Journal of Public Economics 94 (2010) 16–29

Contents lists available at ScienceDirect

Journal of Public Economics

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

Affirmative action in education: Evidence from engineering college admissions in India

Marianne Bertrand [a,b,c,d], Rema Hanna [b,e,f,g,*], Sendhil Mullainathan [b,f]

[a] University of Chicago Booth School of Business, United States
[b] NBER, United States
[c] CEPR, United Kingdom
[d] IZA, Germany
[e] Kennedy School of Government, United States
[f] Harvard University, United States
[g] BREAD, United States

ARTICLE INFO

Article history:
Received 12 February 2009
Received in revised form 27 October 2009
Accepted 5 November 2009
Available online 12 November 2009

Keywords:
Education
Affirmative action
Discrimination
Development

ABSTRACT

This paper examines an affirmative action program for "lower-caste" groups in engineering colleges in India. We study both the targeting properties of the program, and its implications for labor market outcomes. We find that affirmative action successfully targets the financially disadvantaged: the upper-caste applicants that are displaced by affirmative action come from a richer economic background than the lower-caste applicants that are displacing them. Targeting by caste, however, may lead to the exclusion of other disadvantaged groups. For example, caste-based targeting reduces the overall number of females entering engineering colleges. We find that despite poor entrance exam scores, lower-caste entrants obtain a positive return to admission. Our estimates, however, also suggest that these gains may come at an absolute cost because the income losses experienced by displaced upper-caste applicants are larger than the income gains experienced by displacing lower-caste students. Limited sample sizes in our preferred econometric specifications, however, prevent us from drawing strong conclusions from these labor market findings.

1. Introduction

Many countries around the world mandate affirmative action programs in higher education, where groups that have historically faced discrimination are given preferential admissions to colleges and universities. In the context we study (state-controlled colleges in India), more than 50% of admissions slots are typically reserved for the members of lower-caste groups. We propose to explore the effectiveness of one such program in redistributing both opportunity and income to marginalized groups.

We are motivated by the vigorous debate surrounding affirmative action programs (for a more in-depth discussion of the debate, see Fryer et al., 2003; Holzer and Neumark, 2000). The debate is focused on two main issues. The first issue centers on the question of who these programs actually *target*. Many claim that affirmative action only benefits a small sliver of the population: those of the traditionally disadvantaged group that actually come from economically-advantaged backgrounds. Moreover, if the marginal admit of the "disadvantaged" group comes from a richer household than the marginal admit of the "advantaged" group, affirmative action programs may actually be regressive in nature. Put another way, affirmative action programs may take slots away from disadvantaged members of an advantaged *group* and give them to individuals who are more advantaged, but just happen to belong to a disadvantaged *group*.

The second issue is whether and how much these programs actually help those who gain admissions—the so-called *mismatch hypothesis* (see, for example, Alon and Tienda, 2005; Herrnstein and Murray, 1994; Kane, 1998; Loury and Garman, 1993; Rothstein and Yoon, 2007, Rothstein and Yoon, 2008; Thernstrom and Thernstrom, 1997). One side argues that because these programs place minorities in an academic environment for which they are unprepared, those who gain admissions due to affirmative action will do poorly in their classes, be more likely to drop out, and flounder in the job market. In the worst-case scenario, affirmative action may actually make the minority admits worse off, either through lost time spent in college or through a discouragement effect. At the other extreme, some argue that such programs may result in net gains for the disadvantaged groups and for society—that is, minority students may not only benefit from the education they receive, they may actually benefit *more* than the majority students they are taking slots from. Thus, affirmative action might be consistent with an efficient reallocation of educational slots.

* Corresponding author. Harvard University, United States.
E-mail address: rema_hanna@ksg.harvard.edu (R. Hanna).

0047-2727/$ – see front matter
doi:10.1016/j.jpubeco.2009.11.003

Despite the importance of this debate, data restrictions have hampered a thorough empirical investigation.[1] Most existing studies of affirmative action are limited to studies of the outcomes of those who were admitted and matriculated. The lack of follow-up data on those who do *not* gain admissions makes it hard to test several of the above hypotheses, such as the targeting properties of the programs and the costs to those who "lost out." As Fryer and Loury (2005) discuss, there are strong beliefs that affirmative action programs greatly harm non-minority groups and this belief dampens support for such programs. But these beliefs are hard to evaluate, since there is very little real evidence to date on the magnitude (if any) of the harm. The lack of data on those who did not matriculate is particularly problematic. This study assembles the most comprehensive dataset available on affirmative action in higher education (to our knowledge) by collecting detailed family background and follow-up data on a pool of initial applicants, regardless of whether they were admitted or whether they matriculated.

We do so in the Indian context. In traditional Hindu society, individuals are born into a caste group that determines one's rights and responsibilities related to occupation, social standing, and permissible forms of social interactions with members of other caste groups. Since caste is hereditary and also tended to be occupation-specific, occupation was traditionally determined by birth (Osborne, 2001). Munshi and Rosenzweig (2006), for example, show how caste-based labor markets have trapped individuals in narrow occupational categories for generations, and persist even today. In an attempt to reduce these caste-based inequalities, state universities in India reserve seats for individuals from each of the traditionally disadvantaged groups—the Scheduled Tribes (ST), the Scheduled Caste (SC), and the Other Backward Castes (OBC).

The system of admissions and reserved seats is quite straightforward. To apply to a university, each applicant takes a standardized exam. Admission is a deterministic function of one's exam score and caste group. Within each caste group, only those who have scored above a given threshold for their caste group are admitted. Given the simple structure of this process, we can compare various background characteristics and outcomes for applicants above and below the admission threshold in each caste group. This allows us to pursue two main questions. First, we investigate the targeting properties of the affirmative action policy. Second, we ask whether the minority applicants who are favored in admissions by the affirmative action policy actually benefit economically from it, and if yes, how these gains compare to the potential losses of the majority of applicants whom the policy is displacing.[2]

We focus on individuals applying for entrance to engineering colleges in 1996 in one Indian state. We chose to study engineering colleges because they are among the most prestigious types of colleges in India. To implement our research design, we collected two data sets. First, we collected a census of all individuals who took the admissions exam in 1996. The minimum score for admission is roughly 480 out of 900 for upper-caste individuals, 419 for the OBC category, and 182 for the SC category.[3] These score disparities at entrance confirm the plausibility of the hypothesis that lower-caste students will simply not be able to perform in college and, hence, will not benefit from admission because of the mismatch between their basic skill level and the skill requirements of an engineering education. To better understand outcomes across caste groups, we interviewed about 700 households from this census of *applicants* between 2004 and 2006 (approximately 8–10 years after the entrance exam). We surveyed both the applicant and their parents to obtain socio-background information, and gauge various life outcomes including income and occupation.[4]

Our first finding is that even though applicants are positively selected from the population, caste-based targeting does result in the redistribution of education resources to those who are more economically disadvantaged. Lower-caste applicants to the program have incomes that are typically about four times as large as the average minority in their state. However, under a reasonable set of assumptions, the parental income among upper-caste students displaced by the affirmative action policy is Rs 14,088 ($282) per month, compared to Rs 8340 ($167) per month among the displacing lower-caste students. Similarly, 41% of displaced students come from a household in which the head holds at least a master's degree, compared to only 14% of displacing students. Thus, even though those who are aided by affirmative action are richer on average, they are still relatively poorer than the students that they replace. However, caste-based targeting falls short of increasing the representation of all traditionally disadvantaged groups. Specifically, we find that females may lose out as a result of the policy. While women comprise 23% of those displaced by the reservation policy, they comprise only 16% of the "displacers." Thus, gender diversity is reduced because fewer females end up attending engineering school.

Next, we examine labor market outcomes. Our point estimates suggest that despite much lower basic skills (as measured by the score on the entry exam), those who are admitted by affirmative action economically benefit from attending engineering college. Depending on the specifications, attending engineering school increases lower-caste members' income by between Rs 3500 ($70) and Rs 6200 ($124) per month. However, our point estimates also suggest that the affirmative action programs come at an absolute cost, because the negative income *level* effect experienced by displaced upper-caste applicants is much larger than the positive income level effect experienced by displacing lower-caste students. Unfortunately, a limited sample size and large standard errors in the preferred econometric specifications prevent us from drawing strong conclusions from these labor market findings.

The rest of the paper proceeds as follows. Section 2 discusses the history of affirmative action policies in India. Section 3 describes the data and methodologies. Section 4 studies the targeting question, while Section **5** explores the labor market implications. Section 6 concludes with a discussion of our main findings and suggestions for future work.

2. History of affirmative action policies in India

The roots of India's current affirmative action policies date back to the colonial-era reservation programs for government jobs, political representation, and scholarships established under British rule (Kumar, 1992). The first recorded affirmative action policies in India were quota systems for administrative positions in the more progressive state governments, such as Mysore in South India, and

[1] Recent exceptions include Bowen and Bok (1998) who has examined affirmative action in higher education using detailed micro-admissions data similar to our own and Arcidiacono (2005), who has used a structural model to estimate the impact of affirmative action policies in higher education.

[2] Because affirmative action in education has existed in India for over fifty years, it may have had long-term investment effects, with households changing their primary and secondary school decisions and behaviors in response. This would affect the distribution of test scores we observe on the entrance exam in ways that are not obvious. On the one hand, lower caste families may be more likely to invest in secondary education if they believe that they now have a "shot" at getting into an engineering college. On the other hand, lower caste individuals may put less effort into their primary and secondary school work if the probability of admissions in their specific caste category is high even with low test scores. Estimating these longer-term effects is beyond the scope of our analysis: we cannot simulate the full counterfactual of what would happen in the absence of affirmative action.

[3] Applicants from the ST category were dropped from the analysis. Due to the large number of available seats and few candidates from this group, a large fraction of applicants from this category obtained admissions.

[4] Note that there are numerous engineering colleges of varying levels of quality. As we have attempted to survey those around the cutoff, we are in effect measuring the effect of admissions to an engineering college that an individual's score and caste would have qualified him or her for, rather than the effect of any engineering college *per se*.

Baroda and Kolhapur in Western India. Political reservations soon followed: the first caste-based lists of reserved, parliamentary seats were published under British rule in 1936 (Osborne, 2001). With independence from Britain in 1947, Jawaharal Nehru and the Congress Party chose to maintain the British system of reserving seats in the legislature (Osborne, 2001). Article 46 of India's Constitution contained a directive to provide special considerations to members of the Scheduled Castes (SC) and Scheduled Tribes (ST), the two most disadvantaged groups in India. This paved the way for both the central government and the individual states to adopt reservations in college admissions for two of the groups: the SC and ST (Weisskopf, 2003).

Reserved seats for members of "other backward categories" (OBC)—a group that was more advantaged than the SC and ST, but less advantaged than the upper-caste groups—are less prevalent. Although the constitution ostensibly prohibited discrimination against the OBC, it did not institute any specific affirmative action policies for them. Then in 1953, the central government formed a commission to study the situation of the OBC, and the commission recommended that an additional 2399 "backward" castes—roughly 40% of the population—should also be eligible for reservations (de Zwart, 2000). However, this recommendation was not acted upon. Instead, the central government left the power to grant concessions to other backward caste groups to individual states. Thus, while many states began to institute reserved seats for the OBC in state-controlled colleges, there were no reserved seats for OBC in nationally controlled colleges (Baley, 1999). In 1978, the central government formed a second exploratory commission—known as the Mandal Commission—to explore again the situation of the OBC. It identified 3747 castes, or 52% of the population, as backward (de Zwart, 2000; Wolpert, 2006). It recommended a 27% reservation for the OBC for university admissions, public sector jobs and all private sector endeavors that receive financial assistance from the government (Kumar, 1992; Weisskopf, 2003).[5] The report was shelved until 1990, when the central government reported that it would enforce the recommendations of the Commission only for public sector jobs. This announcement led to widespread rioting and protest (Osborne, 2001). In 1992, the Supreme Court upheld the Mandal reservations public sector jobs, and also ruled that the central government and all state governments must exclude the "creamy layer," i.e. richer members, of the OBC from the reservations in public sector jobs (Osborne, 2001; Wolpert, 2006).

The original reservations specified in India's Constitution were set to expire in 1960. They have since been extended several times over, and are now set to expire in 2010. In all colleges controlled by the central government, 7.5% of seats have been reserved for the ST and 15% of seats have been reserved for the SC, for a total of 22.5% of reserved seats. There are no reservations for the OBC in these centrally-controlled colleges. In colleges controlled by state governments, the percentage of seats reserved for ST/SC depends on the approximate proportions of these groups in each state. After the Mandal Commission, many states also implemented the 27% reservations for OBCs to varying degrees in state-controlled colleges.

Beginning in 2005, reservations in education were once again at the forefront of Indian politics. In August 2005, the Supreme Court ruled that the government could not impose quotas in privately-funded colleges. The ruling led to the passage of the 93rd Constitution Amendment Act, which gave the government the power to institute affirmative action policies in all "educational institutions." In May 2006, the government announced a plan to extend the 27% reservation for the OBC to all central universities, resulting in massive protests. Parliament passed the bill in the winter session of 2006–2007. However, in March 2007, the Supreme Court gave a stay order on the bill, citing the lack of data on which groups are indeed economically and socially disadvantaged. Thus, as in many other countries, affirmative action continues to be a contentious topic in the Indian education policy debate.

3. Institutional context, sample construction and survey design

This section details the admission process for an engineering college in the state where our research occurs. It also discusses the sample construction and survey design. We conclude with a discussion of attrition.

3.1. Admission process for an engineering degree

During the start of his or her final year of high school, an individual must take a general school leaving/college admissions examination (called the "10 + 2 examination"). The individual's score on this exam and caste category fully determine his or her admissions to most colleges. To gain admission into very prestigious programs or universities—such as medical and engineering colleges—in a particular state, each individual must first pass the 10 + 2 exam and then take a second-round statewide entrance exam.

In this paper, we use data from the second-round entrance exam to engineering colleges in one Indian state. We chose to study engineering colleges as this is one of the more prestigious educational tracks in India. Those who turn down admission to an engineering college in their state would mainly do so because they obtained admissions to medical school or to a better engineering school in another state. Those who do not gain admission to an engineering school typically still attend some form of higher education, but usually wind up in less prestigious majors in local colleges (the most common being Commerce or Economics), go to 2-year technical schools, or go to private engineering schools (which are not as prestigious as the government schools, are not subject to reservations, and are more costly).

The second-round entrance exam to engineering colleges consists of three sections—math, physics and chemistry—each with 100 multiple choice questions worth 3 points each, for a total of 900 points. After taking the exam, individuals are ranked based on their total score in each caste category (General, OBC, SC, and ST).[6] Starting from the highest ranking in each caste category, individuals are then invited for a counseling session where they are informed about possibilities for admissions to the different engineering colleges. Individuals then choose whether they would like to attend, and if so, which college and type of engineering course they want to pursue. The process continues down the rankings until all seats in the colleges are full. Note that individuals with a higher ranking get a choice about which institute to attend and which subject within engineering to study. Since the reservation percentages cut across each institution and each branch, it is impossible to place all members from the reserved categories in only the worst colleges or branch. As such, the distribution of college quality attended is the same across the different caste groups.

Since states can increase the reserved seats for each caste group, the actual percentage of reserved seats varies from state to state. In the state-year we study, a total of 2643 seats were available, with 2054 seats open to the reservations policy; the remainder were

[5] The Commission would have preferred to recommend that the OBC groups receive a 52% quota, in addition to the 22% already set aside for SC and ST. However, the Supreme Court had already ruled that central-government reservations could not exceed 50%. As such, the Commission recommended a 27% reservation for the OBC (Wolpert, 2006).

[6] For individuals with the same score, individuals are ranked first by math score, then by chemistry score, and finally by physics score. In the circumstance that individuals have the same three scores, the birthday determines the time—the older individual obtains the higher rank. As such, each individual has a unique ranking.

payment seats not covered by the policy. The quotas were determined by the distribution of castes in the state: there was a 16% reservation for the SC, a 21% reservation for the ST, and a 14% reservation for the OBC, for a total of 51% of seats reserved. In each caste category, there were additional reservations for those who entered from a technical stream, were the child or grandchild of a “Freedom Fighter,” or were military personnel (saniks).

As indicated above, we obtained data from the 1996 entrance exam. For all individuals, the data included: individual's name, father's name, address in 1996, scores on each portion of the exam, rank, birth date, caste group, and an indicator for whether the individual qualified for any other reserved category. We also obtained a list of all applicants who actually attended engineering colleges in the state in 1996.

3.2. Sample determination

When determining the research sample, we dropped two types of individuals. First, we excluded individuals who qualified for the special reserved categories (freedom fighter, etc.). Second, we excluded the ST category from the analysis as most of the ST applicants gained admission to the engineering colleges.

Fig. 1 presents test score distributions by caste category. Lower-caste applicants, and especially SC applicants, scored lower on average on the entrance exam. The upper-caste group (general category) and OBC group also had a larger range of test scores than the SC group.

We aimed to survey applicants who were both above and below the admission cutoff in each caste category. However, in this particular context, the “true” cutoff scores for admission are unknown. We therefore need to estimate the cutoff scores. To proceed, we follow the first of the two methods described in Card et al. (2008). In particular, for each caste category, we select as the cutoff the score that maximizes the goodness of fit (R_2) in a regression of engineering college attendance on a constant and a dummy equal to one if one's score is above a particular score threshold, 0 otherwise. Applying this technique to the enrollment data, we estimated score cutoffs of 182 for the SC, 419 for the OBC, and 490 for the upper-caste group.

In practice, Card et al. estimate their cutoff on a randomly selected subset of their data, and use the remaining data for analysis. Our limited sample size precludes us from doing so. However, it is important to note that our set-up is quite different from Card et al. Specifically, they test for the existence of a break-point for their key variable, and then estimate the magnitude of the jump at that estimated break-point for that same variable. This raises concern of specification search bias if the same data is used both to find the tipping point and estimate its magnitude. In our setting, we estimate the break-point in the attendance variable, and then estimate the magnitude of the jump for other variables, such as income level, at that estimated break-point. As such, because our key dependent variables are never used in the estimation of the break-point, it is unlikely that there is a hard-wired relationship.

Another important difference between the Card et al. set-up and ours is that in their context, it is not clear whether or not a discontinuity exists and there is a concern that they might “find” a discontinuity even if none were actually present. In our particular context though, since we know there is an admission threshold (e.g. a score below which no one can gain admission because all seats in the state are filled in), this problem is abated. A more relevant problem in our context is that there might be other discontinuities for each caste category in addition to the admission threshold. Since higher-ranked individuals can choose which school (and major) they will attend, there could, in theory, be discrete jumps in the attendance variable whenever a given school (or major) fills up and lower-ranked admits are forced to consider a lower quality school (or less attractive major). In practice though, it is unlikely that the technique described above would pick such school-specific or major-specific admission thresholds over the arguably larger jump that should occur when all colleges and majors are filled in. This is confirmed in Fig. 2, which graphs, for each caste category, the percentage of individuals who actually enrolled as a function of their test score, as well as the aforementioned estimated score cutoffs (represented by a vertical line).[7] As one can see from Fig. 2, engineering college attendance is virtually 0 for those individuals that scored below the estimated cutoffs. Note, however, that it is not completely zero, as by the time the offers have been made lower down on the list, the applicants might have already accepted an offer for another type of college, and that some applicants who are lower on the list may stake the admissions office waiting for someone to drop out at the last minute.

We selected a sample of applicants above and below their caste category cutoff. For the first round of surveys, we restricted ourselves to individuals who lived in one of the four more populous cities in terms of number of applicants. Conditioning on an address in one of these four cities, we then chose for the SC sample the 190 applicants right above the cutoff and 190 applicants right below the cutoff. This sample represented a higher percentage of the total number of SC applicants than the 190 applicants right above and right below the cutoff for the other groups. We therefore calculated what percentage of the full list of SC applications our sample comprised, and determined ranges of test scores that corresponded to the same percentage for the other two caste groups. The OBC and upper-caste samples were then constructed by randomly choosing individuals above and below the cutoff within these test score ranges. Thus, we made a choice to not obtain more individuals closer to the cutoff for the upper caste and other backward castes in order to obtain more comparable samples across caste categories. In total, we searched for 1292 households across the three caste categories in the first wave of surveying that took place between October 2004 and May 2005. Between May and July 2006, a second wave of surveys was conducted in the next four more populous cities in terms of number of applicants. The research samples were constructed using the same methodology as for the four more populous cities. In this second wave of surveying, we searched for an additional 692 households. Therefore, in total, we searched for 1984 households. Overall, 663 parents and 407 applicants were both found and agreed to participate in the survey, for a total of 721 households.[8] The individuals in the upper-caste category who were located and agreed to participate span the 56th to 100th percentile of the full test score distribution for their caste group. The equivalent figures for the OBC and SC are 56th to 99th percentile and 44th to 88th percentile, respectively.

3.3. Survey design and outcomes

The survey consisted of two parts. The first part was directed at the parents of the applicant, while the second part was directed at the applicants themselves. Many questions, particularly parental background and applicant employment outcomes, were similar on both the parent and applicant surveys, so that if the applicant could not be interviewed, we could still collect their basic outcomes data from surveying their parents. When both the parent and the applicant could be located, we continued to ask each of them both sets of questions to be able to assess the accuracy of the parent's responses on applicant's educational and employment history.

The data we collected in the survey covered various aspects of the policy debate regarding affirmative action. First, we obtained

[7] The test score distribution for each caste group was broken into 50 bins of equal size; we then graph the percentage of individuals who enrolled in each bin.

[8] An additional 172 families were located, but refused to be part of the survey. In most cases, when the individual could not be located, the family had moved and the neighbors did not have contact address. In a few cases, the address did not exist, the family never lived at the address, or the individual in question passed away.

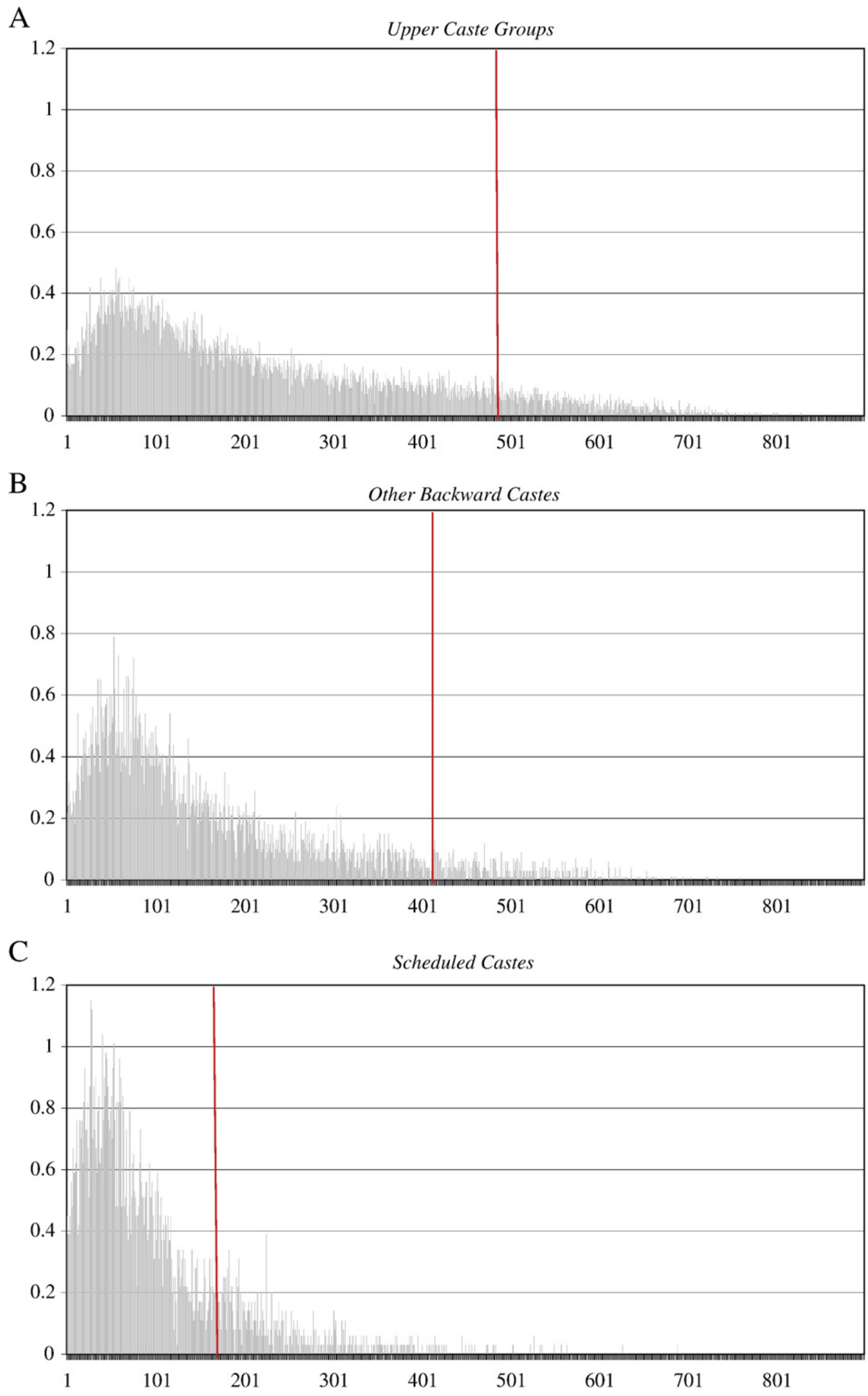

Fig. 1. Score on admissions Exam in 1996, by category. Notes: 1. The figures provide the distribution of test scores for applicants to engineering colleges in 1996, by caste group.

information on the background of parents to allow us to determine the social standing of those aided by the reservations. Second, we documented the educational and employment histories of the applicants themselves to determine the economic consequences of attending an engineering college.

Considerable effort was spent to locate as many households as possible. The enumerators first visited the parents' recorded address as of 1996 to determine if the parents still lived there. If the parents had moved, the survey team looked up the new contact information in the phone book (for within-city moves) or asked the neighbors for the new contact information. The survey was primarily conducted in person, but if the applicant or parents had moved out of the city, the survey was conducted by phone.

Appendix Table A1 compares located households with households that could not be located, by cutoff status and caste category. First, as Panel A shows, our ability to locate a given household was about the same regardless of whether or not that household was above or below the score cutoff; this is true for all three caste groups. In the remaining panels of Appendix Table A1, we compare basic characteristics (as available in the entrance exam data) between located and non-

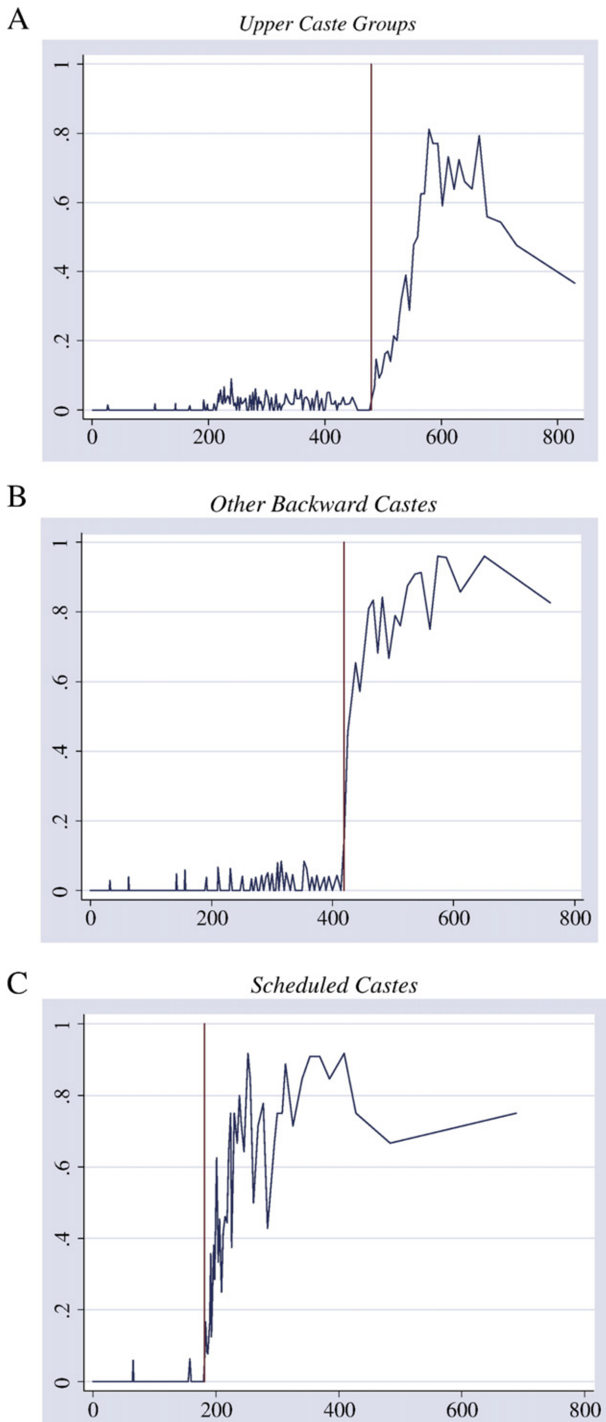

Fig. 2. Percentage of applicants who attended Engineering University in 1996, by score and category. Notes: 1. The figures graph the percentage of applicants who attended engineering college in 1996 by test score for each caste group.

located households. Column 1 presents the difference (between located and non-located households) in the characteristic of interest for those above the threshold, while column 2 presents the same difference for those below the threshold. Column 3 presents the differences between columns 1 and 2, with standard errors reported in parentheses below. While there are slight differences in test score, gender, age, and likelihood of attending engineering college between located and non-located households, these differences are not systematically related to the score cutoffs.

When only the parents were found, we used data from the parents' survey to supplement the applicant's outcome information, and vice versa. To confidently use this method, we would expect that parents and applicants would be able to accurately answer questions regarding one another. Appendix Table A2 confirms that they do, showing high correlations between answers given by the applicants and his/her parent for key questions, such as father's income, applicant's employment status, and applicant's income.

4. Does affirmative action target the poor?

We first study the question of whom affirmative action targets. Affirmative action policies can be instituted for a variety of reasons. First, governments may decide that for reasons having to do with social advancement, diversity or social harmony, it may be preferential to provide resources to previously excluded minorities. Second, governments may intend to target poor households, and choose to use race or minority status as a proxy for income.[9] In many developing countries, it is difficult to cost-effectively ascertain income levels because much of the economy is non-formal, i.e. there exists no clear system of paperwork or computer systems in place that would easily provide income measures to the government. As such, targeting traditionally disadvantaged caste groups may provide an alternative way for the government to successfully identify poor families, while at the same time preventing the behavioral distortions that one may see from targeting methods that rely on household characteristics that are associated with low socio-economic status (i.e. assets, whether children are in school or have access to health care, and so forth). In India, for example, the government adopted caste-based targeting specifically because they believed that income-based targeting would be too expensive and unwieldy, and that the caste-based targeting would result in the desired income-based targeting (de Zwart, 2000; Wolpert, 2006). In fact, much of the current debate in India on whether affirmative action should be continued revolves around the targeting question: many are afraid that only rich minorities are being helped by the program and, even worse, that rich minorities are displacing poor members of the upper caste.

To better understand who the families that benefited from the program are, we start by describing the socio-economic background of our survey respondents, by caste group (Table 1A). For variables on family background, data from the parent survey were used when available; when the parent survey data were missing, data from the applicant survey were used. For variables containing personal information on the applicant, data primarily come from the applicant survey; for cases in which the applicant was not found, the variable takes on the value of the equivalent question in the parent survey. Column 1 reports means for the entire sample; columns 2, 3, and 4 report means for upper-caste, OBC and SC groups, respectively. The *p*-values for comparisons of means across groups are reported in the remaining columns of Table 1A. Panel A focuses on characteristics of the main income earner in each applicant's household (as of 1996), while Panel B provides some background information about the applicants themselves. One star indicates significance at the 10% level, two stars indicate significance at the 5% level, and three stars indicate significance at the 1% level.

Among the applicants in the survey, the upper caste group tends to be better off than the lower-caste groups in terms of socioeconomic characteristics. The upper-caste group had an average monthly income of Rs 12,790 (column 2) as compared to Rs 8947 for the OBC group and Rs 8081 for the SC group (columns 3 and 4,

[9] See, for example, Ravallion (2009), for a general description of the rationale and debate on the use of proxies for income in targeting the poor. See, for example, Leonard (1985), for a discussion of affirmative action, specifically, as a tool for redistribution.

Table 1

A. Comparing pre-characteristics across groups

	Means				Comparison of means (*p*-value)		
	Full sample	Upper caste	OBC	SC	UC vs OBC	UC vs SC	OBC vs SC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Characteristics of main income earner in family in 1996							
Below college degree	0.36	0.21	0.43	0.53	0.00***	0.00***	0.08*
Above master's degree	0.21	0.30	0.16	0.13	0.00***	0.00***	0.46
Employed	0.97	0.97	0.97	0.98	0.83	0.49	0.65
Income	10324.77	12790.86	8947.84	8081.25	0.00***	0.00***	0.26
Log (income)	9.02	9.26	8.90	8.78	0.00***	0.00***	0.08*
Engineer	0.13	0.13	0.14	0.08	0.51	0.27	0.08*
Uses computer in job	0.11	0.14	0.10	0.07	0.23	0.03**	0.31
Number of other children	2.69	2.33	2.76	3.21	0.00***	0.00***	0.00***
Characteristics of engineering applicant							
Attended an English private school	0.50	0.63	0.45	0.32	0.00***	0.00***	0.01**
Attended a government school	0.17	0.11	0.14	0.28	0.44	0.00***	0.00***
Age	18.51	18.20	18.52	18.98	0.00***	0.00***	0.00***
Male	0.84	0.80	0.89	0.86	0.00***	0.06**	0.32

B: Comparing pre-characteristics across groups in the NSS Round 55

	Means				Comparison of means (*p*-value)		
	Full sample	Upper caste	OBC	SC	UC vs OBC	UC vs SC	OBC vs SC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HH below college degree	0.80	0.70	0.89	0.94	0.00***	0.00***	0.00***
HH works regularly	0.81	0.81	0.81	0.79	0.96	0.26	0.28
Monthly consumption expenditures	3452.04	4009.11	2995.30	2491.95	0.00***	0.00***	0.00***
HH engineer	0.01	0.02	0.01	0.00	0.01**	0.02**	0.48
N	2911	1518	977	416	1393	977	1393

Notes:
1. Panel A reports the mean background characteristics of the applicants and their families in 1996.
2. Panel B reports the mean background characteristics of a random sample of households who live in urban areas of the state of India where our study occurs. The data for this table comes from the National Sample Survey of India, 55th Round.
3. Column 1 presents the mean for the full sample, while columns 2–3 report the means by caste groups. Columns 5–7 present *p*-values for the difference in means between each of the indicated groups.
4. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

respectively). The head of the household in the upper caste households tended to have more technically advanced positions than those from lower-caste households. For example, 14% of the heads from upper caste households had used a computer on the job and 13% were engineers. In contrast, only 7% of the heads from SC households used a computer, and only 8% were engineers. The lower family income translates to fewer opportunities for the lower-caste applicants themselves: for example, they are less likely to have attended an English-language private school (which tends to be the highest quality schools) and more likely to have attended a typically lower quality, government school. It is interesting to note that while the OBC applicants are better off on average than the SC applicants, the "economic distance" between these two groups is smaller than the distance between the average upper-caste and average lower-caste applicants. OBC and SC households do not appear significantly different than each other in terms of income, propensity of the household head to have a master's degree and propensity of the household head to be employed (Table 1A, column 7).

Of course, the finding that the applicants from upper class households tend to be better off than those from lower-caste households does not imply that the lower-caste applicants are poor per se. One argument against affirmative action is that the low caste applicants may come from households that are relatively richer than the "typical" lower-caste households. To provide evidence on this topic, we present in Table 1B socio-economic information for households living in urban areas in the Indian state where our research takes place. This table was constructed from the 55th round of the National Sample Survey of India (NSS) that was conducted in 1999–2000. Unfortunately, the NSS does not contain all background variables that we collected in our survey. Thus, Table 1B lists the following variables from the NSS that are most equivalent to the data we collected: a dummy variable for whether the head of the household's education was below a graduate degree, a dummy variable for whether the head of household worked regularly, monthly consumption expenditures (in Rs 1999), and a dummy for whether the head of the household works as an engineer. The column structure of Table 1B is identical to that of Table 1A.

Overall, the applicants in our survey are positively selected compared to other individuals in the state. While 80% of household heads in the NSS did not complete a college degree, this fraction is only 36% among applicants' households (column 1). Applicants to engineering colleges are also more likely to be part of a household in which the head is employed. Finally, there is strong inter-generational correlation in career choice: while only about 1% of the households in the NSS are headed by engineers, 13% of the applicants to engineering colleges originate from households in which the head was an engineer.

A further comparison of the information in Table 1A and B shows that the positive selection among applicants to engineering colleges occurs across all three caste groups. While 89% of the OBC and 94% of the SC in the urban areas of the state we consider have less than a college education (columns 3 and 4 in Table 1B), these figures are respectively 43% and 53% and for the applicants (columns 3 and 4 in Table 1A). Similarly, the monthly consumption rates for the average OBC and SC households (Table 1B) is about a fourth of the monthly income of the applicants in our sample (Table 1A). Taken together, these findings suggest that the minority households that have the ability to take advantage of the affirmative action programs are much better off than the typical minority household. Importantly, though, as we discussed earlier, the positive selection among OBC and SC applicants is not so severe that their background characteristics become similar to those of the upper-caste applicants. Hence, despite minority applicants being positively selected from their group, caste-based targeting still appears to be redistributing resources to relatively *poorer* families, even if it does not redistribute to the *poorest* families.

To show this more formally, we next compare the socio-economic characteristics of the applicants offered a seat in an engineering college in 1996 due to the reservation program to the characteristics of those who were refused a seat but would have been admitted in the absence of the program. To perform this exercise, we make several assumptions. First, we assume that the number of applicants in each caste category, as well as their exam score, would be unaffected by the removal of affirmative action. Second, we assume that the total number of available seats in engineering colleges would remain the same in the absence of affirmative action. Finally, we also need to make an assumption about what the enrollment rate would be among those who would be offered a seat in the absence of affirmative action. The exam data show that the average enrollment rate is about 50% for the upper-caste group. However, admissions are made on a rolling basis and the rate is closer to 60 to 70% for those who were informed earlier of their acceptance. Therefore, we consider two different scenarios: 50 and 70% enrollment rates.

With these assumptions, we can calculate the score on the entry exam that would have been the threshold for admission in the absence of the reservation policy—that is, in an environment in which all applicants are ranked on the same list. This threshold score is, by construction, higher when we assume a 70% enrollment rate than when we assume a 50% rate. Using the respondents' score on the entry exam, we can identify the individuals who would have been admitted in the absence of the reservation policy and those who would not have been admitted. We can then compare the socio-economic background of those who would have been admitted but were not (e.g. the “displaced”) and those who would not have been admitted but were (e.g. the “displacing”).

A first (unsurprising) finding is that if the government simply aimed to increase the number of minority candidates in engineering, they were very successful. In the sample, 49 to 59% of the low caste applicants that we considered (SC and OBC) that gained admissions to the engineering colleges did so as a result of affirmative action (based on the 50 to 70% level cutoff, respectively). Because our analysis excludes the Scheduled Tribe group, of which less than 1% would have been admitted in the absence of affirmative action, these figures understate the true increase in caste diversity achieved due to affirmative action.

We next turn to quantifying how much affirmative action succeeded in redistributing educational opportunities to poorer households, using monthly family income and other measures of wealth as our metrics (Table 2). We assume a 50% enrollment rate in columns 1 through 3, and a 70% rate in columns 4 through 6. Columns 1 and 4 report average background characteristics for displaced applicants, while columns 2 and 5 report average background characteristics for displacing applicants. We report the *p*-values for tests of comparison of means in columns 3 and 6.

Table 2 confirms that the reservation policy is associated with the admission of individuals of a lower socio-economic background. Under the assumption of a 70% enrollment rate, mean parental income among the displaced individuals is Rs 14,088 per month (column 1) compared to Rs 8340 (column 2) among the displacing individuals. Similarly, 41% of displaced individuals come from a household in which the head holds at least a master's degree, compared to only 14% of displacing individuals. In addition to coming from poorer households, displacing minority students come from households that offered fewer educational opportunities at the high school level: 59% of displaced individuals attended an English-language private school, compared to only 35% of displacing individuals. Not surprisingly, the economic magnitude of the difference between displacing and displaced individuals is smaller under the assumption of a 50% rate than under the assumption of a 70% rate, as the former assumption implies a lower threshold for admission and, hence, draws fewer of the displaced students from the (economically better off) general-caste category. Even under the 50% assumption, there is no evidence that the caste reservation policy discriminates against the economically weak, as some opponents of this policy have argued.

However, targeting by minority status may have other unforeseen social impacts. In particular, caste-based targeting appears to reduce gender diversity in the engineering colleges. This pattern emerges regardless of the enrollment rate assumed, but is only statistically significant under the assumption of a 50% rate. Under the assumption of the 50% yield, 73% of those “displaced” by the reservation policy are males, compared to 84% of those “displacing.” This finding is not surprising given that the share of male applicants (80%) is lower in the upper-caste group than in the OBC and SC groups (89% and 86% respectively) in our sample. More generally, individuals from lower-caste groups that complete secondary school in India disproportionately tend to be male (Henriques and Wankhede, 1985). Thus, it follows that the college applicants from these lower-caste groups would also disproportionately tend to be male. This suggests that the targeting of lower-caste groups, on the whole, may exclude women from opportunities in higher education.

In summary, while caste-based targeting did not benefit the poorest members of society, it did appear to be successful in reallocating resources towards relatively less well-off households. Yet at the same time as it substantially increased diversity by caste and income groups, caste-based targeting reduced diversity by gender in the engineering colleges.

5. Labor market outcomes

5.1. Descriptive evidence

Appendix Table A3 summarizes labor market outcomes by caste category across all surveyed applicants, regardless of whether they were admitted or attended the engineering colleges. At the time of the survey, the lower-caste applicants, and especially SC applicants, are on average economically worse off compared to the upper-caste applicants. Of more indirect interest to us, however, is the question of whether lower-caste applicants benefit from gaining access to the engineering colleges, and if so, how this compares to the return for upper-caste applicants.

Table 3 offers descriptive evidence on this topic by reporting, separately for upper- and lower-caste groups, the OLS estimates of the relationship between the applicants' monthly income at the survey time and attending an engineering college in 1996. Income is defined as gross monthly income if employed, and zero otherwise. We report the results of two regression models: one in which we only control for a vector of individual characteristics (column 1) and one in which we additionally control for household characteristics and city of origin (column 2). The individual-level controls are dummies for OBC and SC categories, a gender dummy, the logarithm of age, and dummies for the type of secondary school attended (e.g. English private, Hindi private, or government school). The household-level controls are monthly income as of 1996 (in Rs), dummies for the educational attainment of the household head, a dummy for whether the household head is an engineer, a dummy for whether the household head uses a computer at work, and the number of other children in the household. We also control for whether the data come from the parent or applicant survey in all regressions.

The OLS estimates in Table 3 suggest positive and significant income effects associated with attending engineering college, even

Table 2
Comparing the displaced with the displacing.

	50% Yield			70% Yield		
	Mean for displaced	Mean for displacing	p-Value of difference in means	Mean for displaced	Mean for displacing	p-Value of difference in means
	(1)	(2)	(3)	(4)	(5)	(6)
A. Characteristics of main income earner in family in 1996						
Below college degree	0.37	0.48	0.19	0.24	0.47	0.07
Above master's degree	0.23	0.13	0.12	0.41	0.14	0.01***
Employed	0.97	0.98	0.73	1.00	0.97	0.47
Income	11098.28	7908.05	0.01**	14088.24	8340.00	0.00***
Log (income)	9.07	8.82	0.02**	9.33	8.87	0.00***
Engineer	0.08	0.07	0.92	0.11	0.10	0.92
Uses computer in job	0.05	0.06	0.85	0.06	0.09	0.64
Number of other chlidren	2.49	3.18	0.01***	1.76	3.16	0.00***
B. Characteristics of engineering applicant						
Attended an English private school	0.54	0.33	0.01**	0.59	0.35	0.07*
Attended a government school	0.07	0.28	0.00***	0.00	0.27	0.02**
Age	18.25	18.76	0.02**	18.16	18.73	0.11
Male	0.73	0.84	0.10*	0.74	0.87	0.15

Notes:
1. This table compares mean background characteristics of individuals displaced by affirmative action policies ("displaced") and with the mean for those who gained admissions due to affirmative action policies ("displacing").
2. Columns 1–3 assume a 50% rate of acceptance of an admissions offer, providing a total sample size of 159. Column 1 presents the mean for the displaced, while column 2 reports the mean for the displacing. Column 3 presents the p-value of the difference in means. Columns 4–6 replicate columns 1–3, assuming a 70% yield ($N = 132$).
3. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

among the lower-caste groups. This contradicts the extreme view that seats in engineering colleges are "wasted" on lower-caste candidates. It also appears that attending engineering college is associated with larger average income gains for the upper caste compared to the lower-caste groups (Rs. 5400 vs Rs 3200 in column 1; Rs 3700 vs 2700 in column 2). However, because of large standard errors, we cannot reject the hypothesis of equal gains across the two caste categories.

Table 3
OLS regressions — labor market outcomes.

Outcome:	Income	Income
	(1)	(2)
Panel A: upper caste groups		
Went	5337.69 (2012.20)***	3685.44 (2066.57)*
Observations	273	273
R-squared	0.06	0.18
Panel B: lower caste groups(SC and OBC)		
Went	3118.64 (1133.88)***	2717.76 (1105.12)**
Observations	380	379
R-squared	0.17	0.28
Individual controls	Yes	Yes
Household controls		Yes
City fixed effects		Yes

Notes:
1. Each coefficient contains the result of a separate regression in which the dependent variable is income. The reported coefficient is on "went to engineering college."
2. The sample in Panel A includes upper-caste applicants, while the sample in Panel B includes lower-caste applicants.
3. The individual-level controls are dummies for OBC and SC categories, a gender dummy, the logarithm of age, dummies for the type of secondary school attended (e.g. English private school, Hindi private school or government school), and controls for whether or not the outcome data comes from the parent or child survey. The household-level controls are household head monthly income as of 1996 (in Rs), dummies for the educational attainment of the household head, a dummy variable for whether the household head is an engineer, a dummy for whether or not the household head uses a computer at work, and the number of other children in the household.
4. Standard errors in parentheses. The number of observations and the R^2 of the regression are listed below the standard errors.
5. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

5.2. Empirical strategy

The OLS results reported above are somewhat difficult to interpret. In the OLS model, part of the variation in who attended engineering college is driven by the variation in who scored above and below the admission threshold in their caste group, while another part of the variation is driven by one's decision of whether to attend, *conditional* on having been admitted. This second source of variation is clearly problematic when one tries to measure the causal effect of attending engineering college. First, suppose that only those applicants who saw the greatest return to attending did so. In this case, we would clearly overestimate the effects of attending engineering college. Second, the decision of whether or not to attend, conditional on having been admitted, may be driven by liquidity constraints at the household level. Third, the selection into attending, conditional on having been admitted, may also reflect variation in outside options. Suppose that a seat in medical school is preferred to a seat in engineering. Then, only those who did not make it into medical school will attend an engineering school. Hence, we would have included among those who did not attend engineering school some individuals who went instead to medical school, leading us to underestimate the effect of attending engineering school. Furthermore, if there is systematic variation in the quality of these outside options across caste categories, this would invalidate any comparison of the OLS returns to attending engineering college across caste categories. For example, if the upper-caste applicants systematically have better outside options than the lower-caste applicants, the OLS model might give us an underestimate of the difference in the returns between the two groups.

We propose alternative estimates of the returns to attending engineering college that only rely on the first source of variation listed above: variation in who scored above and below the admission thresholds on the entry exam. Specifically, in our basic IV model, we replicate the specifications in columns (1) and (2) of Table 3, but use the score-cutoff dummies as instruments for attending engineering college.

One remaining issue, of course, is that there might be systematic unobservable differences between those who scored above and below the admission thresholds that would induce differences in their current labor–market outcomes, even in the absence of an engineering education. While this is a reasonable concern, it is important to

Table 4
Comparing pre-characteristics for applicants above and below threshold, by category.

Sample: pre-characteristic	Upper caste		Lower caste	
	Full	Discontinuity	Full	Discontinuity
	(1)	(2)	(3)	(4)
A. Characteristics of main income earner in family in 1996				
Below college degree	−0.062 (0.059)	−0.077 (0.084)	−0.083 (0.052)	−0.082 (0.074)
Above master's degree	0.015 (0.056)	−0.06 (0.091)	0.019 (0.038)	−0.046 (0.056)
Employed	−0.01 (0.022)	−0.018 (0.022)	−0.014 (0.017)	0.003 (0.024)
Income	708.857 (1359.592)	842.545 (2258.617)	724.411 (769.800)	−584.429 (1004.810)
Engineer	0.023 (0.041)	0.006 (0.118)	0.024 (0.031)	0.022 (0.040)
Uses computer in job	0.058 (0.043)	0.084 (0.058)	0.027 (0.030)	0.016 (0.044)
Number of other children	−0.134 (0.130)	−0.048 (0.204)	−0.12 (0.152)	0.067 (0.222)
B. Characteristics of engineering applicant				
Attended an English private school	0.058 (0.060)	0.102 (0.092)	0.052 (0.051)	0.033 (0.075)
Attended a government school	0.021 (0.040)	0.084 (0.057)	−0.006 (0.042)	−0.057 (0.059)
Age	−0.079 (0.115)	−0.054 (0.202)	−0.087 (0.135)	0.023 (0.185)
Male	0.084 (0.047)*	0.126 (0.073)	0.038 (0.033)	0.002 (0.047)
Chi-squared	10.79	15.33	7.85	8.22
P-value	0.37	0.12	0.55	0.61

Notes:
1. Each cell contains the result of a separate regression in which the dependent variable is the pre-characteristic as indicated. The reported coefficient is on a dummy variable for "above cutoff." Standard errors are listed in parentheses below the coefficient.
2. The sample in Column 1 is the full sample of upper-caste groups. The sample in Column 2 is the discontinuity sample for upper-caste groups where applicants with the top 25% test scores and the bottom 25% test scores are excluded from the sample. The remaining columns replicate columns 1–2 for pooled lower-caste groups (columns 3–4).
3. The Chi-squared and *p*-value of a SUR regression for joint significance are listed on the last two rows. Uses computer in job are excluded in order to be able to run the SUR model.
4. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

stress that we attempted to form a homogeneous group of survey respondents, given our survey constraints. Our sample is limited to a group of individuals that not only passed the first-round $10+2$ exam, but further scored in the upper half of the second-round entrance exam. Columns 1 and 3 of Table 4 statistically confirm the balance of observable background characteristics for survey respondents above and below the admission threshold within each caste category. Each cell in Table 4 corresponds to a different regression where the dependent variable is the background characteristic in that row and the independent variable is the relevant cutoff dummy for the group in that column; reported in the cell is the estimated coefficient on the cutoff dummy.[10] There are only very few cases for which we find statistically significant differences between those applicants who are above the cutoff and those who are below, even though standard errors are admittedly sometimes large enough to prevent us from ruling out economically meaningful differences.

Of course, we know that there are meaningful differences in the test score variable for respondents that are below and above the admission threshold. It is reasonable to assume that one's test score would be a predictor of future earnings, even in the absence of an engineering education. The natural approach to deal with this concern would be to focus as much as possible on the *marginal* admit within each caste category. We do this in two ways. Our first strategy consists in constructing for each caste category a "discontinuity sample" in which we drop those survey respondents who scored below the 25th percentile or above the 75th percentile of the score distribution of all those surveyed in their caste category. Thus, the test scores in the general-caste discontinuity sample range from the 85th to the 97th percentile of the distribution of test scores among general-caste entry exam-takers in 1996; the equivalent ranges are 84th to 96th percentile and 77th to 92nd percentile for the OBC and SC samples, respectively. As expected, the differences in pre-characteristics between those above and below the admission threshold are in general even smaller in these discontinuity samples than in the full samples (see columns 2 and 4 of Table 4). As a second strategy, we also estimate on the full sample variants of the basic IV model described above where we include a quadratic polynomial in test score.

Note that our proposed methodology for measuring the marginal effects of affirmative action (using the discontinuity samples and control functions in score) necessarily discards important information. While it provides more causally credible estimates, it focuses on sub-samples of the populations of interest. The groups of applicants who place in (lower caste) and place out (upper caste) of engineering college due to the reservation policy come from a much broader range of test scores than the marginal test takers the methodology focuses on. Relatedly, this focus also means we can only learn about the causal impact of being admitted to the lowest quality engineering colleges (and/or less prized majors within engineering) vs not being admitted at all. The affirmative action policy as a whole, however, contributes to placing lower-caste applicants into higher quality colleges and majors, and conversely moves upper-caste applicants out of higher quality colleges and majors. This discussion suggests a trade-off between studying average returns to admission in broader, but still rather homogeneous and more policy-relevant, samples of applicants (as we do in our basic IV model) vs attempting a more causal measurement of the returns to attending engineering college on admittedly less policy-relevant samples.

6. Results

The results are present in Table 5. Panel A shows the results from the first-stage regressions, while Panel B provides second-stage results. In light of the discussion above, we consider four different models. Models 1 and 2 use the full samples of survey respondents within a given caste category. Model 1 simply consists of regressing "went" on a dummy for the group-specific cutoff (as well as a dummy for the OBC category for the lower-caste group), while Model 2 includes a vector of individual controls, household background controls, and city of origin fixed effects. Model 3 replicates Model 2 (e.g. including the full list of controls) on the

[10] When we pool both lower-caste groups together in Table 4, we also include in each regression a dummy for OBC status. Results for the individual minority groups look very similar, and are available upon request from the authors.

Table 5
Returns to an Engineering Degree.

	Upper caste groups				Lower caste groups			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: first stage								
Cutoff	0.51 (0.05)***	0.53 (0.05)***	0.43 (0.09)***	0.17 (0.10)*	0.68 (0.03)***	0.68 (0.03)***	0.54 (0.05)***	0.48 (0.06)***
Observations	308	308	154	308	413	412	209	413
R-squared	0.27	0.35	0.24	0.34	0.52	0.54	0.48	0.57
B. IV analysis of impact of attending engineering college on income								
Went	13,002.05 (3942.70)***	8972.98 (3902.04)**	12,256.5 (8533.27)	19,961.81 (29,927.45)	6202.71 (1606.85)***	5549.22 (1568.88)***	4071.75 (2690.64)	3571.95 (4180.91)
Observations	273	273	138	273	380	380	193	380
Sample	Full	Full	Disc	Full	Full	Full	Disc	Full
Individual controls		Yes	Yes			Yes	Yes	
Household controls		Yes	Yes			Yes	Yes	
City fixed effects		Yes	Yes			Yes	Yes	
Quadratic score function				Yes				Yes

Notes:
1. In Panel A, each coefficient contains the result of a separate regression in which the dependent variable is "went to engineering college." The reported coefficient is on "above cutoff for admissions." In Panel B, each coefficient contains the result of a separate IV regression in which the dependent variable is either income. The reported coefficient is on "went to engineering college," which is instrumented by "above cutoff for admissions."
2. In columns 1–4, the sample includes upper-caste applicants, while the sample in columns 5–8 includes lower-caste applicants.
3. The individual-level controls are dummies for OBC and SC categories, a gender dummy, the logarithm of age, dummies for the type of secondary school attended (e.g. English private school, Hindi private school or government school), and controls for whether or not the outcome data comes from the parent or child survey. The household-level controls are household head monthly income as of 1996 (in Rs), dummies for the educational attainment of the household head, a dummy variable for whether the household head is an engineer, a dummy for whether or not the household head uses a computer at work, and the number of other children in the household.
4. In the discontinuity sample, applicants with the top 25% test scores and the bottom 25% test scores are excluded from the sample.
5. Standard errors in parentheses.
6. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

discontinuity samples. Finally, in Model 4, we control for the quadratic in the entrance exam test score.[11]

The first-stage regressions reported in Table 5 statistically confirm the association between the likelihood of attending engineering college in 1996 and scoring above the cutoff on the entry exam. However, we observe a sharp drop in the coefficient of the cutoff when controlling for the quadratic in score, particularly for the upper caste group. In all models, we observe a smaller estimate of the cutoff dummy for the upper-caste applicants compared to the lower caste applicants. This indicates that a smaller fraction of upper-caste admits attend engineering college. This could reflect the fact that the general caste admits face a larger set of educational choices.

All of the second-stage point estimates (Panel B) show positive returns to attending engineering college, and tend to be larger than the OLS estimates presented in Table 3. In Models 1 and 2, attending engineering college increases the income of upper-caste individuals by between Rs 9000 and Rs 13,000 (statistically significant); the equivalent figures for lower caste applicants are Rs 5500 and Rs 6500, respectively (both statistically significant). Hence, like in the OLS model, we find no evidence for the extreme view that engineering education is being wasted on lower caste applicants but we also find larger average gains of engineering college attendance for the upper-caste group, consistent with some aggregate economic cost of the reshuffling of seats that is imposed by the reservation policy.

Our two attempts to obtain a purer estimate of the causal effect of engineering education by focusing on the marginal admits (Models 3 and 4) are inconclusive due to large standard errors. Note that despite the large standard errors, the point estimates stay broadly consistent with those of Models 1 and 2.

[11] In a fifth model (not reported here), we include both the quadratic function in test score and the individual, household and city-level controls. The results are qualitatively similar, and obtainable from the authors upon request.

Table 6 replicates the analysis of Table 5, but focuses on job characteristics other than income. For each job characteristic (as listed in the first column), we report second-stage estimates of the impact of attending engineering college (e.g. each cell corresponds to a separate regression). As in Table 5, our attempts to provide estimates for the marginal admits are not very successful (this is particularly true for Model 4). Focusing the discussion on the average differences estimated under Models 1 and 2, we find that attending engineering college has a strictly positive, but smaller, impact on one's likelihood of working as an engineer for lower caste compared to upper-caste individuals (row 1 in Panels A and B, respectively). Both lower- and upper-caste members see their likelihood of being self-employed or working in a family business decline by about the same amount (10 percentage points) after attending engineering college. Most of the extra employment generated among general-caste members is in the private sector. For lower caste members, the picture is less clear, but suggestive of less growth in private sector employment and perhaps more growth in public sector employment. Hence, we cannot rule out the view that part of the labor market gains associated with attending engineering college for lower caste candidates might be linked to other affirmative action programs in the labor market.

6.1. Does family background affect the estimated returns?

In Table 7, we investigate whether there is heterogeneity by socio-economic background in the caste-specific income effects we estimated above. To proceed, we first summarize a given individual's socio-economic background characteristics into a single index. We do this by regressing a given individual's test score on the vector of individual and household background variables introduced above; we then assign each individual a dummy variable ("disadvantaged") that equals 1 if the individual's predicted test score is below the predicted median *within* his or her caste category and 0 otherwise. We then replicate the IV regressions estimated in Models 1, 2, and 4 of Table 5,

Table 6
Job characteristics (2SLS).

	Model 1	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)
Panel A: upper caste groups				
Works as engineer	0.44	0.35	0.37	0.1
	(0.12)***	(0.12)***	(0.26)	(0.72)
Works in private sector	0.39	0.27	0.28	0.83
	(0.12)***	(0.11)**	(0.23)	(0.79)
Works in government sector	−0.01	0.07	0.01	−0.24
	(0.07)	(0.07)	(0.14)	(0.43)
Self-employed/ family-employed	−0.11	−0.11	−0.15	−0.12
	(0.04)**	(0.04)**	(0.08)*	(0.26)
Panel B: lower caste groups (OBC and SC)				
Works as engineer	0.27	0.22	0.22	0.22
	(0.07)***	(0.07)***	(0.13)*	(0.19)
Works in private sector	0.04	−0.01	0.08	0.13
	(0.07)	(0.07)	(0.13)	(0.19)
Works in government sector	0.07 (0.05)	0.08	0.04	0.29
		(0.06)	(0.10)	(0.15)*
Self-employed/ family-employed	−0.09	−0.09	−0.12	−0.12
	(0.04)**	(0.04)**	(0.07)*	(0.10)
Sample	Full	Full	Disc	Full
Individual controls		Yes	Yes	
Household controls		Yes	Yes	
City fixed effects		Yes	Yes	
Quadratic score function				Yes

Notes:
1. Each coefficient contains the result of a separate IV regression in which the dependent variable is as indicated. The reported coefficient is on "went to engineering college," which is instrumented by "above cutoff for admissions."
2. In Panel A, the sample includes upper-caste applicants, while the sample in Panel B includes lower-caste applicants.
3. See Table 5 for a discussion of the included control variables.
4. In the discontinuity sample, applicants with the top 25% test scores and the bottom 25% test scores are excluded from the sample.
5. Standard errors in parentheses.
6. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

but add to the regressions an interaction term between going to engineering college and the "disadvantaged" dummy variable.[12]

The most striking finding is for lower caste applicants (Panel B). Well-off lower caste applicants benefit more from attending engineering college than those from more disadvantaged backgrounds. This pattern holds in all models, including the model with the quadratic control in test score (column 3). In fact, our estimates indicate very small returns to attending engineering school among those lower caste applicants who come from poorer backgrounds. This pattern does not extend to the upper-caste applicants (Panel A).

This last finding somewhat darkens the picture of the reservation policy with regard to its benefits to lower-income groups (and not just lower caste groups). Specifically, while we showed that the targeting of the lower caste groups does lead to the admission of students of a lower socio-economic background, Table 7 suggests that the reservation policy may provide benefits only to those who are already economically better off within the lower caste groups. We can only hypothesize about the possible channels for this. For example, higher SES lower caste members may find it easier to make it through college. Also, higher SES lower-caste members may find the post-graduation job market easier to navigate, or have the social networks needed to take better advantage of their degree.

[12] We of course also include the "disadvantaged" dummy as a direct control. The estimation of the 2SLS models now relies on two instruments: the score threshold dummy and the score threshold dummy interacted with the "below median predicted score" dummy. Note that, we do not replicate Model 3 (the discontinuity sample) because we drop half the observations in this model, and therefore, it becomes difficult to do within group comparisons for each caste group.

Table 7
IV estimates of returns, by socioeconomic group and caste.

	Model 1	Model 2	Model 4
	(1)	(2)	(3)
Panel A: upper caste groups			
Went	6929.28	4890.88	11,924.76
	(5466.10)	(5344.92)	(32,049.23)
Disadvantaged	−8035.22	−8102.64	−8264.22
	(3295.76)**	(4316.32)*	(3604.12)**
Disadvantaged* went	9332.11	11,008.92	8534.02
	(7617.46)	(7583.77)	(9861.07)
Panel B: lower caste groups (OBC and SC)			
Went	9426.9	9581.81	5923.17
	(2188.89)***	(2146.69)***	(4201.15)
Disadvantaged	639.59	3835.14	243.24
	(1878.64)	(2337.27)	(1925.10)
Disadvantaged* went	−9137.85	−8918.19	−8355.83
	(3249.55)***	(3245.04)***	(3392.34)**
Sample	Full	Full	Full
Controls		Yes	
Qudratic score function			Yes
Mean score college			
GPA			

Notes:
1. Each column in a panel presents the results of a separate IV regression in which the dependent variable is income. Disadvantaged is a dummy variable which indicates whether an individual was below the predicted median score for his or her caste group.
2. In Panel A, the sample includes upper-caste applicants (273 observations), while the sample in Panel B includes lower-caste applicants (380 observations).
3. Standard errors in parentheses.
4. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

7. Conclusion

Our analysis confirms that affirmative action policies that target by social group can be a successful tool in both increasing diversity and allocating resources to relatively disadvantaged families. Contrary to the arguments of some critics, the affirmative action program under study here does not merely crowd out economically-disadvantaged upper-caste students to make way for economically-advantaged lower-caste students. The individuals who are displaced by the program come from stronger socio-economic backgrounds than the displacers. Hence, by targeting disadvantaged caste groups, the policy achieves some income targeting without generating any of the behavioral distortions typically associated with income targeting.

We find suggestive evidence that admits from lower castes earned returns from attending engineering college. However, we also identify some potential limitations. First, like other redistributive programs, our point estimates suggest that the affirmative action policy comes at an economic cost, as the general caste admits lose out. Note, however, that sample size considerations limit our ability to fully measure this tradeoff. Second, and most troublesome, our results suggest that among the lower-caste admits, it is those from stronger socio-economic backgrounds who benefit most from the reservation policy. This result somewhat weakens the case that the policy benefits the economically disadvantaged.

This study makes several key contributions. First, it contributes to our knowledge of the returns to higher education in developing countries, for which there is a small but growing literature. Second, by taking advantage of both a straightforward policy experiment and a comprehensive dataset, this study provides a detailed analysis on the targeting implications of affirmative action programs. Third, we provide suggestive evidence on the relative returns to engineering colleges for not only on the individuals who obtained admissions to higher education due to affirmative action, but also on those who were denied admissions as a result of affirmative action.

Our study also provides clear directions for future work. First, despite a large-scale data collection undertaking on our part, the strenuous data requirements of the regression discontinuity design methods coupled with our limited sample size reduced our ability to provide conclusive evidence on the returns to attending engineering school for the marginal admit. Thus, we see great value in the replication of this study in contexts where there could be a potentially larger sample size. Second, our analysis focuses on educational reservations in one state and one field (engineering). We see great value in the replication of this research in other educational fields and other regions: relationships between advantaged and disadvantaged groups vary greatly across regions, and the nature of the educational production function may also vary significantly across fields. Third, this study was conducted in a context in which affirmative action for minorities was also present in public sector employment, in addition to the affirmative action at the university level. It would be interesting to understand whether the returns would be similar in contexts where other affirmative action programs do not exist. Finally, our research suggests the economic and redistributive trade-offs that are involved in affirmative action. Future research should be prepared to further understand these social costs, and to compare affirmative action policies to other approaches with similar social objectives.

This project was funded by the National Science Foundation. We are grateful to Hardy K. Dewan for his assistance in obtaining the data, and IMRB for conducting the fieldwork. We thank Jessica Pan for outstanding research assistance. We thank seminar participants at Bocconi, Pompeu Fabra, the “Labor Markets in Transition and Developing Economies: Emerging Policy and Analytical Issues” conference (University of Michigan), the BREAD/CEPR conference (LSE), Case Western Reserve University, NYU and University of Houston/Rice for helpful comments.

Appendix A

Table A1
Differences in initial characteristics between individuals who were located and not located, by caste group and cutoff.

	Above	Below	Difference
	(1)	(2)	(3)
A. Percent found			
Upper castes	0.380	0.334	0.046 (0.033)
Other backward castes	0.329	0.375	−0.046 (0.039)
Scheduled castes	0.375	0.367	0.008 (0.042)
B. Difference in total score (found–not found)			
Upper castes	4.677	−1.973	6.65 (10.906)
Other backward castes	−0.733	5.998	−6.73 (11.372)
Scheduled castes	−1.030	−6.177	5.15 (7.877)
C. Difference in percent female (found–not found)			
Upper castes	0.037	0.068	−0.03 (0.054)
Other backward castes	−0.042	0.005	−0.047 (0.055)
Scheduled castes	0.006	−0.021	0.027 (0.064)
D. Difference in age (found–not found)			
Upper castes	0.043	0.007	0.036 (0.174)
Other backward castes	−0.614	−0.020	−0.593 (0.878)
Scheduled castes	0.018	0.246	−0.228 (0.320)
E. Difference in attended (found–not found)			
Upper castes	0.072	0.013	0.059 (0.058)
Other backward castes	0.039	0.027	0.013 (0.051)
Scheduled castes	−0.012	0.000	−0.012 (0.062)

Notes:
1. In Panel A, we report the percentage of individuals found above (column 1) and below (column 2) the cutoff for each caste group. The difference in these means is reported in column 3. The standard error is presented below the difference in parentheses.
2. In Panels B–E, we report the difference in means of pre-characteristics for those found and not found during the surveying. We report this difference for those above the cutoff in column 1 and for those below the cutoff in column 2. The difference in these means is reported in column 3. The standard error is presented below the difference in parentheses.
3. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

Table A2
Correlations between parent and applicant outcomes.

Answered by applicant	Answered by parent				
	Father income	Applicant employed	Applicant income	Applicant log (income)	Applicant income<median
Father income	0.62 (0.16)***				
Applicant employed		0.87 (0.03)***			
Applicant income			0.56 (0.05)***		
Applicant log (income)				0.86 (0.03)***	
Applicant income<median					0.8 (0.04)***

Notes:
1. Each coefficient contains the result of a separate OLS regression in which the dependent variable is the variable indicated as answered by the applicant. The reported coefficient is on the same variable as reported by the parent.
2. Standard errors in parentheses.
3. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

Table A3
Mean of outcomes, by category.

	Means					Comparison of means (*p*-value)
			Lower caste groups			
	Full sample	Upper caste	Total	OBC	SC	Upper vs lower
	(1)	(2)	(3)	(4)	(5)	(6)
A. Employment and income						
Employed	0.74	0.81	0.69	0.79	0.58	0.00***
	(0.44)	(0.39)	(0.46)	(0.41)	(0.49)	
	721	308	413	213	200	
Income	12,637.22	17,706.02	8995.69	12,361.04	5630.35	0.00***
	(14,055.61)	(15,988.59)	(11,162.42)	(12,844.89)	(7880.14)	
	653	273	380	190	190	
B. Employment characteristics						
Works as engineer	0.41	0.47	0.37	0.46	0.28	0.01***
	(0.49)	(0.50)	(0.48)	(0.50)	(0.45)	
	716	306	410	211	199	
Works in government sector	0.13	0.09	0.16	0.14	0.20	0.01***
	(0.34)	(0.29)	(0.37)	(0.34)	(0.40)	
	719	306	413	213	200	
Works in private sector	0.46	0.57	0.37	0.47	0.27	0.00***
	(0.50)	(0.50)	(0.48)	(0.50)	(0.45)	
	719	306	413	213	200	
Self-employed/family-employed	0.05	0.04	0.06	0.08	0.05	0.11
	(0.22)	(0.19)	(0.24)	(0.26)	(0.22)	
	719	306	413	213	200	
Family or friends helped find job	0.10	0.11	0.09	0.12	0.07	0.64
	(0.30)	(0.31)	(0.29)	(0.33)	(0.25)	
	706	304	402	206	196	

Notes:
1. This table reports mean outcome variables, by caste group. Standard deviations are in parentheses; the sample size for each variable is listed below each mean and standard deviation.
2. Column 1 reports the mean for the full sample. Column 2 reports the mean for the upper-caste groups. Column 3 presents the mean outcomes for the pooled lower-caste groups, while columns 4 and 5 report the means for the two main lower-caste categories, OBC and SC respectively. Column 6 reports the *p*-value of the difference in means between the upper and lower-caste groups.
3. Significance at 10% level is represented by a *, at the 5% level by a ** and at the 1% level by ***.

References

Alon, Sigal, Tienda, M., 2005. Assessing the 'mismatch' hypothesis: differentials in college graduation rates by institutional selectivity. Sociology of Education 78, 294–315.

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

Baley, S., 1999. Caste, society and politics in India from the eighteenth century to the modern age. The New Cambridge History of India, vol. IV. Cambridge University Press, Cambridge.

Bowen, W.G., Bok, D., 1998. The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions. Princeton University Press, Princeton, NJ.

Card, D., Mas, A., Rothstein, J., 2008. Tipping and the dynamics of segregation, Quarterly Journal of Economics. The Quarterly Journal of Economics 123 (1), 177–218.

De Zwart, F., 2000. The logic of affirmative action: caste, class and quotas in India. Acta Sociologica 43, 235–249.

Fryer, R.G., Loury, G.C., 2005. Affirmative action and its myths. Journal of Economic Perspectives 19, 16–28.

Fryer, R.G., Loury, G.C., Yuret, T., 2003. Color-blind affirmative action. NBER Working Paper no. 10103.

Henriques, J., Wankhede, G.G., 1985. "One Step Forward Yet Two Steps Behind: a study of wastage and stagnation in education of scheduled castes and scheduled tribes in Maharashtra." Mimeograph.

Herrnstein, R.J., Murray, C.A., 1994. The Bell Curve: Intelligence and Class Structure in American Life. The Free Press, New York.

Holzer, H., Neumark, D., 2000. Assessing affirmative action. Journal of Economic Literature 38, 483–568.

Leonard, J.S., 1985. Affirmative action as earnings redistribution: the targeting of compliance reviews. Journal of Labor Economics 3 (3), 363–384.

Kane, T.J., 1998. Racial and ethnic preferences in college admissions. In: Jencks, C., Phillips, M. (Eds.), The Black–White Test Score Gap. Brookings Institution Press, Washington D.C.

Kumar, D., 1992. The affirmative action debate in India. Asian Survey 32, 290–302.

Loury, L.D., Garman, D., 1993. Affirmative action in higher education. American Economic Review 83, 99–103.

Munshi, K., Rosenzweig, M., 2006. Traditional institutions meet the modern world: caste, gender, and schooling choice in a globalizing economy. American Economic Review 96, 1225–1252.

Osborne, E., 2001. Culture, development and government: reservations in India. Economic Development and Cultural Change 49, 659–685.

Ravallion, M., 2009. How relevant is targeting to the success of an antipoverty program? The World Bank Research Observer 24 (2), 205–231.

Rothstein, J., Yoon, A.H., 2007. "Mismatch in Law School," Working Paper.

Rothstein, J., Yoon, A.H., 2008. Affirmative action in Law school admissions: what do racial preferences do? University of Chicago Law Review 75 (2), 649–714.

Thernstrom, S., Thernstrom, A.M., 1997. America in Black and White: One Nation, Indivisible. Simon & Schuster, New York.

Weisskopf, T., 2003. Globalization and affirmative action. Economic and Political Weekly 38, 2818–2819.

Wolpert, S. (Ed.), 2006. Encyclopedia of India. Thomson Gale.