged 20 and above was literate (whereas 66% of the population was), only 30% had attained an education level higher

☆ I am grateful to Gani Aldashev, Denis Cogneau, Christelle Dumas, Esther Duflo, Hemanshu Kumar, Eliana La Ferrara, Sylvie Lambert, Andreas Madestam, Marco Manacorda, Talia Brun, Annamaria Milazzo, Kaivan Munshi, Barbara Petrongolo, Rohini Somanathan, Vincenzo Verardi and Yanos Zylberberg, as well as participants in the seminar and conferences in which I presented this paper. 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. I thank Rohini Pande for having shared her data. I thank Fred Arnold at DHS for having helped me understand the coding of jati names in the DHS data. 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 and from MIS-FNRS 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 gs4: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.

E-mail address: guilhem.cassan@unamur.be.

[1] The "Scheduled Castes" are the castes traditionally the most discriminated against. They represent approximately 17% of the population in 2011.

https://doi.org/10.1016/j.jdeveco.2018.10.001
Received 7 February 2017; Received in revised form 26 September 2018; Accepted 3 October 2018
Available online 5 October 2018

than primary (versus 41%), and only 6% had gone beyond secondary schooling (versus 11%). Hence, the current focus on the effect of quotas in universities for higher education students tends to neglect the weakest segment of the Indian population, the majority of SC which never reach university. This tendency to focus on the impact of affirmative action in higher education institutions on the marginal beneficiaries more than on the entire target population is present in other affirmative action literature in settings such as the United States of America (Arcidiacono, 2005; Hinrichs, 2012) or Brazil (Francis and Tannuri-Pianteo, 2012). In the Indian context, this issue is particularly important, as affirmative action policies have other dimensions than just quotas in higher education.

This paper makes two new contributions. First, it measures the average effect of affirmative action policies at every level of education, instead of the marginal effect at higher levels of education. I find an overall important but imprecisely estimated effect of affirmative action, which may have led to an increase in literacy of 10 percentage points and in secondary attainment of 7 percentage points. This highlights that the focus on higher education quotas had led us to overlook the important and beneficial impact of affirmative action policies at lower levels of education. However, although quotas in higher education probably change the level of education of individuals benefiting from them, they do not appear to directly affect the average level of education of SC in a meaningful way: I find no statistically significant increase in higher education.[2] But if quotas in higher education do not seem to affect strongly higher education attainment, they may still have an indirect impact on education: by providing role models or by increasing educational attainment at lower levels, via an incentive to pursue longer studies.[3] Note that these positive findings on the effects of affirmative action for SC contrast with the research on the impact of quotas in legislative assemblies on SC. Indeed, Pande (2003), Chin and Prakash (2011) and Jensenius (2015) find very little effect of electoral quotas on policies targeted to SC, poverty reduction and public goods provision. As a consequence, this paper allows to nuance our view on the impact of affirmative action policies on SC.

Second, by emphasizing the gender differential in the impact of affirmative action, this paper underlines the heterogeneous effect of affirmative action policies within the treated population. The overall increase in education is entirely captured by men. This finding relates to the literature on intersectionality (Crenshaw, 1989, 1991) by showing that individuals at the intersection of different discriminated groups (such as women of low castes) may not be sufficiently protected by policies that ignore this cumulative discrimination. More generally, it relates to the literature on gender discrimination in education (Jensen, 2012) and in particular to the literature on the asymmetric effect of social policies in developing countries (Foster and Rosenzweig, 2003; Rosenzweig and Schultz, 1982; Ashraf et al., 2015).

The research design of this paper relies on a natural experiment which creates variation within jati - the subcaste that is the relevant reference group in daily life in India - and within states in access to SC status. At Independence, a list of jatis to be considered SC was drawn for each state. There were some discrepancies across states: the same jati could be considered SC in one state, but not in a neighboring state.[4] In 1956, the borders of the Indian states were redrawn along linguistic lines, while leaving the lists of SC unchanged. This aggravated the discrepancies: a jati could have been considered SC in one part of the state but not in an other part of the same state. This situation lasted until 1976, when the lists of SC were harmonized within each state, giving 2.4 million individuals (Government of India, ed, 1978) access to the SC status. This historical variation creates a natural experiment for assessing the impact of affirmative action policies. For all jatis that were considered SC after 1976, I distinguish two groups, a control group of "early SC" (ESC) who were considered SC at Independence, and the treated group, "late SC" (LSC), that became SC after 1976. To avoid any risk of confounding due to congestion effects,[5] I will use two additional control groups: the members of jatis considered SC in a state different from their own and the OBC. The former, members of the jatis considered SC somewhere else (SCSE) are likely to be exposed to the same discrimination but never benefit from affirmative action. The latter are composed of jatis also discriminated against by high castes, who were subsequently targets of affirmative action in the 1990s. Both represent a useful counterfactual: they are not exposed to affirmative action policies during the period of interest, but are arguably not too dissimilar from SC jatis. Figs. 1 and 2 summarize the variation used in the paper. They show the evolution of the SC status of three jatis, J, K and L, spread across two states whose borders will change in 1956. Jati J becomes SC at Independence in state A only, and jati K in state B only, while jati L will get the OBC status in the 1990s. As the borders of states A and B change in 1956, the SC status of members of jatis J and K now becomes different within the same state B' depending on their region of residence. In 1976, with the removal of area restriction, the SC status of jatis J and K becomes consistent within state (but not across). Ideally one would also want to use the variation generated by the creation of the SC status in 1950. In 1950, the ESC get access to the SC status while the LSC do not: the LSC thus provide an ideal counterfactual. However, given the long delay between the data collection and the 1950 natural experiment, I can not fully exploit this variation and only use it as suggestive evidence. Table 1 summarizes the control and treatment status of each jati, based on the notation in Figs. 1 and 2.

As the treatment status varies across jatis, and within jati across time and space, this paper innovates by coding the treatment status of individuals based on their jati name, instead of their declared beneficiary status.[6] By using the jati name I can use demanding specifications, which identify the effects of the policy within a single jati, something that, to the best of my knowledge, has never been done in the literature.

The data used in the paper are the second round of the DHS (also know as NFHS), collected in 1998–9. To my knowledge, this is the only dataset containing at the same time a large enough sample size, the precise jati name and district identifiers. A major drawback of this dataset is the lack of information on migration: I know where the respondents reside at the time of the survey, but have no information on their residence at the time of the policy change. Hence, selective migration may be an issue for the results. I specifically address this concern, along with concerns of selective identity manipulation as in Cassan (2015) to show that they can not drive the results.

In the first section of this paper, I present the context and the natural experiment. I then describe the data and the empirical strategy, leading to the results, discussion and various robustness checks (exploitation of the first access to the SC status, differential threshold for treatment, migration and identity manipulation). Finally, I explore the heterogeneity of the findings to suggest some possible channels through which access to SC status may lead to increased education.

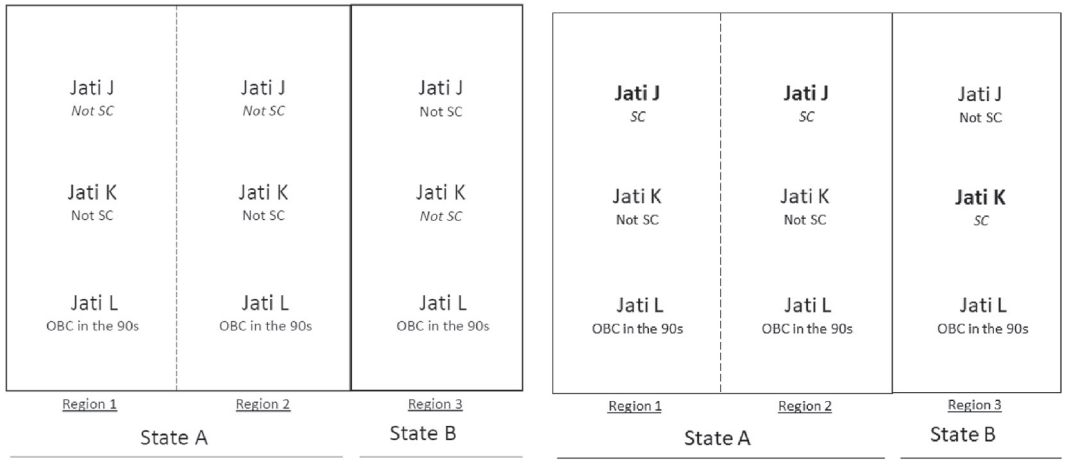
[2] Note however that quotas in higher education should, by construction, increase the average level of education among SCs if the individuals entering higher education with quotas would not have entered higher education otherwise. The fact that I do not observe this pattern in the data only means that this effect is too small to be captured, as very few SC reach higher education.

[3] On that point, see for example Khanna (2013).

[4] There were also within state discrepancies, but they were relatively minor.

[5] If the resources allocated to SC do not change in 1976, but the number of SC increases, some of the convergence between ESC and LSC may be due to the decrease in resources allocated to ESC. This would still be proof that access to SC status mattered for educational attainment, but the interpretation of the results would be different.

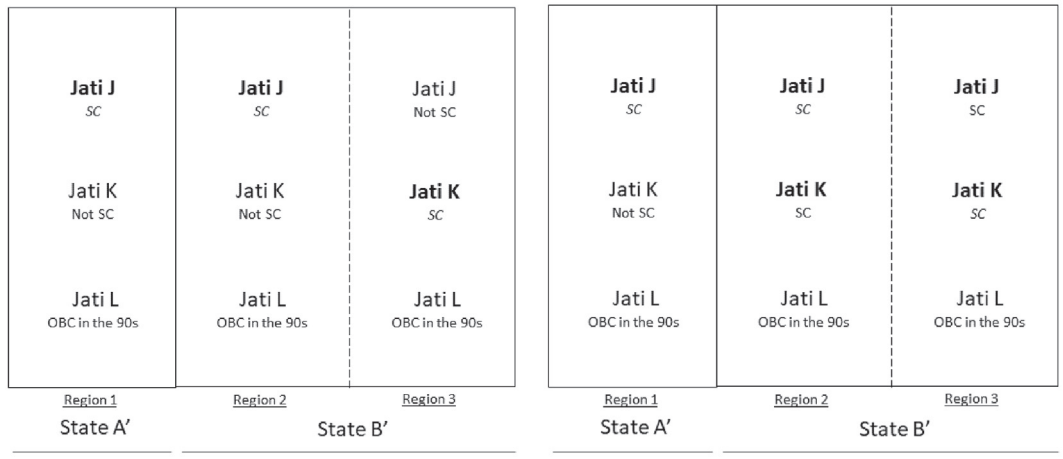
[6] Indeed, most of the micro level literature on affirmative action in India uses the household's declaration of its SC status as a basis for identifying the "treatment" group (Khanna, 2013; Hnatkovska et al., 2012, 2013; Prakash, 2009).



(a) SC Status before Independence: No SC status

(b) SC Status at Independence: SC status not harmonized across states

Fig. 1. Treatment status before and after 1950.



(a) SC Status in 1956: SC status not harmonized within state

(b) SC Status in 1976: SC status harmonized within but not across states

Fig. 2. Treatment status before and after 1976.

2. Context

2.1. Affirmative action in India

There are three categories of castes that receive affirmative action in India: the Scheduled Castes, the Scheduled Tribes and the Other Backward Classes. While the first affirmative action policies for SC and ST were implemented under British rule, no systematic policy was implemented until after Independence. Affirmative action, or "reservations", has three main dimensions: political representation, education and public employment. There are electoral quotas for SC in all federal and state assemblies, as well as in panchayat (municipal) elections (since the mid 90s for the latter). A share of new public employment positions is reserved for SC. Finally there are various affirmative action policies in education: quotas in higher education institutions, free secondary schooling, scholarships, specific schools, hostels and free mid day meals.[7] Affirmative action 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 facilitate post-secondary education. The quotas in public employment increase 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. Which dimension of reservation policies is more likely to contribute to the results is discussed in Section 6 via the analysis of heterogeneous effects.

For OBCs, while the Indian constitutions explicitly mentions that they should also be protected, it is not until the 1990s that they obtained affirmative action benefits. For the period studied in this paper, they are not systematically targeted by such policies.

2.2. The definition of the Scheduled Castes

How are the beneficiaries of these affirmative action policies chosen? As underlined by Galanter (1984), one of the main difficulties in creating a list of Scheduled Castes is the criteria of inclusion. While

[7] See Galanter (1984), chapter 3.B for a presentation of the main educational policies targeting low castes apart from quotas in higher education.

Table 1
Treatment and control status of Jatis.

	Region 1	Region 2	Region 3
	First Experiment: Independence		
Jati J	Treatment	Treatment	Control
Jati K	Control	Control	Treatment
Jati L	Control	Control	Control
	Second Experiment: Area Restriction Removal		
Jati J	Control	Control	Treatment
Jati K	Control	Treatment	Control
Jati L	Control	Control	Control

"untouchability" is the prerequisite, defining "untouchability" is not straightforward. The Constitution avoids any definition and only provides a procedure of designation.[8] The SC list enacted in 1950 reproduced the British colonial authorities' list from 1936, which "*reflected definitions of untouchability with an admixture of economic and educational tests and considerations of local politics*" (Galanter, 1984). As the SC are defined at the state level, this allowed for inconsistencies across and within states,[9] as certain jatis would be considered SC only in certain areas within a state. Despite those inconsistencies, the lists were revised only four times following Independence. The first change, in 1951 was to correct small anomalies. The second change in 1956 mainly allowed all Sikh untouchable jatis to claim SC status, while the fourth change of 1990 gave Buddhists access to SC status in all states. The most important change, the third, took place in 1976 and is described in more details in the following subsection. 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.

2.3. The Scheduled Castes and Scheduled Tribes Orders (Amendment) Act of 1976

In 1956, India reorganized its states' borders along linguistic lines[10], but the state SC lists remained unchanged. This led to a large increase of within state discrepancies in the SC lists. Hence, from 1956, the number of jatis considered as SC in one part of a state and not in an other part of the same state vastly increased until the SC lists were harmonized within states by the SC and ST (Amendment) Act of 1976.[11] This act was also known as the "Area Restriction Removal Act"[12] and it removed almost all intra-State restrictions, increasing the SC population by 2.4 million.[13] I will refer to areas of restriction as "regions".

Individuals from the same jati in the same linguistic area would have seen their SC status differ: some would get access to SC status at Independence, while others would only get access after the 1976 Area Removal Restriction Act. Was this difference only de jure, while de facto, all the members of the jati would have been granted the benefits of reservations? We will explore empirically this question in subsection 5.1 and in Appendix D, but a reading of the Reports of the Commissioner for SC and ST helps clarifies that there was no such thing as a de facto SC status. For example, in its 24th report, the Commissioner for SC and ST writes: "*The removal of area restrictions in the respect of Scheduled Castes and Scheduled Tribes will no doubt enable the members of these communities, who were deprived so far of the benefits and concessions given by the Central and State Governments to get their due share of educational, economic and political safeguards.*" (Government of India, ed, 1978). The following section will detail how this variation in exposure to the SC status is empirically exploited.

3. Data and descriptive statistics

The National Family and Health Survey of 1998–99 (NFHS2) is, to my knowledge, the only dataset offering both the jati name of respondents, their district of residence and a sufficient sample size to perform the type of analysis done in this paper.[14]

The jati names in the NFHS2 are given only in raw form, so I tried to clean the names of the jatis in a systematic manner. I contracted a data entry team based in India to match the raw names from the NFHS2 with the list of jatis and their synonyms from Kitts (1885), and a second list of jatis and their synonyms. Both of these lists had been used for the coding of the first round of the NFHS.[15] I manually refined this matching using the state wise list of SC and OBC synonyms taken from states' lists of SC and OBC, and corrected remaining errors. I identified a total of 287 jatis, among which 137 SC jatis (see Online Appendix 1 and 2 for the list of jatis present in the data and their coding as ESC, LSC, OBC or SCSE). Using the cleaned jati names, the 1971 and 1981 Census lists of SC and the district of residence of households, I can identify the households that were granted SC status in 1976.[16] Note that the attribution of the treatment status implicitly assumes that there was no migration across regions and that there was no identity manipulation. As a result, miscoding of the treatment status may be an issue for this analysis. In subsection 5.3, I explicitly rule out that this could drive the results.

[8] "*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.*"

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

[10] In 1960, the states of Maharashtra and Gujarat were created from the former state of Bombay.

[11] According to Galanter (1984) the number of intrastate restrictions dropped from 1126 to 64.

[12] 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).

[13] 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. 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.

[14] For example, the 2004–5 round of the NFHS does not contain information on the district of residence of the respondents, while the IHDS data, if it contains both jati names and district identifiers, only has half the sample size of the NFHS2.

[15] The first round of the NFHS (1992–3) also offers jati and district identifiers. This project initially planned to use both first and second NFHS rounds. However, the first round turned out to be unusable as the documentation on the codes of jatis (which is based on Kitts (1885)) does not correspond to the codes present in the data. A different list of jati codes had been used for the first round of NFHS, overlapping Kitts (1885) but adding many other jati codes, and which I managed to acquire. However, this second list appeared to be unknown to the central organization of NFHS and also to not correspond perfectly to the codes present in the raw data either. As a consequence, the first round of the NFHS could not be used.

[16] I match the 1999 districts of the NFHS2 data to their 1971 counterparts using the Indian Administrative Atlas of 2001, which follows district changes over time.

Table 2
Descriptive Statistics: Male population.

	Late SC		Early SC		OBC and SCSE	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education variables						
Education level	2.019	1.473	2.131	1.446	2.425	1.363
Literacy	0.660	0.474	0.689	0.463	0.773	0.419
Primary Completion	0.687	0.464	0.719	0.449	0.799	0.401
Secondary Completion	0.502	0.500	0.536	0.499	0.616	0.486
Higher Education Completion	0.141	0.348	0.153	0.360	0.209	0.407
Control variables						
Hindu	0.990	0.101	0.896	0.305	0.981	0.137
Sikh	0.003	0.058	0.062	0.241	0.019	0.137
Buddhist/Neo Buddhist	0.007	0.082	0.042	0.200	0.000	0.009
N	882		9400		13,670	
Number of jatis	49		107		184	

Sample: Male population aged 18 and above at the time of the survey and born from 1950 onwards. Data source: NFHS2.

Table 3
Descriptive Statistics: Female population.

	Late SC		Early SC		OBC and SCSE	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education variables						
Education level	1.001	1.424	1.117	1.427	1.360	1.504
Literacy	0.328	0.470	0.369	0.483	0.445	0.497
Primary Completion	0.350	0.477	0.403	0.490	0.476	0.499
Secondary Completion	0.231	0.422	0.247	0.431	0.318	0.466
Higher Education Completion	0.067	0.250	0.062	0.241	0.088	0.283
Control variables						
Hindu	0.988	0.107	0.894	0.307	0.982	0.131
Sikh	0.002	0.048	0.060	0.237	0.018	0.131
Buddhist/Neo Buddhist	0.009	0.096	0.046	0.209	0.000	0.009
N	864		9262		13,644	
Number of jatis	54		107		185	

Sample: Female population aged 18 and above at the time of the survey and born from 1950 onwards. Data source: NFHS2.

The summary statistics for the variables used in the paper are in Tables 2 and 3. All variables are self explanatory, except "education level", which I created based on the level of education declared,[17] and which provides a useful summary variable. I restrict the sample to individuals born from 1950 onwards and aged at least 18 at the time of the survey.[18] Since Muslims, Christians, Atheists and Jains do not have access to SC status, they are removed from the sample. Note that other religions such as Sikhs and Buddhists were not granted access to the SC status systematically: the Sikhs could claim the SC status only from 1956, and the Buddhists only from 1956 in Maharashtra and from 1990 in the rest of India. Since they are a small fraction of the population and have access to the SC status for most of the period of interest,[19] I keep them in the sample, but results are robust to keeping only Hindus.

4. Identification strategy and results

4.1. Identification strategy

As the SC are a very specific population, much poorer and more discriminated against than the rest of the population, they do not have an obvious counterfactual. In particular, as many pro-poor policies were put in place in the mid 70s, comparing the SC to the entire Indian population would probably confound the effect of the affirmative action policies with other social policies. The natural experiments at hand alleviate this issue by allowing the comparison of the differential evolution of educational attainment as a function of the timing of access to the SC status for individuals within jati and within states.[20] To my knowledge, this study is the first to go beyond the comparison of trends between SC and other social groups in India to assess the impact of affirmative action on the average beneficiary (Desai and Kulkarni, 2008; Hnatkovska et al., 2012).

To use the notations of Figs. 1 and 2 and summarized in Table 1, in 1976, Jati K gains access to the SC status in region 2 of state B′ while

[17] 0 if no education, 1 if primary is completed, 2 if secondary is completed and 3 if higher education is completed. Note that I did not consider literacy in that education level variable since very few individuals declare to be literate and to not have completed primary education (144 individuals in a sample of 47,801). The opposite case of individuals declaring to be illiterate while having completed primary education is more frequent, therefore, studying literacy separately from primary school completion remains informative.

[18] Except when analyzing the impact of the access to the SC status in 1950, in which case I restrict to individuals aged 65 and above at the time of the survey.

[19] There are almost no Buddhists in SC jatis outside of Maharashtra.

[20] Prior to 1956 the borders of the states were different. Hence, while it is correct to present the analysis as within state, one may worry that parts of states coming from different political entities prior to 1956 may follow a different trajectories. My preferred specifications include regions∗cohort fixed effects, so that all regional trends are flexibly controlled for.

Jati J gain access to it in region 3 of state B′. This setting allows me to measure the impact of the access to the SC status within state B′: Jati J (resp. K) in region 2 (resp. 3) is the control group and Jati J (resp. K) in region 3 (resp. 2) is the treated group. Hence, because there is within-jati variation across time and space as well as within-region variation across jatis, I can control for jati FE, jati∗cohort FE, region FE, region∗cohort FE as well as region∗jati FE.

With this setting I can compare within the same jati the fate of those who had access to the SC status early against those who had access to it later. This calls for a difference in differences specification. The first difference would compare cohorts too old to benefit from the access to the SC status in 1976 to those that were young enough. The second difference would compare treated segments of jatis to untreated segments of jatis. This approach has the advantage of relying solely on within jati within state variation.

However, a concern with comparing LSC to ESC is congestion effects. If the budget for affirmative action policies did not increase in 1976, per capita spending would have decreased. The control group would be contaminated, and (negatively) treated. This is not a worry for the identification strategy but it would change the interpretation of the coefficients. In addition, it is useful to understand what is driving the results: is the convergence observed between male ESC and LSC solely due to an increase in LSC's education, or to a combination of an increase for LSC and a decrease for ESC's? To answer that question, I use triple differences strategies. Indeed, the double differences just described do not take advantage of all the variation at hand. In particular, they do not use jati L nor do they use jati K in Region 1, that is OBCs and members of jatis residing in a state in which their jati is not considered SC while it is in an other. Jati K in Region 1 would provide a very useful counterfactual: it is the same jati as jati K in state B′, but never obtains access to the SC status. Jati L is also interesting to look into: while it is a different jati, we know that it is a relatively worse off jati and subject to discrimination, as it will gain access to affirmative action policies in the 1990s.

4.2. Fully flexible specifications

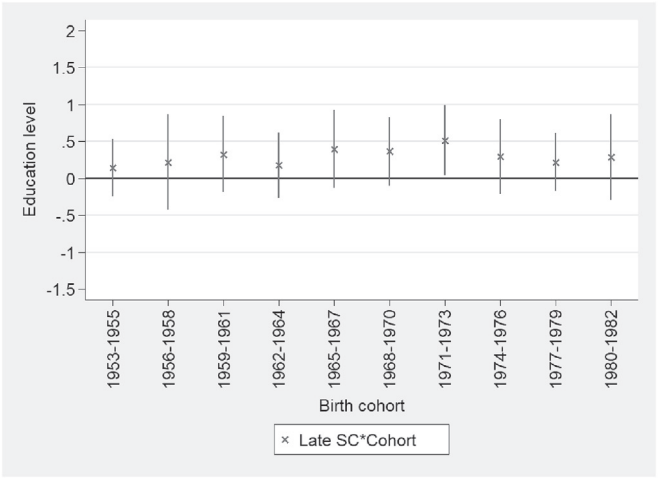
In that setting, the standard approach is to run fully flexible specifications as in Duflo (2001), where the effect of the treatment is allowed to vary by year. Given the sample size, I can not allow the coefficient to vary by year, but will allow it to vary by 3 years cohorts. I run the following regression for the double differences setting:

$$Edu_{ijdt} = \sum_{t=1951}^{1981} \delta_t * Late\_SC_{jd} * cohort_t + \sum_{t=1951}^{1981} \gamma_t * cohort_t + \lambda X_{dt} + \omega Y_{jt} + \iota Z_{dj} + \eta W_{ijdt} + constant + \epsilon_{ijdt} \quad (1)$$

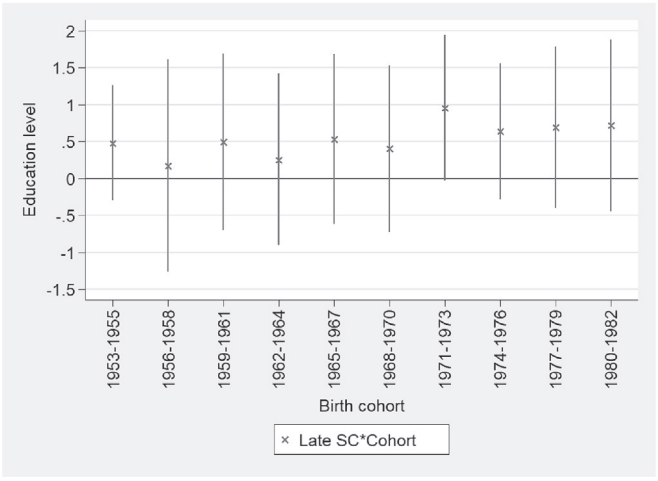
and, for the triple differences setting:

$$Edu_{ijdt} = \sum_{t=1951}^{1981} \beta_t * SC_{jd} * cohort_t + \sum_{t=1951}^{1981} \delta_t * Late\_SC_{jd} * cohort_t + \sum_{t=1951}^{1981} \gamma_t * cohort_t + \lambda X_{dt} + \omega Y_{jt} + \iota Z_{dj} + \eta W_{ijdt} + constant + \epsilon_{ijdt} \quad (2)$$

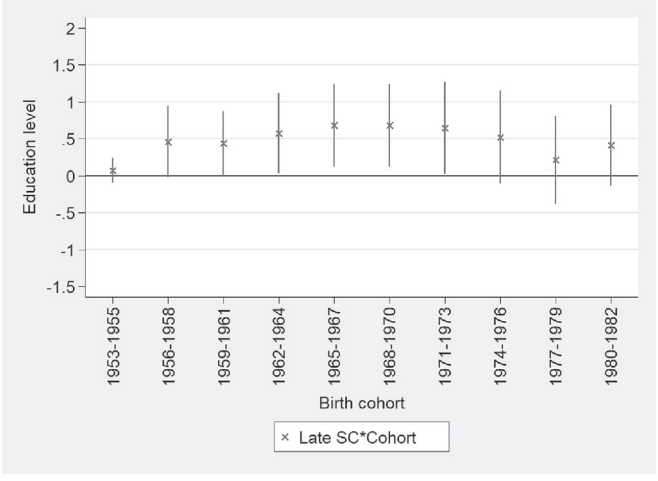
Where Edu_{ijdt} is a measure of educational attainment of individual i of jati j born in year t and residing in region d, $Late\_SC_{jd}$ a dummy indicating whether jati j in region d will be added to the SC list in 1976, SC_{jd} is a dummy indicating whether jati j in region d will be added to the SC list either at Independence or in 1976, $cohort_t$ a set of 3 years cohort of birth dummies and X_{dt} is a set of region∗5-years cohorts fixed effects, Y_{jt}, jati∗5-years cohorts fixed effects, Z_{dj} region∗jati fixed effect and W_{ijdt} a set of fixed effects: religion, region, birth year, jati and gender. Figs. 3 and 4 presents the results of such specifications



Both genders

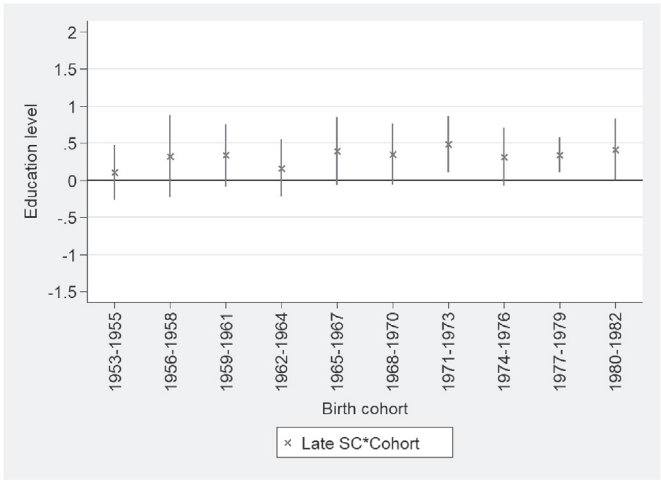


Males

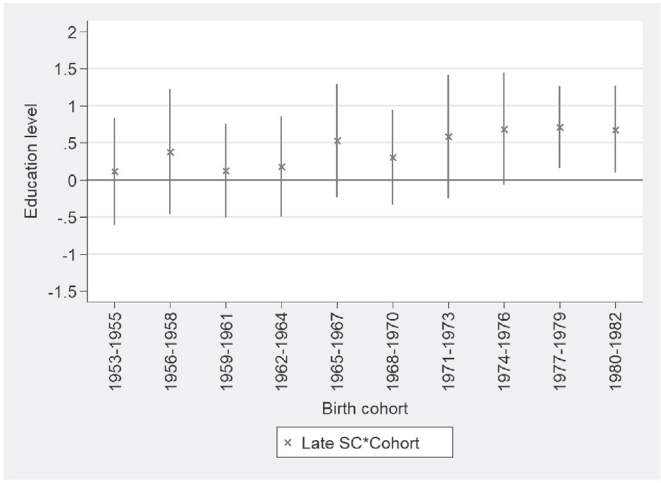


Females

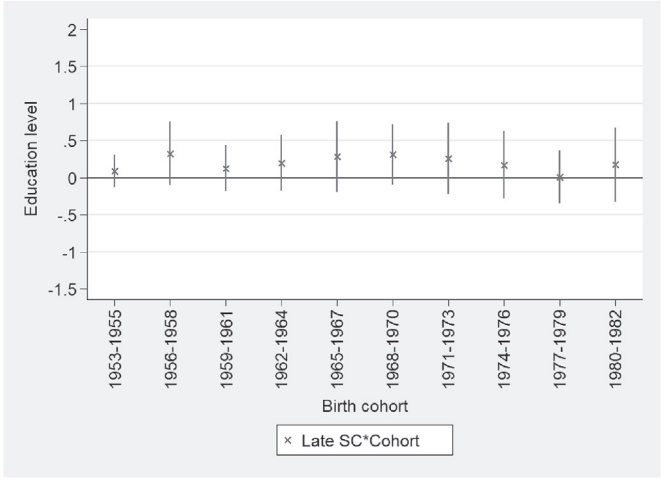
Fig. 3. 1976 natural experiment: cohort specific effects, double differences estimates.
Data source: Round 2 of the Indian DHS. Sample: SC population born from 1950 onwards and aged 18 and above at time of survey. Cohorts born from 1965 onwards are younger than 12 at the time of the reform, and are likely to have benefited from it. The control cohorts are 1950–1952.



Both genders



Males



Females

Fig. 4. 1976 natural experiment: cohort specific effects, triple differences estimates.
Data source: Round 2 of the Indian DHS. Sample: SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey. Cohorts born from 1965 onwards are younger than 12 at the time of the reform, and are likely to have benefited from it. The control cohorts are 1950–1952.

by gender, for the double and triple differences approach respectively, presenting each δ_t coefficient and their 95% confidence interval.

Both the double differences and the triple differences approaches point to the same result: there may be an increase in education for LSC upon accessing SC status from cohorts born from 1965 onwards (that is, aged 12 or below at the time of implementation of the reform), and in particular, from cohorts born from 1971 onwards (that is, aged 6 or below at the time of the reform). However, this aggregate effect hides heterogeneity across genders: for males, there is an increase in education for the cohorts born from 1965 onwards. While for cohorts born from 1953 to 1967, no trend is visible, and the level is centered around zero, there is an increase of roughly 0.5 education levels for cohorts 1965 to 1970 and of roughly 0.7 education levels for cohorts born from 1971. For females there does not seem to have been an important impact. It is also clear from the visual inspection of the graphs that differential trends across treatment groups do not seem to be present. Finally, the graphs of the β_t coefficients from the triple difference estimation can be found in Appendix A They show that ESCs do not seem to be affected by the arrival of LSCs. This underlines that congestion effects are probably not an issue here. It is obvious from the graphs that almost no coefficient is statistically significant: given the small sample size and the very demanding specifications, it is no surprise that slicing the treatment effect across 3-years cohorts would not lead to precisely estimated coefficients. The δ_t coefficients for the triple differences estimates of the regressions using literacy, primary, secondary and higher education as outcomes are presented in Appendix B.

4.3. Double and triple differences estimations

To test for statistical significance, I have to impose more structure on the data, and implement blunter double and triple differences specifications. In practice, this is similar to running a test of joint significance of the individual coefficients of the fully flexible approach. I will thus run double differences regressions of the type:

$$Edu_{ijdt} = \delta_1 Late\_SC_{jd} + \delta_2 Late\_SC_{jd} * post\_1965_t + constant + \gamma_1 post\_1965_t + \lambda_1 X_d + \lambda_2 W_{ijdt} + \epsilon_{ijdt} \quad (3)$$

Where $post\_1965_t$ is a dummy taking value 1 if the cohort t is young enough to be affected by the treatment. I consider as young enough the individuals who were below secondary school age at the time of the reform (that is, born from 1965 onwards). X_d is a set of region fixed effects and W_{ijdt} a set of individual controls (jati fixed effect, birth year fixed effect, religion, gender). In Section 5.2, I show that the results are robust to considering an alternative threshold, that of individuals below primary school age at the time of the reform (born from 1971 onwards). Finally, triple differences specifications of the following type can also be used:

$$Edu_{ijdt} = \beta_1 SC_{jd} + \beta_2 SC_{jd} * post\_1965_t + \delta_1 Late\_SC_{jd} + \delta_2 Late\_SC_{jd} * post\_1965_t + constant + \gamma_1 post\_1965_t + \lambda_1 X_d + \lambda_2 W_{ijdt} + \epsilon_{ijdt} \quad (4)$$

I run two specifications of these regressions: one that includes only jati FE, region FE and birth year FE, and one that adds jati∗5 years cohort FE, region∗jati FE and region∗5 years cohort FE. That is, relatively "naive" double and triple differences estimations, before turning to estimations that effectively control for confounding trends in a non parametric manner. The latter is my preferred specification. Since the treatment is attributed at the region∗jati level, the variable $Late\_SC_{jd}$ is fully collinear with the set of region∗jati FE. However, the coefficient on that variable has an interesting interpretation: it is the average difference in educational attainment between ESC and LSC before the LSC had access to the SC status but after the ESC benefited from it. That is, this coefficient measures the impact of the access to the SC status on

Table 4
Double difference estimates, both genders.

	Education level	Literacy	Primary	Secondary	Higher
	A: Without jati-region, region-cohort and jati-cohort FE				
Late SC	−0.012	−0.038	−0.025	−0.004	0.018
	(0.091)	(0.045)	(0.040)	(0.034)	(0.028)
	[−0.16, 0.15]	[−0.11, 0.04]	[−0.10, 0.05]	[−0.06, 0.05]	[−0.02, 0.06]
Late SC∗Post 1965	−0.001	0.038	0.012	0.011	−0.024
	(0.070)	(0.032)	(0.030)	(0.027)	(0.027)
	[−0.14, 0.14]	[−0.02, 0.10]	[−0.06, 0.08]	[−0.04, 0.06]	[−0.07, 0.02]
N	20,438	20,438	20,438	20,438	20,438
	B: With jati-region, region-cohort and jati-cohort FE				
Late SC∗Post 1965	0.142	0.098	0.056	0.070	0.016
	(0.081)	(0.027)	(0.039)	(0.031)	(0.046)
	[−0.02, 0.30]	[0.04, 0.16]	[−0.03, 0.13]	[0.01, 0.13]	[−0.05, 0.08]
N	20,247	20,247	20,247	20,247	20,247

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE, gender FE and religion FE. Panel B also includes jati∗5 years cohort FE, region∗5 years, and jati-region FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: SC population born from 1950 onwards and aged 18 and above at time of survey.

Table 5
Double difference estimates, males.

	Education level	Literacy	Primary	Secondary	Higher
	A: Without jati-region, region-cohort and jati-cohort FE				
Late SC	−0.098	−0.097	−0.069	−0.051	0.021
	(0.088)	(0.055)	(0.039)	(0.039)	(0.034)
	[−0.27, 0.08]	[−0.19, −0.00]	[−0.15, 0.01]	[−0.12, 0.02]	[−0.04, 0.08]
Late SC∗Post 1965	0.073	0.119	0.074	0.053	−0.053
	(0.086)	(0.045)	(0.032)	(0.037)	(0.043)
	[−0.10, 0.23]	[0.02, 0.21]	[−0.01, 0.14]	[−0.02, 0.12]	[−0.12, 0.01]
N	10,276	10,276	10,276	10,276	10,276
	B: With jati-region, region-cohort and jati-cohort FE				
Late SC∗Post 1965	0.306	0.178	0.125	0.137	0.044
	(0.115)	(0.041)	(0.057)	(0.045)	(0.061)
	[0.08, 0.56]	[0.07, 0.29]	[−0.01, 0.26]	[0.05, 0.23]	[−0.05, 0.14]
N	10,041	10,041	10,041	10,041	10,041

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE and religion FE. Panel B also includes jati∗5 years cohort FE, region∗5 years, and jati-region FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: male SC population born from 1950 onwards and aged 18 and above at time of survey.

ESC. Obviously, as this is a simple difference, this interpretation relies on the assumption that the level of education of both ESC and LSC was similar prior to Independence. This issue will be explored further in subsection 5.1. Note that since the treatment is attributed at the jati and region level, it is necessary to two way cluster the standard errors along those two dimensions. There are only 38 non overlapping regions. There is no consensus on what the minimum number of clusters should be, but 38 may be too small. Hence, I provide the confidence intervals generated by the cgmwildboot (Caskey, 2015) ado file, bootstrapping on the regions: the statistical significance of the results remain almost entirely unaffected.

Finally, all regressions are weighted using the survey weights, but results are very similar when regressions are not weighted.[21] Tables 4–6 present the results of the double differences estimates for the whole sample and then by gender. The visual inspection of Fig. 4 is confirmed: having SC status increases overall educational attainment, but this is entirely driven by men. Men seem to gain in literacy, primary and secondary but not higher education. These gains are important: they represent around 20% of the mean. Women do not seem to benefit at all. In fact, the coefficient is sometimes even negative. Note how the inclusion of the two way FE increases the coefficients for both men and women. This underlines how omitted variables may be an issue and bias the results downwards should jati and region specific trends as well as jati-region fixed effect not be controlled for.

Note also that the coefficient on Late SC is negative for men. That is, ESC males prior to 1976 had a higher level of education than that of LSC (but this is not precisely estimated). This is what one would expect if SC status increases education: ESC have SC status prior to 1976 while LSC do not. See subsection 5.1 for a detailed analysis of the effect of the access to the SC status on the ESC.

The triple differences estimates presented in Tables 7 and 8 confirm the results of the double differences. Note how the inclusion of two way fixed effect effectively nullifies the coefficient on SC∗Post 1965. This indicates the importance of these two way fixed effects to control for confounding trends which would have led us to underestimate the effect of the access to SC status. This small and non significant coefficient on SC∗Post 1965 allows me to rule out congestion effects as a driver of the findings. In addition, the comparison of the double and

[21] Unweighted results can be computed by the reader using the replication data and code.

Table 6
Double difference estimates, females.

	Education level	Literacy	Primary	Secondary	Higher
	A: Without jati-region, region-cohort and jati-cohort FE				
Late SC	0.091	0.029	0.024	0.048	0.018
	(0.101)	(0.044)	(0.044)	(0.039)	(0.027)
	[−0.07, 0.24]	[−0.04, 0.09]	[−0.05, 0.10]	[−0.01, 0.12]	[−0.02, 0.06]
Late SC∗Post 1965	−0.102	−0.054	−0.061	−0.038	−0.004
	(0.059)	(0.028)	(0.031)	(0.025)	(0.016)
	[−0.23, 0.02]	[−0.11, 0.00]	[−0.13, 0.01]	[−0.09, 0.01]	[−0.04, 0.04]
N	10,121	10,121	10,121	10,121	10,121
	B: With jati-region, region-cohort and jati-cohort FE				
Late SC∗Post 1965	0.109	0.047	0.032	0.041	0.037
	(0.120)	(0.044)	(0.064)	(0.052)	(0.046)
	[−0.10, 0.34]	[−0.04, 0.13]	[−0.09, 0.15]	[−0.06, 0.14]	[−0.04, 0.11]
N	9859	9859	9859	9859	9859

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE and religion FE. Panel B also includes jati∗5 years cohort FE, region∗5 years, and jati-region FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: female SC population born from 1950 onwards and aged 18 and above at time of survey.

Table 7
Triple differences estimates, males.

	Education level	Literacy	Primary	Secondary	Higher
	A: Without jati-region, region-cohort and jati-cohort FE				
SC	−0.184	−0.090	−0.081	−0.069	−0.034
	(0.032)	(0.018)	(0.022)	(0.015)	(0.017)
	[−0.29, −0.08]	[−0.14, −0.04]	[−0.13, −0.26]	[−0.12, −0.02]	[−0.07, 0.00]
Late SC	−0.093	−0.097	−0.073	−0.048	0.028
	(0.093)	(0.063)	(0.044)	(0.037)	(0.033)
	[−0.28, 0.09]	[−0.21, 0.01]	[−0.17, 0.01]	[−0.12, 0.02]	[−0.03, 0.08]
SC∗Post 1965	0.077	0.041	0.056	0.032	−0.011
	(0.033)	(0.011)	(0.011)	(0.017)	(0.017)
	[0.02, 0.14]	[0.02, 0.06]	[0.04, 0.07]	[−0.00, 0.07]	[−0.03, 0.01]
Late SC∗Post 1965	0.082	0.125	0.081	0.052	−0.051
	(0.090)	(0.047)	(0.034)	(0.036)	(0.044)
	[−0.09, 0.25]	[0.03, 0.22]	[−0.00, 0.17]	[−0.02, 0.12]	[−0.12, 0.02]
N	23,952	23,952	23,952	23,952	23,952
	B: With jati-region, region-cohort and jati-cohort FE				
SC∗Post 1965	0.034	−0.002	0.015	0.015	0.004
	(0.081)	(0.032)	(0.035)	(0.035)	(0.028)
	[−0.10, 0?17]	[−0.07, 0.06]	[−0.05, 0.08]	[−0.06, 0.09]	[−0.04, 0.05]
Late SC∗Post 1965	0.369	0.168	0.146	0.168	0.055
	(0.076)	(0.042)	(0.033)	(0.040)	(0.034)
	[0.22, 0.53]	[0.10, 0.24]	[0.07, 0.22]	[0.10, 0.24]	[−0.02, 0.13]
N	23,394	23,394	23,394	23,394	23,394

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE and religion FE. Panel B also includes jati∗5 years cohort FE, region∗5 years, and jati-region FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: male SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey.

triple differences estimates show that the double differences strategy may have led to an underestimation of the effect for males. Indeed, the increase in primary and secondary education is slightly higher with this estimation strategy. Finally, Appendix C shows that the coefficient on Late SC∗Post 1965 for male is statistically significantly different from that for females: there is a different effect of access to SC status by gender. In the interest of space, the rest of the paper will only present triple differences estimates by gender[22] with the full set of fixed effects.

Why such difference across genders? Women in India are generally at a disadvantage in terms of schooling. 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. As a result, women's returns to education are often perceived as nonexistent. Moreover, as marriage practices are generally patrilocal, parents' often see investment in daughters' education as unnecessary. 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

[22] Double differences estimates give similar results and can be computed by the reader from the replication files.

Table 8
Triple differences estimates, females.

	Education level	Literacy	Primary	Secondary	Higher
	A: Without jati-region, region-cohort and jati-cohort FE				
SC	−0.046	−0.017	−0.034	−0.006	−0.006
	(0.080)	(0.033)	(0.038)	(0.034)	(0.011)
	[−0.17, 0.08]	[−0.08, 0.04]	[−0.09, 0.02]	[−0.07, 0.06]	[−0.03, 0.01]
Late SC	0.106	0.043	0.036	0.048	0.021
	(0.089)	(0.043)	(0.042)	(0.037)	(0.019)
	[−0.04, 0.25]	[−0.03, 0.12]	[−0.04, 0.12]	[−0.01, 0.11]	[−0.01, 0.05]
SC∗Post 1965	−0.018	−0.008	0.01	−0.008	−0.02
	(0.035)	(0.013)	(0.012)	(0.019)	(0.012)
	[−0.08, 0.04]	[−0.04, 0.02]	[−0.02, 0.04]	[−0.04, 0.02]	[−0.04, 0.00]
Late SC∗Post 1965	−0.104	−0.057	−0.065	−0.036	−0.003
	(0.060)	(0.029)	(0.032)	(0.026)	(0.016)
	[−0.23, 0.02]	[−0.11, 0.00]	[−0.14, 0.01]	[−0.09, 0.02]	[−0.05, 0.03]
N	23,770	23,770	23,770	23,770	23,770
	B: With jati-region, region-cohort and jati-cohort FE				
SC∗Post 1965	−0.04	−0.049	−0.024	−0.005	−0.011
	(0.071)	(0.040)	(0.041)	(0.025)	(0.016)
	[−0.15, 0.07]	[−0.11, 0.02]	[−0.09, 0.04]	[−0.05, 0.03]	[−0.04, 0.16]
Late SC∗Post 1965	0.041	0.004	0.002	0.004	0.035
	(0.095)	(0.038)	(0.056)	(0.040)	(0.029)
	[−0.14, 0.24]	[−0.08, 0.08]	[−0.11, 0.11]	[−0.08, 0.09]	[−0.02, 0.09]
N	23,222	23,222	23,222	23,222	23,222

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE and religion FE. Panel B also includes jati∗5 years cohort FE, region∗5 years, and jati-region FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: female SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey.

parents to be of somewhat uncertain value, if not quite 'pointless'."[23,24]

In this context, a gender neutral increase in returns to education and decrease in costs of education might have different effects by gender. Indeed, child education is costly, in terms of direct cost and opportunity costs. As a result, parents' perception that returns might be gender-asymmetric (because they think they are non-existent for girls, or at least less valuable to them than the return for boys) might prevent girls from benefiting from affirmative action.

5. Robustness check

5.1. 1950 as a natural experiment

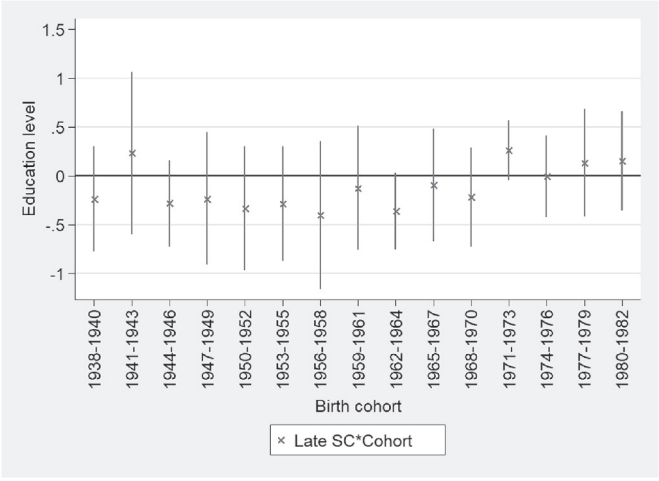
As discussed in the introduction, there are two natural experiments: 1976 and 1950, when SC status was attributed for the first time. This paper can not fully exploit the natural experiment of 1950 for two reasons. First, as the data was collected in 1998–9, there are few individuals old enough in 1950 to not be affected by access to SC status, so all the regressions using these cohorts are imprecisely estimated. They may also be subject to bias due to differential survival rates across groups. Second, as the state borders changed in 1956, it can not be said that the analysis is exploiting only within state variation for the cohorts born prior to 1950. Nevertheless, bearing in mind the limits of the exercise, it is informative to exploit the full variation, by analysing the data on an extended sample, with individuals aged between 18 and 65 at the time of the survey, that is, with individuals born in 1933–34 and after. This larger sample allows the exploitation of both natural experiments: that of 1950 and that of 1976. For the 1950 experiment, the LSC are the control group (since their reservation status does not change at the time), and if SC status leads to an increase in education, it is expected that they should lose relative to the ESC (who gain access the SC status). For the 1976 experiment, the LSC are the treated group, and they should gain relative to ESC. Note that in 1950, following Independence, many policies change apart from the systematization of the SC status. In particular, education policies changed dramatically, affecting the level of education of the entire population. As a consequence, it is hard to have a prior about the relative evolution of ESC compared to OBC and SCSE. The only prediction is that if access to SC status indeed affects education, then LSC should lose relative to ESC upon the creation of SC status. Fig. 5 presents the results of a non parametric specification equivalent to that of subsection 4.2 on that sample. LSC males seem to lose relative to the ESC for the cohorts born between 1944 and 1964 compared to cohorts born earlier. For females, no clear picture emerges apart from an absence of effect: access to SC status does not seem to affect their level of education. The results of triple differences regression exploiting this variation can be viewed in Appendix D. They confirm the visual analysis: there is an increase in the education gap between ESC and LSC males (but not females) for the cohorts born from 1950 onwards, but the coefficients are less precisely estimated for the 1950 experiment. Hence, it appears that the increase in education within jatis is linked to the timing of access to SC status: within a jati, those that got access to SC status in 1950 saw their education increase compared to those that got access later. When the LSC finally get SC status, this gap is filled. This pattern in the data is a strong confirmation that the observed increase in education observed is due to gaining SC status.

5.2. Alternative age thresholds

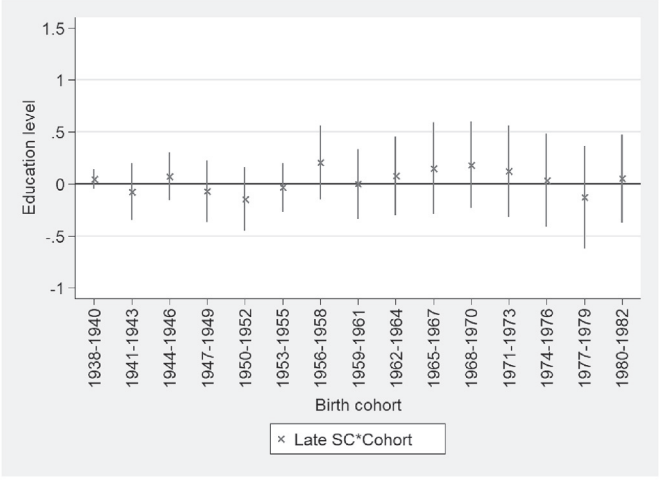
I have so far assumed that SC status could only be beneficial to individuals aged 12 or below at the time of the implementation of the

[23] Dreze and Sen (2002) also underline a second important reason for the low enrolment of women: the low quality of schooling infrastructures: 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.

[24] 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, 2017).



Males



Females

Fig. 5. 1950 and 1976 natural experiment: cohort specific effects, triple differences.
Data source: Round 2 of the Indian DHS. Sample: SC, OBC and SCSE population aged 18 to 65 at the time of the surveys. Cohorts born from 1965 onwards are younger than 12 at the time of the reform, and are likely to have benefited from it. Cohorts born from 1938 onwards are younger than 12 at the time of creation of SC status and are likely to have lagged behind ESC. The control cohorts are 1933–1937.

Area Restriction Removal Act. However, there is ambiguity about the specific age at which the treatment will begin: many children could already be out of school at age 12 and not benefit from SC status. As a robustness check, I present here the results of regressions in which the age threshold is fixed at 6 years old at the time of the reform (born from 1971 onwards) instead of 12. That is, I allow SC status to affect children below school age at the time of the reform. These children are likely to benefit from the full effect of the reform, since they could not yet have dropped out from school. The results are in Table 9. The results remain qualitatively unaffected by the change in the threshold year, even if the coefficients are larger with 1971 as a threshold.

As previously discussed, children aged 6 and below at the time of the reform are likely to benefit more from SC status than children aged 7 to 12, since all of them would still be at school at the time of the reform. It is therefore not surprising that the estimated effect is larger in this specification than in the specification using 1965 as a threshold.

5.3. Selective migration and identity manipulation

There are two main threats to the identification strategy. The first is migration. Due to the limitations of the data, I attribute the LSC and ESC status based on the district of residence at the time of the survey. Therefore, I do not know if there has been selective migration: it is easy to imagine that before 1976, there was an incentive to migrate to regions in which one would have SC status. However, it is well documented that migration is relatively low in India (Munshi and Rosenzweig, 2009) and particularly so before the 1970s. 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.[25]

The second threat to identification would be identity manipulation (Cassan, 2015). Indeed, I attribute the treatment and control status based on the declared jati name. Since SC status effectively depends on the jati, people may have had an incentive to manipulate their jati identity to obtain SC status before 1976. To my knowledge, there are no estimate of the extent of this type of identity manipulation.[26] In the context of colonial Punjab, Cassan (2015) shows that a jati based policy with arguably stronger impact on day to day life led to an identity manipulation movement of up to 7.5% of the population.

Both mechanisms are empirically very similar: they result in coding as control individuals that should have been in the treatment group. It is hard to think of a scenario in which those mechanisms would bias the results upwards: households choosing to migrate/manipulate would probably be the ones that would have benefited the most from SC status had they not migrated/manipulated. This means that, if anything, migration/identity manipulation is likely to bias the estimates downwards, as I would code those households as "ESC". In addition, it is extremely difficult to think about a miscoding of the treated individuals that would lead to the patterns observed in Fig. 5: similar level of education for cohorts born prior to 1943, lower for cohorts born between 1943 and 1966 and similar again after. As the NFHS2 data does not offer useful migration information,[27] it is not possible to directly verify if the migrating/manipulating households are different in any way from the other households.

The exercise that I propose is to reallocate from control to treatment the observations that contribute the most to the results. That is, for male, I identify non LSC individuals born prior to the treatment cohort with an above median education level, and the ones born after the treatment cohort with a below median education level (and symmetrically for female). Among these, I randomly draw a number corresponding to 33%, 66% and 100% of the total potential migrating/manipulating population,[28] reallocate them to the "LSC" status and re-run 1000 times the triple difference model with the full set of fixed effects, on the cohorts born from 1950 onwards and aged 18 + at the time of the survey. Note that what is in principle a worry is the selection on gains in education obtained via migration/manipulation. However, I do not observe gains in education but only the actual level of education so this robustness checks only looks into levels. This is arguably a stronger test on the data, since a high *gain* in education may lead to a relatively low *level* in education, so displacing individual with high levels of education instead of high gains in education is likely to affect more the results. In addition, note that selection on *gains* may lead to a downward bias (if individuals with high gains

[25] This is an overestimation of the type of migration that concerns us, since a bias might emerge if those migrants moved in a way that would change their SC status, which is not systematically the case when crossing a district border since regions are larger than districts.

[26] In other contexts, affirmative action policies have been shown to influence racial identification (Francis and Tannuri-Pianto, 2013).

[27] And obviously no information about identity manipulation.

[28] I take the largest population between the estimates of migration and manipulation. That is migrants/manipulators that would account for 7.5% of the total male "LSC" and 11.1% of the total female "LSC".

Table 9
Robustness check: 12 years old as threshold, triple differences estimates.

	Education level	Literacy	Primary	Secondary	Higher
	A: Males				
SC∗Post 1971	0.033	0.01	0.02	0.027	−0.014
	(0.093)	(0.043)	(0.038)	(0.048)	(0.022)
	[−0.16, 0.22]	[−0.07, 0.09]	[−0.05, 0.09]	[−0.08, 0.13]	[−0.06, 0.03]
Late SC∗Post 1971	0.408	0.178	0.169	0.196	0.043
	(0.106)	(0.069)	(0.051)	(0.054)	(0.044)
	[0.22, 0.62]	[0.07, 0.30]	[0.08, 0.27]	[0.11, 0.28]	[−0.05, 0.13]
N	23,394	23,394	23,394	23,394	23,394
	B: Females				
SC∗Post 1971	0.08	0.021	0.031	0.027	0.021
	(0.094)	(0.040)	(0.039)	(0.046)	(0.014)
	[−0.09, 0.24]	[−0.05, 0.09]	[−0.04, 0.11]	[−0.05, 0.10]	[−0.00, 0.05]
Late SC∗Post 1971	−0.064	−0.066	−0.059	−0.011	0.006
	(0.114)	(0.046)	(0.055)	(0.039)	(0.033)
	[−0.27, 0.15]	[−0.15, 0.02]	[−0.15, 0.04]	[−0.09, 0.07]	[−0.05, 0.06]
N	23,222	23,222	23,222	23,222	23,222

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE, religion FE, jati∗5 years cohort FE, region∗5 years and region∗jati FE. Regressions are weighted with sample weights. Sample, Panel A: male SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey. Sample, Panel B: female SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey.

end up having relatively lower levels of education), while what is a worry is an upward bias. Indeed, an upward bias could lead us to conclude that affirmative action has benefits it doesn't actually have. Fig. 6 plots the density of the coefficients obtained in those estimations as well as the density of the lower bound of the 95% confidence intervals of these coefficients. It can be seen that the results are quite robust, since even under these strong assumptions about the education profile of migrants, the male result remains always positive and mostly statistically significant even if 100% of migrants/manipulators had an above the median education. The same is true for women: even in the case of 100% of migrants have migrated to gain SC status, the coefficient, while increasing importantly, remains non-significant in virtually all regressions.

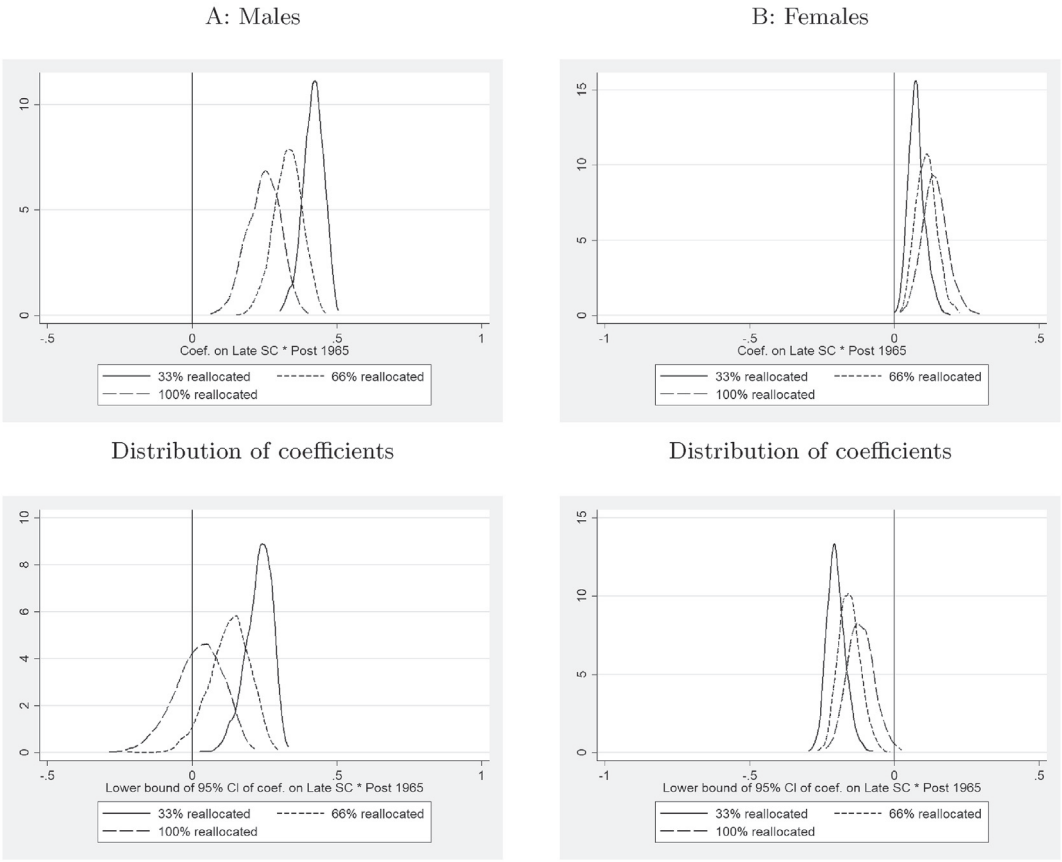


Fig. 6. Reallocation of potential migrants and robustness of the results: distribution of coefficients.

Table 10
Heterogenous effects, triple differences. Males sample.

	Education level		
SC∗Post 1965	−0.041	0.064	0.151
	(0.360)	(0.103)	(0.101)
	[−0.24, 0.15]	[−0.14, 0.26]	[−0.06, 0.36]
Late SC∗Post 1965	0.184	0.242	0.317
	(0.503)	(0.068)	(0.242)
	[−0.02, 0.39]	[0.04, 0.43]	[−0.18, 0.81]
SC∗Post 1965∗Job	0.169		
	(0.097)		
	[−0.01, 0.35]		
Late SC∗Post 1965∗Job	0.272		
	(0.264)		
	[0.03, 0.50]		
SC∗Post 1965∗School		−0.116	
		(0.072)	
		[−0.29, 0.05]	
Late SC∗Post 1965∗School		0.287	
		(0.125)	
		[−0.01, 0.58]	
SC∗Post 1965∗Social			⋮
Late SC∗Post 1965∗Social			−0.093
			(0.197)
			[−0.49, 0.31]
N	20,957	22,078	14,643

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE, religion FE, jati∗5 years cohort FE, region∗5 years and region∗jati FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: male SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey.

6. Heterogeneous effects

What are the specific dimensions of the package of affirmative action policies that are driving the results? To address this question, I will look at the heterogeneity in the results. As there is no exogenous variation on other dimension than the exposure to the SC status, and as there can not be a pre-analysis plan for this type of study, this heterogeneity analysis - as any other such analysis - is to be taken with a fair

Table 11
Heterogenous effects, triple differences. Females sample.

	Years of Education		
SC∗Post 1965	−0.032	−0.085	0.012
	(0.088)	(0.079)	(0.070)
	[−0.15, 0.09]	[−0.21, 0.04]	[−0.11, 0.13]
Late SC∗Post 1965	−0.015	0.009	0.002
	(0.097)	(0.194)	(0.093)
	[−0.20, 0.17]	[−0.29, 0.31]	[−0.22, 0.23]
SC∗Post 1965∗Job	−0.037		
	(0.111)		
	[−0.18, 0.11]		
Late SC∗Post 1965∗Job	0.086		
	(0.137)		
	[−0.14,0.33]		
SC∗Post 1965∗School		0.069	
		(0.036)	
		[−0.03, 0.17]	
Late SC∗Post 1965∗School		−0.046	
		(0.282)	
		[−0.46, 0.38]	
SC∗Post 1965∗Social			0.051
			(0.051)
			[−0.05, 0.13]
Late SC∗Post 1965∗Social			−0.004
			(0.111)
			[−0.23, 0.23]
N	20,825	21,808	15,017

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE, religion FE, jati∗5 years cohort FE, region∗5 years and region∗jati FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: female SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey.

degree of skepticism. Results can at best be interpreted as suggestive correlation and at worst as evidence of data mining, depending on the mood of the reader. As underlined in subsection 2.1, I have so far estimated the effect of the package of affirmative action policies on educational attainment. This package has three main dimensions: electoral quotas, employment quotas and education quotas and policies. It is unlikely that electoral quotas have a direct effect on educational attainment, given the small number of individuals who can benefit from these quotas.[29] However, job quotas may change the incentives to pursue education by increasing its returns (the "job quota" channel), as might policies such as free hostels and scholarships for SC, as well as any SC specific policy which would decrease the (relative) cost of education (the "social policies" channel). Finally, general education expenses may be important: SC status is useless if there are no schools one can attend (the "school supply" effect). To see which channel may be the most important, I merge my dataset with the data compiled by Pande (2003). This data contains state-year level information on the share of government jobs reserved for SC from 1958 onwards. I'll use this as a proxy for the job quota channel. It also contains the share of spendings going to education from 1958 onwards, which I'll use as a proxy for the "school supply" channel. Finally, this data contains information on share of government spendings devoted to SC welfare from 1975 onwards, which includes "among others, group housing projects, hostels for students belonging to these groups, and the provision of public goods in scheduled caste […] hamlets" (Pande, 2003): a proxy for the "social policies" channel. From this data, I construct for each spending item a simple indicator of whether in a given year t, state s is above the median spendings for that item. The descriptive statistics for these variables are in Appendix E. I match the individuals in my dataset to these variables based on their age. The budget of year t in state s is attributed to all children aged 12 in year t in state s, in line with the age threshold already used throughout the paper. I run the following regression:

$$\begin{aligned} Edu_{isjdt} = {} & \beta_1 SC_{jd} + \beta_2 SC_{jd} * post\_1965_t + \beta_3 SC_{jd} * proxy_{st} \\ & + \beta_4 SC_{jd} * post\_1965_t * proxy_{st} + \delta_1 Late\_SC_{jd} \\ & + \delta_2 Late\_SC_{jd} * post\_1965_t + \delta_3 Late\_SC_{jd} * proxy_{st} \\ & + \delta_4 Late\_SC_{jd} * post\_1965_t * proxy_{st} + \omega_1 proxy_{st} \\ & + \omega_2 post\_1965_t * proxy_{st} + constant + \gamma_1 post\_1965_t \\ & + \lambda X_{ijdt} + \epsilon_{isjdt} \end{aligned} \tag{5}$$

With the same notation as before, with s referring to the state and $proxy_{st}$ the value taken by a given proxy in state s at time t. Table 10 shows that for males, the "school supply" and the "job quota" channels may have been important contributors to the increase in education (but their statistical significance is weak[30]). However, the "social policy" channel does not seem to be relevant. Finally, note also that these channels do not seem to be necessary conditions: the coefficient on Late SC∗Post 1965 is positive and large in all specification, even if not precisely estimated. Note also that due to the smaller sample size and very demanding combination of FE, the coefficient on SC∗Post 1965∗Social could not be estimated for males. Table 11 shows that none of these channels seem to explain the lack of impact of SC status on females educational attainment. However, let me once again underline that all the estimates in this section are pure correlations and can not be interpreted causally.

7. Conclusion

This paper studies the impact of affirmative action policies for SC in India. It shows that they positively affect schooling attainment, in particular literacy and secondary schooling. Affirmative action policies have changed the educational attainment of the average SC population. In addition, there are stark differences across genders: males benefit from these policies at all levels of education below higher education, and females appear not to. This suggests that individuals at the intersection of different discriminated groups may not be sufficiently well protected by public policies. More generally, this paper underlines the need to focus on the role of accumulation of discrimination, the question of "intersectionality", which so far tends to be neglected by policy makers.

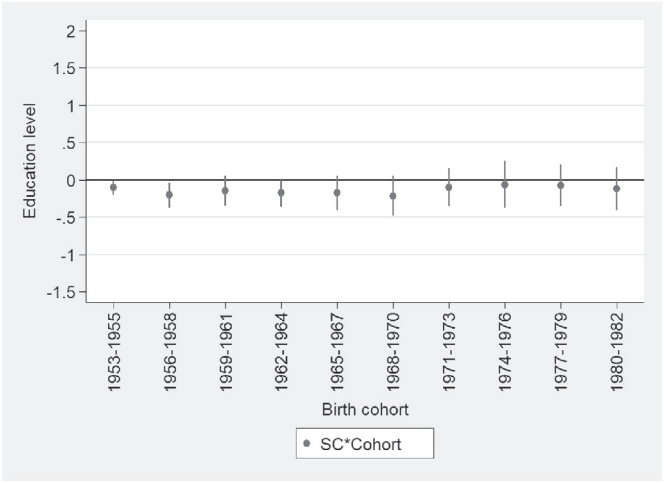
Appendices.

A. *Fully flexible specifications for Early SC*

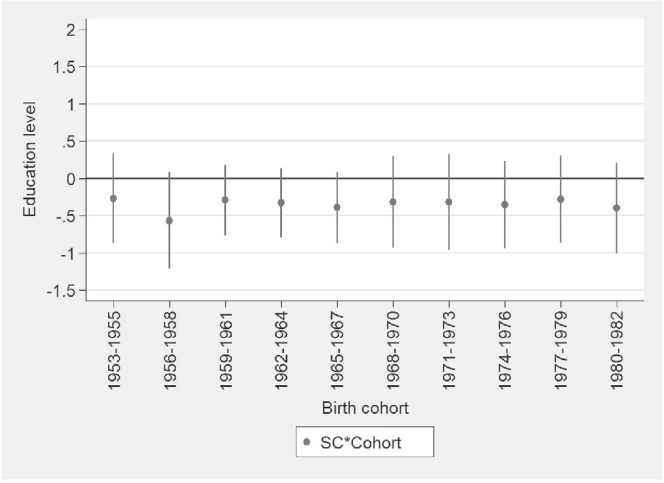
Fig. 7 presents the β_t coefficients of regression described in Model 2. It can be seen that the level of education of ESC is relatively stable across the 1950–1981 cohorts, pointing to the fact that crowding out does not seem to be an issue.

[29] This is not to say that the persons elected under these quotas can not themselves affect policies which may lead to a change in educational attainment as in Pande (2003).

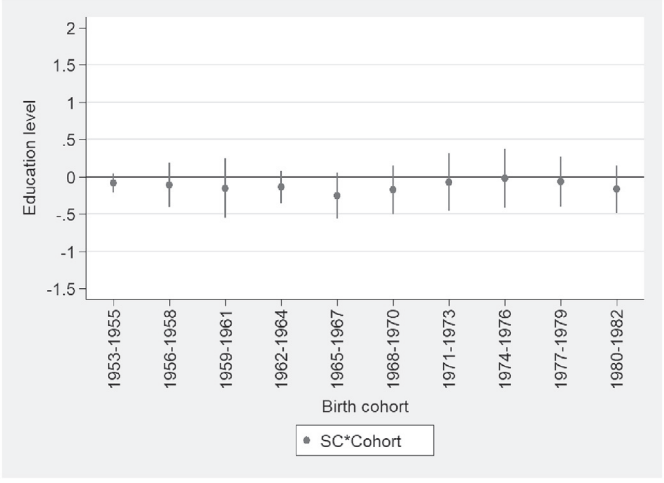
[30] The coefficient on Late SC∗Post 1965∗School is less significant once standard errors are bootstrapped, while the coefficient on Late SC∗Post 1965∗Job gains on statistical significance with the bootstrapped standard errors.



Both genders



Males

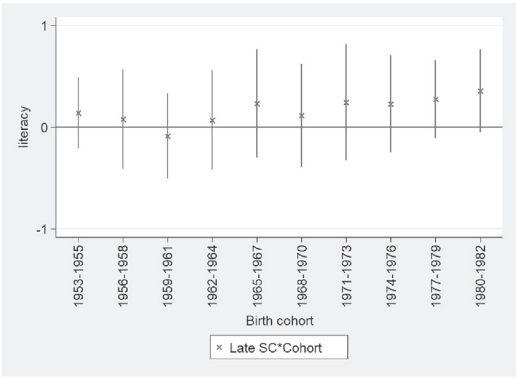


Females

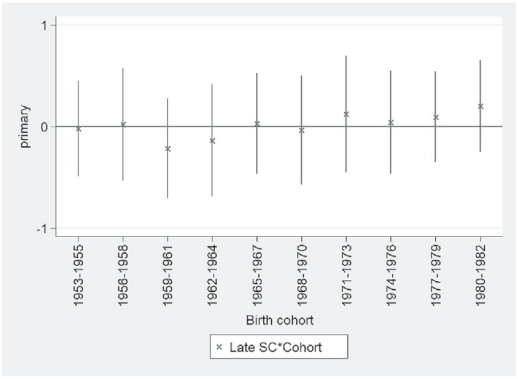
Fig. 7. Cohort specific effects, triple differences estimates.
Data source: Round 2 of the Indian DHS. Sample: SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey. The control cohorts are 1950–1952.

B. Fully flexible specifications, other outcomes

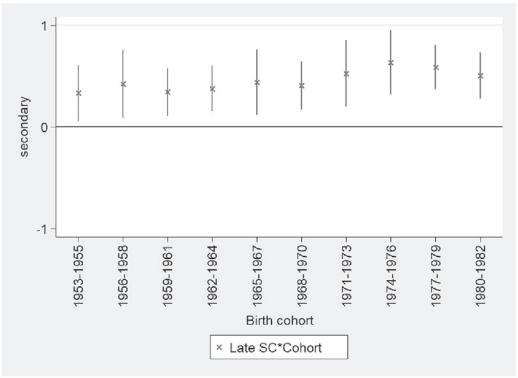
Figs. 8 and 9 present the coefficients of interest of the regression described in Model 2, using literacy, primary, secondary and higher education completion as outcomes.



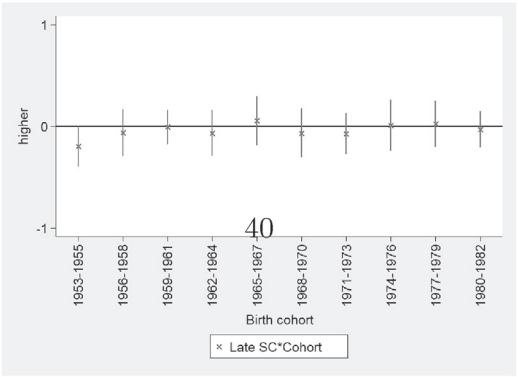
Literacy



Primary

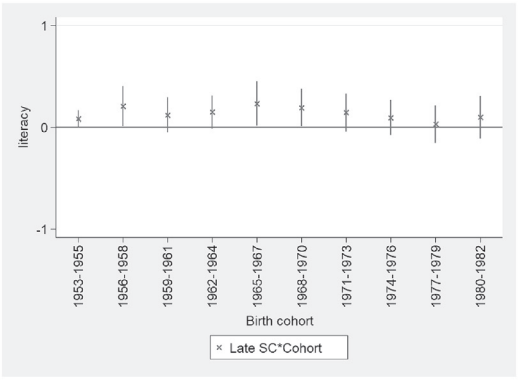


Secondary

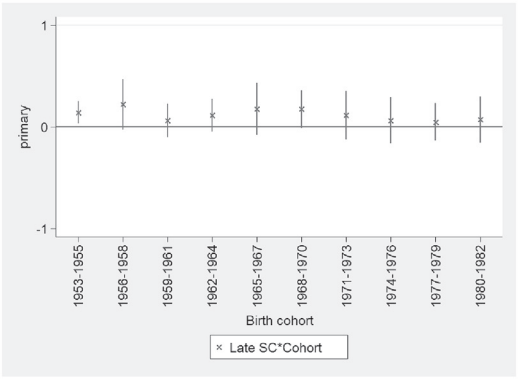


Higher

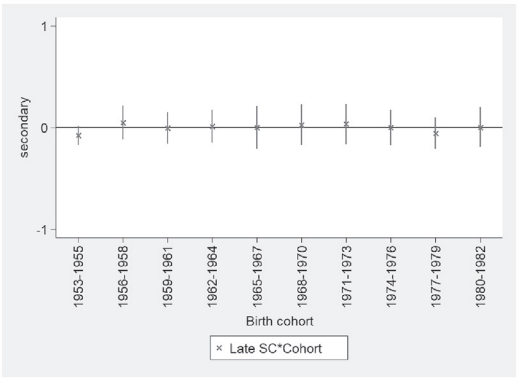
Fig. 8. 1976 natural experiment: male cohort specific effects, triple differences estimates, alternative outcomes.
Data source: Round 2 of the Indian DHS. Sample: male SC population born from 1950 onwards and aged 18 and above at time of survey. Cohorts born from 1965 onwards are younger than 12 at the time of the reform, and are likely to have benefited from it. The control cohorts are 1950–1952.



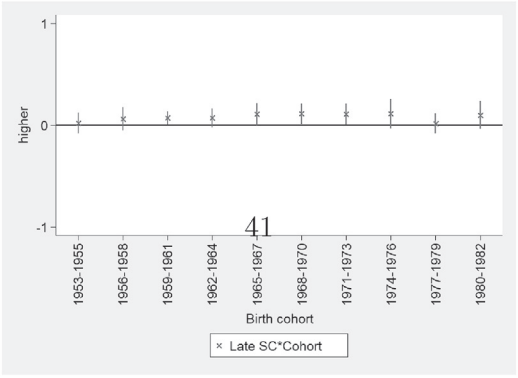
Literacy



Primary



Secondary



Higher

Fig. 9. 1976 natural experiment: female cohort specific effects, triple differences estimates, alternative outcomes.
Data source: Round 2 of the Indian DHS. Sample: female SC population born from 1950 onwards and aged 18 and above at time of survey. Cohorts born from 1965 onwards are younger than 12 at the time of the reform, and are likely to have benefited from it. The control cohorts are 1950–1952.

C. Quadruple differences with gender heterogeneity

This section reproduces the triple differences estimates of Section 4.3, allowing the coefficients to vary by gender. That is, it is a test of the statistical significance of the observed difference in the treatment effect by gender. Note that all controls and fixed effects are also allowed to vary by gender. It can be seen in Table 12 that indeed, the effect on the education of males is significantly different from that on females.

Table 12
Quadruple differences estimates, gender heterogeneity.

	Education level	Literacy	Primary	Secondary	Higher
SC∗Post 1965	−0.04	−0.049	−0.024	−0.005	−0.011
	(0.034)	(0.165)	(0.019)	(0.028)	(0.737)
	[−0.14, 0.06]	[−0.11, 0.01]	[−0.08, 0.03]	[−0.05, 0.04]	[−0.03, 0.01]
SC∗Post 1965∗male	0.074	0.048	0.039	0.02	0.015
	(0.101)	(0.049)	(0.052)	(0.042)	(0.031)
	[−0.09, 0.24]	[−0.04, 0.13]	[−0.04, 0.12]	[−0.05, 0.09]	[−0.03, 0.06]
Late SC∗Post 1965	0.041	0.004	0.002	0.004	0.035
	(0.147)	(0.014)	(0.051)	(0.033)	(0.027)
	[−0.10, 0.17]	[−0.05, 0.07]	[−0.09, 0.09]	[−0.06, 0.07]	[−0.01, 0.08]
Late SC∗Post 1965∗male	0.328	0.163	0.145	0.164	0.019
	(0.104)	(0.050)	(0.065)	(0.052)	(0.038)
	[0.14, 0.52]	[0.07, 0.25]	[0.03, 0.27]	[0.07, 0.26]	[−0.05, 0.09]
N	46,616	46,616	46,616	46,616	46,616

Standard errors two way clustered at the jati-gender and region-gender level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati∗gender FE, region∗gender FE, year of birth∗gender FE, religion∗gender FE, gender∗jati∗5 years cohort FE, gender∗region∗5 years cohort FE and gender∗region∗jati FE. Regressions are weighted with sample weights. The number of observation reported reports to the number of observations effectively used in the regression (ie. ignoring observations dropped due to collinearity). Sample: SC, OBC and SCSE population born from 1950 onwards and aged 18 and above at time of survey.

D. Using the 1950 natural experiment

This section presents the results of a parametric estimation of the impact of the access to the SC status both in 1950 and in 1976. See subsection 5.1 for more details. I run the following regression:

$$\begin{aligned} Edu_{ijdt} =& \beta_1 SC_{jd} + \beta_2 SC_{jd} * post\_treatment\_1950_t + \beta_3 SC_{jd} * post\_1965_t \\ &+ \delta_1 Late\_SC_{jd} + \delta_2 Late\_SC_{jd} * post\_treatment\_1950_t + \delta_3 Late\_SC_{jd} * post\_1965_t \\ &+ constant + \gamma_1 post\_treatment\_1950_t + \gamma_2 post\_1965_t + \lambda X_{ijdt} + \epsilon_{ijdt} \end{aligned} \tag{D.1}$$

With the same notation as earlier and $post\_treatment\_1950_t$ a dummy indicating if a cohort was young enough to benefit from the access to the SC status in 1950, that is, aged 12 or below. It can be seen in Table 13 that the effect of the access to the SC status in 1950 seems comparable to that of 1976, but less precisely estimated, due to the smaller number of observations in these cohorts. The descriptive statistics for this sample are in Tables 14 and 15.

Table 13
Triple differences estimates, combined 1950 and 1976 natural experiments.

	Education level	Literacy	Primary	Secondary	Higher
	A: Males				
SC∗Post 1938	0.034	−0.056	−0.05	0.057	0.027
	(0.062)	(0.055)	(0.041)	(0.027)	(0.023)
	[−0.11, 0.19]	[−0.16, 0.06]	[−0.15, 0.05]	[−0.02, 0.13]	[−0.01, 0.07]
Late SC∗Post 1938	−0.262	−0.132	−0.05	−0.198	−0.014
	(0.224)	(0.115)	(0.124)	(0.109)	(0.045)
	[−0.63, 0.10]	[−0.35, 0.09]	[−0.29, 0.17]	[−0.37, −0.02]	[−0.08, 0.06]
SC∗Post 1965	0.014	−0.011	0.011	0.009	−0.005
	(0.072)	(0.035)	(0.034)	(0.033)	(0.023)
	[−0.11, 0.14]	[−0.07, 0.05]	[−0.06, 0.08]	[−0.06, 0.08]	[−0.04, 0.03]
Late SC∗Post 1965	0.290	0.145	0.116	0.145	0.029
	(0.064)	(0.032)	(0.039)	(0.033)	(0.020)
	[0.18, 0.41]	[0.09, 0.20]	[0.03, 0.20]	[0.09, 0.19]	[−0.02, 0.08]
N	28,942	28,942	28,942	28,942	28,942
	B: Females				
SC∗Post 1938	0.128	0.042	0.054	0.065	0.009
	(0.065)	(0.038)	(0.039)	(0.024)	(0.013)
	[0.17, 0.24]	[−0.03, 0.11]	[−0.01, 0.12]	[0.02, 0.11]	[−0.02, 0.03]
Late SC∗Post 1938	0.022	0.012	0.008	0.011	0.003
	(0.064)	(0.041)	(0.045)	(0.031)	(0.010)
	[−0.09, 0.13]	[−0.05, 0.08]	[−0.06, 0.08]	[−0.05, 0.07]	[−0.02, 0.03]

(continued on next page)

Table 13 *(continued)*

	Education level	Literacy	Primary	Secondary	Higher
SC∗Post 1965	−0.02	−0.03	−0.007	0.001	−0.015
	(0.081)	(0.041)	(0.044)	(0.027)	(0.017)
	[−0.15, 0.11]	[−0.09, 0.04]	[−0.07, 0.07]	[−0.04, 0.04]	[−0.04, 0.01]
Late SC∗Post 1965	0.044	−0.014	−0.011	0.013	0.042
	(0.118)	(0.041)	(0.056)	(0.041)	(0.036)
	[−0.16, 0.25]	[−0.09, 0.06]	[−0.11, 0.09]	[−0.07, 0.10]	[−0.02, 0.10]
N	28,741	28,741	28,741	28,741	28,741

Standard errors two way clustered at the jati and region level in parentheses. 95% confidence interval generated with the cgmwildboot procedure in bracket. All regressions include: jati FE, region FE, year of birth FE, religion FE, jati∗5 years cohort FE, region∗5 years and region∗jati FE. Regressions are weighted with sample weights.

Table 14
Descriptive Statistics: Male population aged 18 to 65.

	Late SC		Early SC		OBC and SCSE	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education variables						
Education level	1.857	1.493	1.948	1.481	2.268	1.415
Literacy	0.613	0.487	0.637	0.481	0.733	0.442
Primary Completion	0.642	0.480	0.667	0.471	0.759	0.428
Secondary Completion	0.446	0.497	0.477	0.500	0.561	0.496
Higher Education Completion	0.125	0.331	0.133	0.340	0.186	0.389
Control variables						
Hindu	0.988	0.109	0.898	0.303	0.980	0.139
Sikh	0.004	0.061	0.062	0.240	0.020	0.139
Buddhist/Neo Buddhist	0.008	0.091	0.041	0.198	0.000	0.008
N	1081		11,691		16,882	

Sample: Male population aged 18 to 65 at the time of the survey. Data source: NFHS2.

Table 15
Descriptive Statistics: Female population aged 18 to 65.

	Late SC		Early SC		OBC and SCSE	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education variables						
Education level	0.858	1.351	0.946	1.365	1.188	1.464
Literacy	0.286	0.452	0.314	0.464	0.391	0.488
Primary Completion	0.303	0.460	0.343	0.475	0.420	0.494
Secondary Completion	0.192	0.394	0.205	0.404	0.271	0.444
Higher Education Completion	0.055	0.228	0.051	0.219	0.074	0.261
Control variables						
Hindu	0.987	0.114	0.894	0.308	0.983	0.131
Sikh	0.003	0.053	0.061	0.239	0.017	0.130
Buddhist/Neo Buddhist	0.010	0.101	0.045	0.208	0.000	0.008
N	1070		11,528		16,866	

Sample: Female population aged 18 to 65 at the time of the survey. Data source: NFHS2.

E. Policy channel variables

Table 16 presents the descriptive statistics of the variables used in Section 6. These variables are computed from the database of Pande (2003), and are presented here at the state∗year level, which is their level of variation. Note that "Years available" refer to the years that I can exploit: for example, I did not use the data beyond 1993, since the latest cohort used in the paper is 1981.

Table 16
Policy Variables.

Variable	Obs	Mean	Std. Dev.	Min	Max	Years available	Birth Years
Social	223	0.50	0.5	0	1	1975–1991	1963–1980
School	406	0.50	0.5	0	1	1963–1993	1950–1981
Jobs	399	0.48	0.5	0	1	1963–1992	1950–1980

Appendix F. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jdeveco.2018.10.001.

References

Arcidiacono, Peter, September 2005. Affirmative action in higher education: how do admission and financial aid rules affect future earnings? Econometrica 73 (5), 1477–1524.

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

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

Bertrand, Marianne, Hanna, Rema, Mullainathan, Sendhil, 2010. Affirmative action in education: evidence from college admissions in India. J. Publ. Econ. 94 (1–2), 16–29.

Caskey, Judson, 2015. Cgmwildboot. Stata ado file.

Cassan, Guilhem, 2015. Identity based policies and identity manipulation: evidence from colonial Punjab. Am. Econ. J. Econ. Policy 7 (4), 103–131.

Cassan, Guilhem, Van de Walle, Lore, 2017. The Social Cost of Social Norms: Gender Quotas Meet Gender Norms. Graduate Institute of International and Development Studies Working Paper HEIDWP18-2017.

Chakravarti, Uma, 1993. Conceptualising brahmanical patriarchy in early India: gender, caste, class and state. Econ. Polit. Wkly. 28 (14).

Chalam, K.S., 1990. Caste reservations and equality of opportunity in education. Econ. Polit. Wkly. 25 (41).

Chin, Aimee, Prakash, Nishith, 2011. The redistributive effects of political reservation for minorities: evidence from India. J. Dev. Econ. 96 (2), 265–277.

Chitnis, Suma, 1972. Education for equality: case of scheduled castes in higher education. Econ. Polit. Wkly. 7 (31–33).

Crenshaw, Kimberle, 1989. Demarginalizing the intersection of race and sex: a black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. Univ. Chicago Leg Forum 1989 (1).

Crenshaw, Kimberle, 1991. Mapping the margins: intersectionality, identity politics, and violence against women of color. Stanford Law Rev. 43 (6), 1241–1299.

Desai, Sonalde, Kulkarni, Venna, May 2008. Changing educational inequalities in India in the context of affirmative action. Demography 45 (2), 245–270.

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

Duflo, Esther, 2001. Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment. Am. Econ. Rev. 4 (91), 795–813.

Field, Erica, Pande, Rohini, Jayachandran, Seema, 2010. Do traditional institutions constrain female entrepreneurship? A Field experiment on business training in India. In: American Economic Review Papers and Proceedings.

Foster, Andrew, Rosenzweig, Mark, 2003. Technological change and the distribution of schooling: evidence from green-revolution India. J. Dev. Econ. 74 (1), 87–111.

Francis, Andrew, Tannuri-Pianteo, Maria, 2012. Using Brazil's racial continuum to examine the short-term effects of affirmative action in higher education. J. Hum. Resour. 47 (3), 754–784.

Francis, Andrew, Tannuri-Pianto, Maria, 2013. Endogenous race in Brazil: affirmative action and the construction of racial identity among young adults. Econ. Dev. Cult. Change 61 (4), 751–753.

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

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

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

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

Hinrichs, Peter, August 2012. The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. Rev. Econ. Stat. 94 (3), 712–722.

Hnatkovska, Viktoria, Lahiri, Amartya, Paul Sourabh, B., 2012. Castes and labor mobility. Am. Econ. J. Appl. Econ. 4 (2).

Hnatkovska, Viktoria, Lahiri, Amartya, Paul Sourabh, B., 2013. Breaking the caste barrier: intergenerational mobility in India. J. Hum. Resour. 48 (2).

Jensen, Robert, 2012. Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. Q. J. Econ. 2 (127), 153–192.

Jensenius, Francesca, 2015. Development from representation? A study of quotas for scheduled castes in India. Am. Econ. J. Appl. Econ. 7 (3), 196–220.

Kapadia, Karin, 1997. Mediating the meaning of market opportunities: gender, caste and class in Rural South India. Econ. Polit. Wkly. 35.

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

Kingdon, Geetah, 1998. Does the labour market explain lower female schooling in India? J. Dev. Stud. 35 (1).

Kitts, Eustace J., 1885. A Compendium of the Castes and Tribes Found in India. Education Society Press, Byculla.

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

Mencher, Joan P., 1988. Women's work and poverty: women's contribution to household maintenance in South India. In: Dwyer, Daisy, Bruce, Judith (Eds.), Home Divided. Stanford University Press.

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

Pande, Rohini, 2003. Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India. Am. Econ. Rev. 93 (4), 1132–1151.

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

Rosenzweig, Mark, Schultz, T. Paul, 1982. Market opportunities, genetic endowments, and intrafamily resource distribution: child survival in Rural India. Am. Econ. Rev. 72 (4), 803–815.

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