NBER WORKING PAPER SERIES

AFFIRMATIVE ACTION IN HIGHER EDUCATION IN INDIA:
TARGETING, CATCH UP, AND MISMATCH AT IIT-DELHI

Verónica C. Frisancho Robles
Kala Krishna

Working Paper 17727
http://www.nber.org/papers/w17727

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2012

This project was started with the late Professor Sanghamitra Das. We would also like to acknowledge Dr. Vibha Arora who was also involved in the exit survey but not in the rest of the data or its compilation or analysis, and Dr. Sunil Kale for his help in obtaining the data. We are grateful to Jun Xiao for research assistance at an early stage of the project and to Susumu Imai for comments on an earlier draft. Kala Krishna is grateful to the Human Capital Foundation (www.hcfoundation.ru), and especially Andrey P. Vavilov, for support of the Department of Economics, the Center for the Study of Auctions, Procurements, and Competition Policy (CAPCP, http://capcp.psu.edu/), and the Center for Research in International Financial and Energy Security (CRIFES, http://crifes.psu.edu/) at Penn State University. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch at IIT-Delhi
Verónica C. Frisancho Robles and Kala Krishna
NBER Working Paper No. 17727
January 2012
JEL No. I20,J15,J31,J7

ABSTRACT

Affirmative action policies in higher education are used in many countries to try to socially advance historically disadvantaged minorities. Although the underlying social objectives of these policies are rarely criticized, there is intense debate over the actual impact of such preferences in higher education on educational performance and labor outcomes. Most of the work uses U.S. data where clean performance indicators are hard to find. Using a remarkably detailed dataset on the 2008 graduating class from the Indian Institute of Technology (IIT) in Delhi we evaluate the impact of affirmative action policies in higher education on minority students focusing on three central issues in the current debate: targeting, catch up, and mismatch. In addition, we present preliminary evidence on labor market discrimination. We find that admission preferences effectively target minority students who are poorer than the average displaced non-minority student. Moreover, by analyzing the college performance of minority and non-minority students as they progress through college, we find that scheduled caste and scheduled tribe students, especially those in more selective majors, fall behind their same-major peers which is the opposite of catching up. We also identify evidence in favor of the mismatch hypothesis: once we control for selection into majors, minority students who enroll in more selective majors as a consequence of admission preferences end up earning less than their same-caste counterparts in less selective majors. Finally, although there is no evidence of discrimination against minority students in terms of wages, we find that scheduled caste and scheduled tribe students are more likely to get worse jobs, even after controlling for selection.

Verónica C. Frisancho Robles
Department of Economics
306 Kern Building
The Pennsylvania State University
University Park, PA 16802
vfrisancho@psu.edu

Kala Krishna
Department of Economics
523 Kern Graduate Building
The Pennsylvania State University
University Park, PA 16802
and NBER
kmk4@psu.edu

1 Introduction

Affirmative action (AA) policies in higher education are used in many countries to try to socially advance historically disadvantaged groups. Although the underlying social objectives of these policies are rarely criticized, there is intense debate over the actual impact of minority preferences in higher education on educational performance and labor outcomes. The debate has mainly focused on three issues: *targeting*, *mismatch*, and *catch up*.

It is well known that family income is a strong predictor of performance. Thus, there is great concern about the fairness of *targeting* based on race, ethnicity, or caste rather than on income. If admission preferences only allow richer students within the minority group to traverse the (lower) hurdles required for admission, then they may be displacing poor students from the non-minority or general group. This is also called the "creamy layer problem" in India.

The second issue is *catch up*. Students admitted to college under preferences often start off far behind those admitted under regular admission criteria. But how does the gap between these two groups change as both progress through college? Do they catch up or fall further behind? If those admitted under preferences can catch up, even part of the way, then the case for preferences is clearly stronger than if they fall further behind.

Opponents of AA also claim that the actual gains for the intended beneficiaries of the policy may not exist. In the extreme case, minority students may even be worse off if they are unprepared for the academic environment they obtain access to through the policy. This argument is known as the *mismatch hypothesis*: students who do not qualify for ordinary admission would do better if they enrolled at schools and/or majors which are more in line with their credentials. If there is severe mismatch, then preferences may even do more harm than good.

Most of the studies to date are narrowly focused on the effects of AA on US minorities' college performance and labor outcomes. The US, we think, is a poor setting in which to look for such evidence. In most US higher education settings, selection criteria are relatively nebulous. While institutions do want good students, they pay attention to much more than grades or SAT scores in deciding whom to admit.[1] SAT scores, extracurricular activities, essays,

[1] This is partly explained by a large number of people having close to or perfect SAT scores. The best schools could easily fill their seat with only such candidates. However, based on the U.S. experience, there is reason to believe that this would result in a worse entering class (Blau et al., 2004 and Bowen and Bok, 1998).

alumni ties, interviews, the perceived likelihood of the student coming[2] and donations all matter. Moreover, AA policies in the US are themselves relatively nebulous: even in their heyday, they basically consisted of adding some "points" for race. There were rarely quotas or large and well documented differences in admission standards.[3] Finally, American students have a huge amount of choice over courses while in college. For example, if smart/serious students take harder courses where good grades are more difficult to obtain, while poor students take the "gut" courses where an A- is ensured with minimal effort, then grades may provide little information on actual academic performance. For all these reasons, the U.S. may not be the best place to evaluate the effects of AA.

We argue that other countries, with transparent selection criteria and rigid course structure, provide much more fertile ground for evaluating the effect of AA policies on minority students. The evidence presented here is particularly important due to its focus on India, which provides a better setting than the US. In India, admission criteria are clear: performance in an open admission exam or in the school leaving exam is all that matters. Moreover, admission preferences imposed by AA in India are far greater than the ones given to African American or Hispanic applicants in the US. India has very strict and binding quotas in higher education in favor of scheduled castes (SC) and scheduled tribes (ST). These groups include what were known as the "untouchable" castes, which used to be relegated to the most menial occupations, as well as tribal populations who were isolated from the mainstream and often treated as badly as the SC.[4] The quotas result in very large differences in admission standards, which provide a nice natural experiment.[5] Thus, it is not likely that the empirical results for the Indian case are confounded by the program being a marginal one. Our focus on India also helps us overcome selection problems present in US college data. Most higher education institutions in India have a very strict curriculum which minimizes the issue of self selection into easier courses. In this

[2]Admissions officers are often rewarded on the basis of acceptance rates.

[3]The Texas top 10% law, which guaranteed admission to the top 10% of graduates from all Texas high schools to any state or public University may be one of the few exceptions. See the Texas Higher Education Opportunity Project (THEOP) for more on this. The law was loosened in June 2010.

[4]Lower castes in India represent a greater share in total population than any minority in the US. Even if we only consider the most disadvantaged castes, SC and ST, their 22.5% share surpasses the 13% share of African Americans in the US.

[5]In fact, the quotas are so much in favor of these disadvantaged groups that even with huge differences in admissions cutoffs, some elite schools are not able to fill their quotas.

setting, grades are a very good indicator of college performance.

Using detailed data[6] on the 2008 graduating class from the Indian Institute of Technology (IIT) in Delhi, this article tries to cast some light on the effects of AA on Indian minorities. In particular, we look at income and grade distributions of minority and non-minority students at IIT-Delhi to provide some basic evidence on targeting. We find that SC/ST students are in general poorer than the others at IIT. Using a supplementary data set on *all* applicants (in 2009) with information on scores, caste and place of residence, we show that AA seems to be effectively targeting minority students who are poorer than the average displaced general (GE) student. By analyzing the college performance of minority and non-minority students as they progress through college, we find no evidence of catch up: SC and ST students, especially those in more selective majors, actually seem to fall *behind* their same-major peers.

We also test the mismatch hypothesis using labor market outcomes and students' self reports on emotional and social well-being while at IIT. Without controlling for selection into selective and non-selective majors, it looks like students in more selective majors earn more. Propensity score matching methods that control for selection in observables reduce the estimated effect of major selectivity on wages, though it remains positive and significant for general students. However, if the wage and selection equations are jointly estimated to take into account the role of unobservables, the positive effect among general students goes away, suggesting it was driven by selection. In other words, general students earn more and choose more selective majors because they are better in terms of unobservables. Even more interesting, minority students in more selective majors end up earning significantly less than their same-race counterparts in less selective majors, which supports the mismatch hypothesis. We also identify some evidence in favor of social mismatch: even after controlling for selection, being enrolled in a more selective major increases stress levels and feelings of not belonging among SC/ST students but the effect goes in the other direction among general students.

Although there is no evidence that wages are lower for SC/ST students once selection,

[6]An exit survey designed by Professor Sanghamitra Das, Professor Kala Krishna, and Professor Vibha Arora was administered to students in the class that graduated in 2008. This data was input using funding from the Pennsylvania State University. We are grateful to Dr. Sunil Kale for allowing this data to be collected and input under his supervision. Data on semester by semester grades was provided by the then Dean for Undergraduates at IIT Delhi, Dr. Sunil Kale. All personal identifers are removed from this data. All other data was collected from public sources.

grades, and background characteristics are controlled for, we conclude by asking whether labor market discrimination may be operating in more subtle ways. We find preliminary evidence of discrimination against minority students in terms of the types of jobs they are able to find once they graduate from IIT. Controlling for selection into majors and grades, SC/ST students are less likely to get placed in highly rewarded jobs in the areas of finances and management consulting. This could be due to choices made by the students themselves, in which case the implications of this pattern are benign. For example, if SC/ST students are more risk averse, they may choose a job based on their core competence that pays less but that is more secure than one in finance.

The rest of the paper proceeds as follows. Section 2 sketches out the major findings in the literature so far. Much of this work is on the U.S. and is plagued by data problems. Section 3 describes the IIT context and the reservation policies mandated by the Indian government in higher education admissions. Section 4 describes the data. Section 5 presents the evidence we have put together on targeting, catch up, mismatch, and discrimination. Section 6 concludes and describes the limitations of the study, as well as directions for future research.

2 The Evidence to Date

AA policies are meant to help historically disadvantaged minorities. However, if only richer students within the minority group benefit from AA preferences, displacing poor students from the advantaged group, the fairness of the policy comes under scrutiny. In the US, for example, the use of race as a proxy for income when targeting the poor has been strongly questioned over the last decade. The current debate focuses on a shift from race-based to economic-based affirmative action policies as proposed by Kahlenberg (1996, 2004) among others. Even though it is true that racial diversity has increased in top colleges in the US, income inequality may have increased as competition to get in has, by all accounts, intensified. Carnevale and Rose (2003) find that 74% of the students at the top 146 colleges in the US came from families in the richest economic quarter while only 3% came from the least advantaged quarter. There is also some indication that AA policies that favor African Americans have disproportionately benefited richer minority students. At the 28 selective colleges studied by Bowen and Bok (1998), 86% of

African Americans were middle or upper class students.

In India, however, work by Bertrand et al. (2010) suggests that affirmative action successfully targets financially disadvantaged students applying to engineering colleges. Even though caste-based targeting did not benefit the poorest SC/ST students, admission was successfully reallocated from richer to poorer households. In particular, upper-caste applicants displaced by AA were richer than the lower-caste applicants taking their place.

Besides targeting, two related issues plague the debate on the appropriateness of AA policies: catch up and mismatch. In India, where quotas in some states reach 50% of college admissions, a large burden comes from the expected negative effect on the average quality of students graduating from higher education institutions. If colleges are forced to admit students from scheduled castes until the quota is met, large reductions in students' average ability are expected. The magnitude of these reductions will be determined not only by the level of the quota, but by the initial differences in performance between general and minority students (Kochar, 2010).

However, if minority students catch up while in college, this particular cost of AA policies can be greatly reduced. Alon and Tienda (2007) evaluate the 10% rule in Texas (where the top 10% of the graduating class in public high schools was ensured admission to UT Austin) and argue that the likelihood of graduation rose after the policy was implemented, and did so significantly for blacks. Their evidence is consistent with previous studies that find that those admitted under the top 10% rule outperform those who were admitted with a lower class rank but higher SAT scores, suggesting that SAT scores are not a good indicator of college performance and that there may well be considerable catch up. However, Alon and Tienda's (2007) results are far from conclusive if those admitted from worse schools were less prepared for college and are likely to choose less challenging majors. If this is the case, graduation rates per se may be less than fully informative. Sander (2004) finds that the average performance gap between blacks and whites at selective law schools is large and, more importantly, tends to get larger as both groups progress through college. He also finds that boosting black applicants into more selective schools lowers their probability of graduation mostly though reduced grades.

It is clear that minority students targeted by AA policies have initial academic credentials that are significantly weaker than those of their non-minority peers. If minority students are not able to close the gap, AA policies that allow them into more selective colleges and/or majors may end up hurting them. If minority students attending more selective schools due to AA policies obtain lower grades than the ones they would have obtained in less selective environments, their labor market outcomes could be worsened by admission preferences.

Attempts to empirically evaluate the "mismatch hypothesis" in the US provide mixed evidence. Rothstein and Yoon (2009) and Sander (2004) find evidence of mismatch in law school. Loury and Garman (1993, 1995) find that blacks in the US get considerable earning gains from attending more selective schools but these gains are offset for black students by lower performance both in terms of grades and probability of graduation. Alon and Tienda (2005) assess the effect of college selectivity on the graduation probability. Using both propensity score matching methods as well as bivariate probit models, they reject the mismatch hypothesis suggesting that blacks and Hispanics in the US are able to catch up. We must keep in mind though that US colleges tend to have low performance graduation requirements so that graduation rates may not be a good measure of academic success. Arcidiacono (2005) estimates a structural model that incorporates application decisions, admission decisions, attendance decisions and future earnings. He argues that removing affirmative action reduces the presence of minority students, especially in top schools, but it does not affect income or college attendance by much.

Bertrand et al. (2009) is one of the first studies analyzing the mismatch hypothesis in India. They find that the marginal effect of caste-based admission preferences in Indian engineering colleges is positive for minority students: i.e., they do earn more as a result. However, they gain less than what the students they displace lose. Though their data is better than ours as they have information on accepted and rejected students, they have no information on grades, which account for a large part of the differences in earnings. At the very least, their results are unable to distinguish between the pure gains from graduating from more selective institutions and the loss arising from poorer grades in these institutions.

3 IIT Admission Process and Reservation Policies

The IITs are engineering and technology-oriented universities of national importance. They were initially created to train scientists and engineers that could contribute to India's industrial development after its independence. Today, there are fifteen IITs in the country which function as autonomous universities with their own curricula, although they are linked to each other through a common central government council in charge of their administration. All institutes offer programs leading to a Bachelor's degree but some IITs also offer Dual Degrees, Integrated Master of Technology, Master of Science and Master of Arts degrees.

Admissions to the undergraduate programs and some graduate programs are conducted through the Joint Entrance Examination (JEE). The admission process is very competitive both because of the difficulty of the open competitive exam and the high number of test takers. The undergraduate acceptance rate through the JEE is less than 1 in 60: over 300,000 annual test takers compete for 5500 seats in undergraduate programs. Only 4000 seats are offered by undergraduate programs at IIT and the rest of the seats correspond to other institutions that also use the JEE to evaluate applicants.

The JEE tests the candidate's knowledge of 3 subjects: Chemistry, Mathematics and Physics. After the exam is administered, the average of the marks scored by all candidates is computed for each of the three subjects. These averages give the Minimum Qualifying Marks (MQM) for Ranking in each subject. All students above the MQM in each subject are ranked in terms of their *aggregate* score to construct a *common merit list*. This merit list contains as many students as the number of seats available in all undergraduate and graduate programs offered by the IITs. The aggregate score of the last candidate admitted from this list gives the general cut-off score for admission.

Although minority students take the JEE with the rest of the students, India's law entitles them to preferences. In traditional Hindu society, caste is hereditary and it used to be occupation-specific. Thus, lower caste individuals were trapped in less attractive occupations, both in terms of prestige and wages. Although reservation policies in India were first applied in labor markets, they soon appeared in higher education as a way to reduce the inequalities generated and perpetuated by the caste system (see Bertrand et al., 2009). After independence from Britain, all central government higher education funded institutions were mandated to comply with reservation requirements for traditionally disadvantaged castes. In particular, the central government requires that 15% of the students admitted to universities must be from SC while 7.5% have to come from ST, reflecting their share of the general population.[7]

To comply with the affirmative action policies imposed in higher education admissions, the IIT has implemented caste-based reserved quotas since 1973. After the common merit list is constructed, separate merit lists for SC and ST candidates are prepared. If the number of candidates in each minority list is at least 1.4 times the number of seats available for that caste, the merit list contains all these candidates. If the number of qualified minority candidates is less than 1.4 times the number of available SC or ST seats, the general admission cut-off is reduced by up to 50 percentage points to get the number of candidates as close as possible to 1.4 times the number of seats. Thus, if the general cutoff is 97%, the SC/ST cutoff could be as low as 47%. However, even after extreme relaxation of the cut-off scores for minority students, the aggregate quota of 22.5% for SC and ST students is not always met.[8]

Each program offers a fixed number of seats for general, OBC, SC, and ST students. Once a student qualifies into the relevant merit list, she submits a preference ranking over majors and IIT combinations. Within each merit list, JEE scores can be thought of as bids for a particular program. Placement is offered until the reserved number of seats for a caste group is filled or until all applicants in the corresponding merit list are placed. Ex-post, this system generates major and caste specific cut-off scores. More prestigious majors with higher salaries in the labor market tend to be more competitive and hence have higher JEE cut-off scores for both general and minority students. This allocation process generates assortative matching within each major: top students from the general group are matched with top students in the minority group. However, as the quotas are major by major and the aggregate quota is not filled, SC/ST students are more likely to be in selective majors so that the difference in performance between SC/ST and general students tends to be greatest in selective majors. Note that once at the IIT all students are evaluated under the same criteria both in terms of grades and graduation requirements though SC/ST students usually take longer to graduate.

[7]Reservations for other backward classes (OBC) were recommended in 1978 and implemented in 1989 in private unaided institutions as well as high-end government jobs for minority communities. IIT did not make any changes to its reservations policy until 2008. Since then, OBCs have also been provided with a 27% reservation, although their share in India's population is about 50%. However, it has been argued that the OBC group is not "backward" and some privileged castes have made it on to this list.

[8]Since 2008, a separate merit list is also constructed for OBC students. However, the relaxation in marks applied to the admission cut-off for this group is at most 10%.

4 Data

The sample of students used in this study corresponds to the graduating class of 2008 from IIT-Delhi, which includes 354 Bachelor students and 97 Dual Degree students.[9] The data set contains institutional records and some background information obtained from an exit survey administered to all graduating students.[10] IIT records contain data on GPA and number of credits completed at IIT on a semester by semester basis, as well as some basic information on the students such as gender, caste, age, and major. Additionally, the exit survey data provides information on previous schooling, family income, land, and property ownership, parents' and siblings' educational levels and occupations, expenditures in coaching to prepare for the JEE, as well as some information about placement, such as type of job and first salary after graduation. The survey also collects detailed information on academic life experience, hostel life at IIT-Delhi[11], and extracurricular activities.

A major limitation of the data is that it does not include students' scores in the JEE as a pre-college performance measure. However, this problem can be partly circumvented by using the cumulative GPA (CGPA) of the student at the end of the first year in the IIT. This is a good proxy for the JEE score because the courses taken during the first year of Bachelor's and Dual Degree programs share a common structure across majors and the material covered closely reflects the material evaluated through the JEE.[12] However, there is anecdotal evidence that some students slack off in the first year. In what follows, we define major selectivity based on students' average performance in the major in the first year. According to this criterion, Bachelor's programs in Computer Science and Electrical Engineering (Power) as well as dual degrees in Computer Science and Electrical Engineering are "selective" majors.[13]

The second data limitation is the large number of students with missing values in one or more of the variables obtained from the survey. Only 56% of the sample has complete data for all the 31 variables we use. Almost 36% of the students have missing values in one or two

[9] Dual degrees programs integrate undergraduate and postgraduate studies in selected areas of specialization. They are completed in five years, only one more year compared to conventional Bachelor's degrees.

[10] The survey was only administered in 2008.

[11] Students at IIT live in hostels located in campus throughout their stay in the IIT.

[12] Common courses usually include basic Electronics, Mechanics, Chemistry, and Physics.

[13] See Table C.1 in Appendix C.

variables, and the remaining 8% of the sample has between 3 and 14 variables with missing values. As column 1 in Table A.1 in Appendix A shows, most of the variables have missing values for a few individuals. The variables with the greatest percentage of missing data points are wages after graduation and type of school attended but even in these cases the problem is not severe. To avoid dropping variables or observations, we rely on multiple random imputation methods to generate an imputed complete data set. Assuming the data is missing at random, a common assumption in most imputations methods, the completed data set is obtained using sequential generalized regression models. A complete description of the procedures used is given in Appendix B.

We also rely on a supplementary database on the scores of the universe of JEE *applicants* in 2009 (384,977 students). These records were obtained from the JEE website in September 2010. In addition to each applicant's aggregate score and scores by subject, the records contain information on the caste and place of residence as indicated by the All India Postal Index Number (PIN). Using the PIN code identifiers, we merged the JEE applicant data with district level poverty data in urban areas as well as data on the share of SC/ST population and the share of rural population.[14]

5 Empirical Evidence

5.1 Targeting

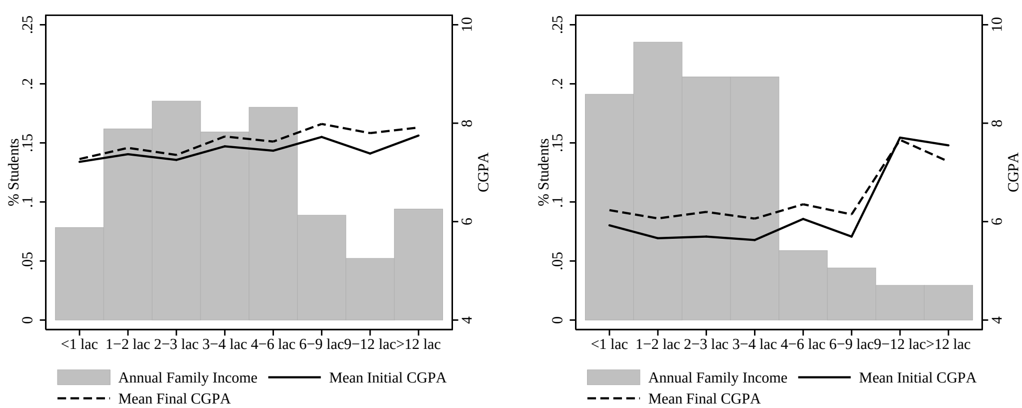
AA policies are usually implemented for two main reasons. First, governments may directly target minority groups pursuing objectives related to diversity, social harmony, and social advancement of historically excluded groups. Second, race (or caste in India) can be a substitute for income to target social interventions. As mentioned above, minority groups in India have been historically trapped in less prestigious occupations with lower wages. Consequently, the government adopted caste-based targeting policies as a substitute for income-based targeting policies with the hope that the former was a more efficient way to identify disadvantaged groups

[14]Poverty data comes from Table A2 in Chaudhuri and Gupta (2009) while data on minority and rural population at the district level is obtained from the Indian Census 2001.

as backward caste status is harder to falsify than income.[15] Verifying income is especially difficult in India since 92% of the labor force works as informal workers (NCEUS, 2009).

Nevertheless, race or ethnicity-based targeting strategies raise questions about the fairness of the preferences. Although it is true that the proportion of people below the poverty line among scheduled castes and tribes is about 50% higher than that among the general category (see Chakravarty and Somanathan, 2008), a common argument against AA is that low caste applicants may be far richer than the average low caste household. Even worse, advantaged minority students may take slots away from poorer students from the non-minority group. Since we do not have data on IIT-Delhi's initial pool of applicants, we cannot address targeting issues by comparing the background characteristics of displaced general students to displacing minority students. Nevertheless, Figures 1 and 2 provide some evidence in our data on the existence of a creamy layer effect resulting from IIT minority quotas.

Figure 1: Is there a Creamy Layer Effect?



(a) General Category

(b) SC/ST

Source: Survey data from IIT-Delhi's graduating class, 2008.

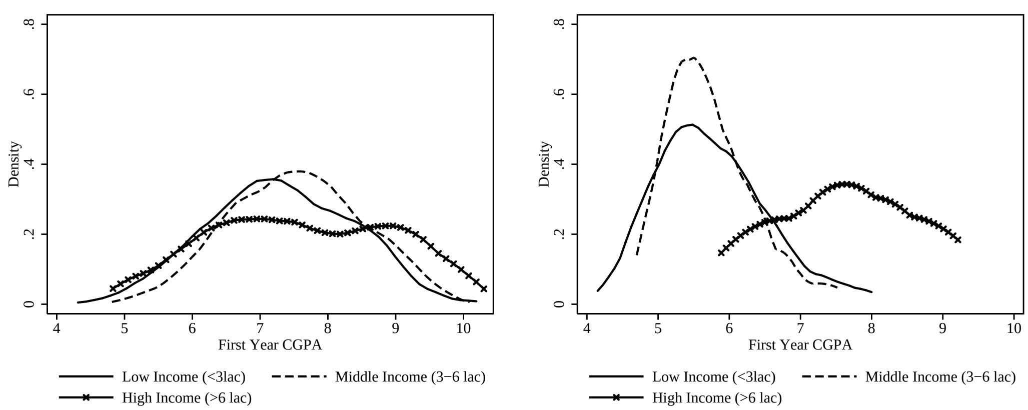
Figure 1 shows that minority students admitted to the IIT are on average poorer than students from the general category (non SC/ST). However, there are a few lower caste students who come from advantaged backgrounds. If we look at the mean CGPA on entrance and at

[15]SC/ST status is documented by certificates issued by the Indian government. Given the widespread income tax evasion in India, income is likely to be underestimated, especially in non-salaried employment.

graduation, we notice that, on average, these "advantaged" students from the "disadvantaged" group have much higher CGPAs both in the first year and in the remaining periods. While there is at best a weak positive link between CGPA and family income for general students, there is a clear jump in CGPA for higher income minority students. Rich SC/ST students perform much like their counterparts in the general category, while poor ones look very different, suggesting that it is the interaction of poverty and SC/ST status that is most harmful.

Figure 2 is consistent with this insight. Panel (a) shows that higher income students from the general group have slightly higher average grades. However, both first order and second order stochastic dominance of the distribution of richer general students is rejected. Panel (b) in Figure 2 shows that high income students among the SC/ST have much higher grades than low income ones. In addition, while the general category has some rich students with poor GPAs, the SC/ST does not. A Kolmogorov-Smirnov type test cannot reject the null of second order dominance of the grade distribution for rich SC/ST students over that for poorer minority students, which is consistent with the higher mean and lower dispersion for the rich observed by eyeballing the data.

Figure 2: First Year CGPA Distribution by Caste and Income Level



(a) General Category (b) SC/ST

Source: Survey data from IIT-Delhi's graduating class, 2008.

Results from a regression of initial CGPA on income, caste, and controls confirm that the interaction between high income (above 9 lacs) and SC/ST is still positive and significant even after adding additional regressors and with very few SC/ST students in the highest income group (see Table C.2). These results confirm the evidence presented above: poor minority students start off lagging behind general and rich minority students.

To summarize, better-off minority students who benefit from caste-based preferences represent a very small proportion of all the SC/ST students admitted into IIT-Delhi (see histogram in Figure 1) and they look like students in the general category. These two facts suggest that the extent of the creamy layer problem at IIT-Delhi might be small. However, we must keep in mind that we are only comparing *admitted* students across income groups and caste and we know nothing about general applicants who are not placed due to the minority preferences.

Fortunately, we can learn more about targeting at the IITs by using the 2009 JEE applicant database. (see Appendix D for more details on district-level patterns identified in the 2009 applicant data.) As described in Section 3, admission to the IITs through the JEE is a deterministic function of the student's score and caste. Within their caste, students are ranked according to their exam score and only those above a caste-specific threshold make it into their group's merit list. Descending through the caste-specific score ranking, students in each merit list are progressively placed until all available seats for that caste are filled or until all students in the merit list are placed. Given this system, we can evaluate the targeting properties of the AA policy. A basic test would be to compare the economic background of minority students admitted under the preferences ("displacing" students) to that of students who are denied a seat but who would have been accepted into an IIT in the absence of the preferences ("displaced students"). If displacing students are less advantaged than displaced students we can conclude that AA is redistributing resources to relatively poorer students.

As mentioned before, the JEE 2009 applicant data lacks of data on IIT placement. However, we approximate displaced and displacing groups using the information on the students in each caste-specific merit list as well as the number of seats available for each caste group. In 2009, all IITs offered 8295 seats.[16] Out of these seats, 4784 were assigned to general students, while 1594, 1282, and 635 were reserved for OBC, SC, and ST students, respectively. We construct

[16]Additionally, 251 seats were reserved for students with physical disabilities but since we do not have these applicants in the JEE data we exclude them from the analysis.

the displaced GE group with all general students who would have obtained a seat at an IIT if the 8295 seats were allocated without caste-specific preferences. These are general students who made it into the common merit list but who were ranked worse than the 4784$^{\text{th}}$ *general* student in this list and so ended up without a seat. The displacing SC/ST group consists of two groups: i) SC students who did not qualify into the common merit list but who are better ranked than the 1282$^{\text{nd}}$ student within the SC merit list and ii) ST students who did not qualify into the common merit list but who are better ranked than the 635$^{\text{th}}$ student within the ST merit list. In 2009, 93% (92%) of the SC (ST) applicants in the SC (ST) merit list are displacing students.[17]

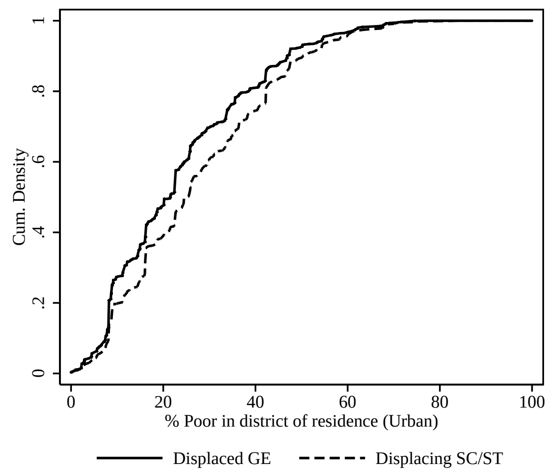
Assuming that neither the number of applicants in each caste nor the number of seats available would change in the absence of the AA policy, we can compare district poverty rates across displaced and displacing groups to evaluate the targeting properties of the program. Figure 3 plots the cumulative distributions of the percentage of poor in urban areas in the applicants' districts of residence for both groups. It is clear that displacing minority students come from poorer districts compared to displaced general students. The null of first or higher order stochastic dominance is rejected, but when rich districts with poverty rates below 4% are dropped, the distribution of poverty rates of displacing students second order dominates that of displaced students. This suggests that the use of minority admission preferences at the IIT is effectively redistributing educational opportunities to students who live in poorer districts. However, lack of individual data on income or consumption in the 2009 JEE applicant database does not allow us to compare living standards of displacing SC/ST's households to that of the average minority household.

5.2 Catch up

The use of minority preferences on college admission in India is justified by the under-representation of SC and ST students in higher education institutions. AA policies were introduced as a way to alleviate the legacy of the caste system and allow SC and ST groups to catch up to the educational and labor outcomes of upper castes. In turn, opponents of affirmative action argue that reservation schemes undermine the quality of education since the lack of preparation of

[17] Although we might mislabel some applicants due to non-enrollment after placement, we expect this potential bias to be small. IITs are top higher education institutions in the country so students who are offered a seat in one of them are not very likely to reject it.

Figure 3: IIT 2009: Does AA Target Students from the Poorest Districts?



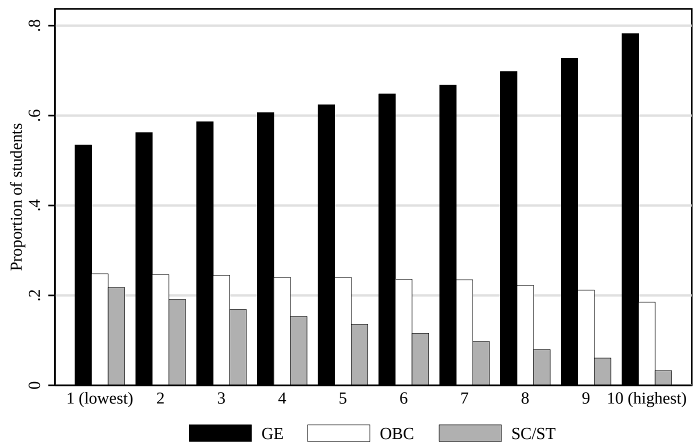
Source: JEE Applicant data, 2009.

minority students during childhood cannot be amended at such a late stage. They argue that minority students who benefit from preferential admission start off college far behind students accepted under regular criteria and that competition while in college will then translate into poor performance among SC and ST students.

In fact, data on test results for the universe of students taking the JEE in 2009 reveal a huge difference in initial caste-based academic levels. If we look at the distribution of JEE scores in Figure 4, we find that SC/ST students account for 22% of those in the lowest decile. Moreover, their participation in higher deciles steadily declines and their contribution in the top decile only reaches 3%. In turn, almost 80% of the students in the top decile come from the general category. Even more striking is the participation of backward castes in the common merit list: out of the 8,295 students who made it into the common merit list in 2009, only 78 come from SC while 15 come from ST.

But how is the educational gap between general and SC/ST students changing once at IIT-Delhi? First of all, what do we mean by educational gap? One approach might be to look at how these two groups fare at the end of their time at IIT relative to the beginning. Figure 5 compares the cumulative GPA of minority and non-minority students at the end of the first

Figure 4: IIT 2009: Proportion of general and SC/ST students by JEE deciles



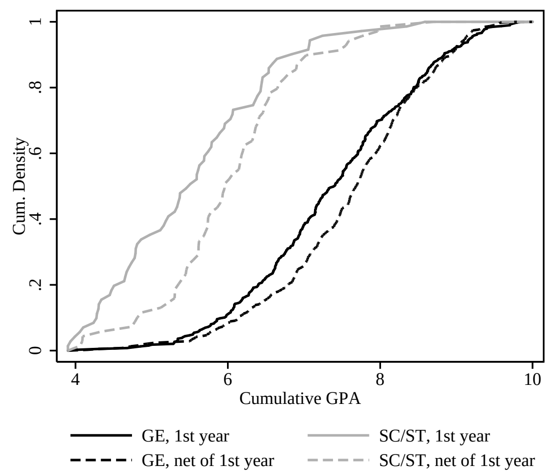
Source: JEE Applicant data, 2009.

year and by the end of their programs, net of their initial performance. While it is clear that the distribution of grades among general students always first order dominates that of minority students (a Kolmogorov-Smirnov type test confirms it), a first look might lead one to think that SC/ST students are catching up. For both SC/ST and the GE group, the CGPA distribution at the end seems to improve relative to that in the first year, and more so for minority students. By this measure, the gap in average CGPA between general and SC/ST students shrinks by 15%.[18] But we argue below that this is not comparing like to like as we need to control, at the very least, for major choices and differences in grading across majors.

Figure 6 presents the evolution of grades over years by caste. Among general students, the distribution of grades does not move much over time. However, the distribution of grades for SC/ST students seems to be improving as time in IIT goes by. While the gap in average CGPAs between the first and the last year contracts by 10% among general students, the reduction is close to 15% among minority students.

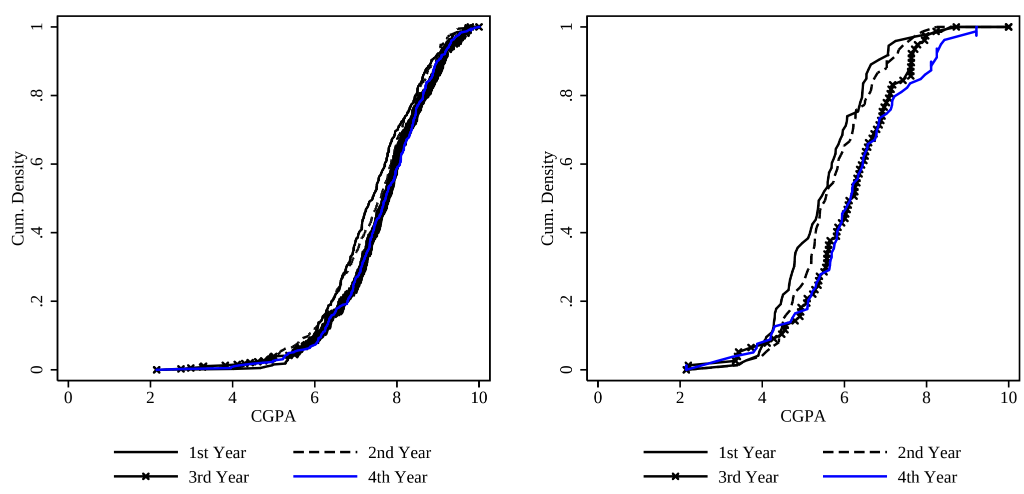
[18]Notice that for each time period and caste group, average CGPA is given by the area above the cumulative distribution. The gap between GE and SC/ST average scores in the first year is just the area between the two solid curves. Similarly, the gap between the two groups at the end of their careers is the area between the dashed curves. The contraction of the gap in averages scores is just the percentage change in these areas between the first and the last year at IIT.

Figure 5: Performance Gap between GE and SC/ST Students at IIT-Delhi



Source: Survey data from IIT-Delhi's graduating class, 2008.

Figure 6: Evolution of GPA over Time



(a) GE (b) SC/ST

Source: Survey data from IIT-Delhi's graduating class, 2008.

It is tempting to interpret the evidence in Figures 5 and 6 as preliminary evidence in favor of catch up, but it is inappropriate to do so without controlling for the major or without taking into account relative instead of absolute grades. First, if GPAs tend to go up in later semesters, say because faculty grade more leniently in advanced courses, we could see CGPAs drifting upwards for all students over time. Moreover, if some majors do more of this than others, and SC/ST students select into these majors, then we might see the above even if there was no catch up. Finally, if we are interested in catch up, we should not be looking at the absolute grades, but at the *relative* grades within each program.

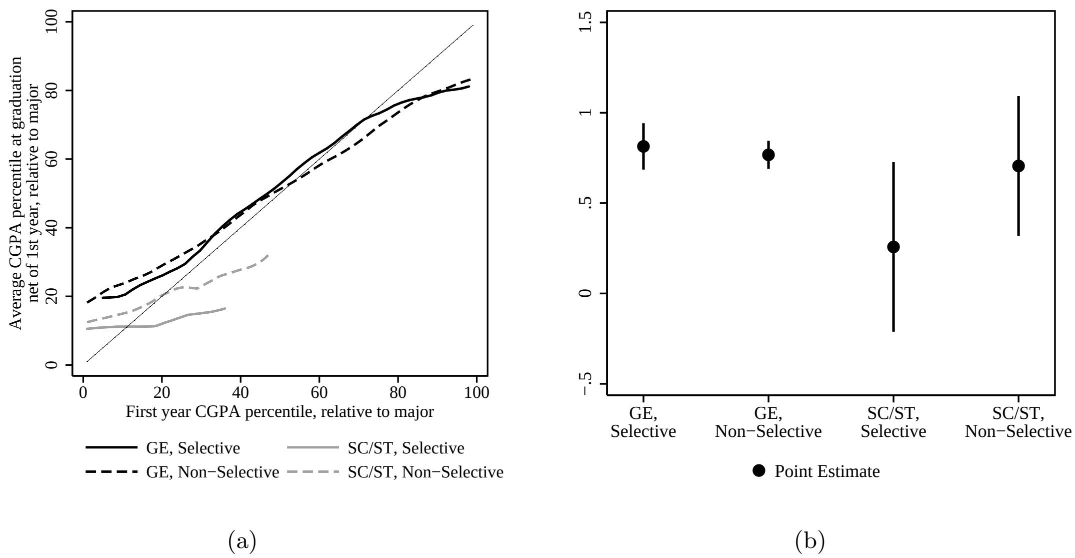
To control for major and relative grades, we look at how students who were in a particular percentile in terms of first year CGPA, *relative to all those in their major*, fared in terms of their CGPA after the first year, *relative to all those in their major*. If students who were in the 30th percentile of their major in terms of their first year CGPA are on average in the 40th percentile in terms of their CGPA in the next 3-4 years of their careers, catch up may be occurring. In turn, if those students average at the 20th percentile after the first year, they are falling back instead of catching up. Since we are considering how students fare relative to others in their major, we eliminate the effect of different grading standards across majors. Moreover, by considering their ordinal standing rather than the level of CGPA, we eliminate the effect of differences in grades in early versus late semesters.

Panel (a) in Figure 7 plots average major ranks at graduation versus first year major ranks both by caste and major selectivity. Average percentiles at the end are calculated through separate locally linear regressions within each group.[19] Thus, for example, general students in non-selective majors who were in the 20th percentile of their majors in terms of their fist year CGPA are, on average, in the 28th percentile of their major at graduation, measured in terms of their final CGPA net of the first year.

In general, the curves presented in Panel (a) show that the slope of the final average ranking with respect to the initial ranking is less than unity, especially at the top and bottom. This is to be expected as those initially at the bottom have no place to go but up and those at the

[19]To avoid misleading patterns due to outliers, the five SC/ST students who ranked above the 50th percentile in the first year are removed. Two of these dropped observations belong to the group of minority students who have high family income.

Figure 7: Average Ranking at Graduation by Initial Ranking: Major-Specific Percentiles of CGPA



Source: Survey data from IIT-Delhi's graduating class, 2008.
Notes: Bachelor programs in CS and EE Power as well as dual degree programs in CS and EE are classified as Selective Majors. Average final rankings are locally mean-smoothed using a kernel-weighted local polynomial smoother.

top have no place to go but down. In other words, reversion to the mean should be present. Nevertheless, in the general category, irrespective of the selectivity of the major, these curves are very close to the 45 degree line suggesting that, on average, general students stay in the major-specific percentiles they start off in. However, this is *not* the case for SC/ST students. Panel (a) shows that although the slope for selective and non-selective majors seems to be below one for minority students, it is even flatter for selective majors, suggesting that SC/ST students are *falling behind* over time but more so in more selective majors.

Panel (b) in Figure 7 presents the results of separate regressions by caste and major selectivity of the final major rank on the initial percentile relative to the major. The round marker shows point estimates of the coefficient of the initial ranking while the vertical line represents the 95% confidence interval of each coefficient. The estimates confirm the pattern observed in panel (a): the slope of final ranking with respect to the initial ranking is close to one for general students, and even so for SC/ST students in less selective majors. In fact, we cannot reject the null of the coefficient being equal to unity for the latter. In turn, the slope for SC/ST students in more selective majors is around 0.25 and unity is clearly far outside the confidence interval.

The evidence presented in Figure 7 suggests that the catch up we seemed to find in the aggregate was an illusion. When we look at how SC/ST students do *relative* to those in their major over time, a different picture emerges. Although minority students in less selective majors are able to at least keep up, SC/ST students in more selective majors seem to be falling behind rather than catching up. This is not surprising considering that minority students who get into IIT start college lagging far behind non-minority students. Consequently, the gap between general and SC/ST students is likely to increase as both groups progress through college, especially in more selective majors. Even by running as fast as they can, SC/ST students can hope, at best, to stay in the same place they started but those in more competitive majors cannot even do this and fall even further behind their general category peers.

5.3 Mismatch

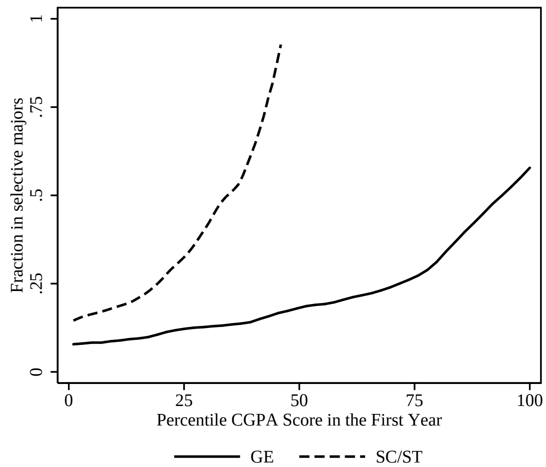
The evidence so far seems to suggest that minority students are not catching up in terms of academic performance. On the contrary, SC/ST students in more selective majors tend to fall behind in their major-specific rankings as they progress through college. This pattern seems to be supportive of the mismatch hypothesis, which predicts lower success rates for minority students who enroll in more selective colleges and/or majors relative to those in colleges and/or majors where their academic credentials are better matched to the average.

We first explore the extent to which AA policies motivate minority students to aim for more selective majors at IIT. Figure 8 displays the fractions of general and minority students in IIT-Delhi who are accepted into selective majors as functions of their first year CGPA percentile.[20] The fraction of students in selective majors at each percentile score is computed from separate locally linear regressions within each group. Although the fraction of students who attend selective majors is increasing in initial performance for both general and SC/ST students, minority students are much more likely to attend selective majors for all levels of first year CGPA. This evidence is consistent with Rothstein and Yoon's (2009) findings on school choice for white and

[20] Again, the five minority students above the 50th percentile in terms of first year CGPA are dropped.

black students in law school and it suggests that the reservation schemes in India are playing a significant role in the major choice of SC/ST applicants.

Figure 8: Fraction of Students Attending Selective Majors by Caste and Initial CGPA Percentile

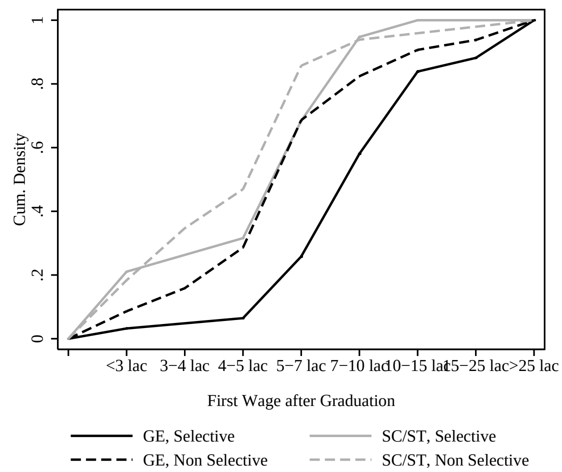


Source: Survey data from IIT-Delhi's graduating class, 2008.
Notes: Bachelor programs in CS and EE Power as well as dual degree programs in CS and EE are classified as Selective Majors. Fractions are locally mean-smoothed using a kernel-weighted local polynomial smoother.

Figure 9 shows preliminary evidence on the effect of minority groups' preferences towards more selective majors. In particular, we plot the cumulative distribution of wages at graduation by caste and major selectivity. It is clear that the distribution of wages for non-minority students in selective majors first order dominates the distribution of wages in all other groups.[21] Whether this is due to ability differences or whether this is because of high returns to enrollment in more selective majors, we cannot say yet. Notice that the gap between selective and non-selective majors seems to be much smaller among minority students than it is among general students. This could be partially explained by a more intense ability screening for non SC/ST students in more selective majors. It could also be explained by SC/ST students gaining less from being in selective majors, i.e., mismatch. However, although SC/ST students do not seem to catch up while in college, it may still be possible that the preferences on admission translate into higher

[21] That of the GE category in non selective majors slightly dominates that of the SC/ST group in selective majors, which is close to that of SC/ST students in non selective majors.

Figure 9: Cumulative Distribution of Wages by Caste and Major Selectivity



Source: Survey data from IIT-Delhi's graduating class, 2008.

wages for those who benefit from the reservations and enter highly selective majors.

The rest of this sub section will evaluate the effect of major selectivity on labor market outcomes for both general and SC/ST students. Taking into account that allocation to selective and non-selective majors is not random, we try to assess the causal link between major selectivity and first wage after graduation. At IIT, admission to selective majors is only driven by performance in the JEE. Thus, the way in which seats in selective and non-selective majors are allocated is not exogenous with respect to future academic performance and wages.

To evaluate the mismatch hypothesis, we compare the wages of students in selective majors with their same-caste counterparts in less selective majors. This task requires taking into account that graduating from a selective major depends on being admitted to that major, which ultimately depends on ability (maybe partially unobserved) and other individual characteristics that also determine wages.

Let S_i^* be a latent continuous variable which is increasing in the selectivity of i's major:

$$S_i^* = Z_i\gamma - \mu_i$$

where Z_i denotes observed individual characteristics such as gender, first year CGPA (proxy for

JEE score), family income, father's education, school type, a dual degree dummy, and type of job. Define $S_i = 1$ if $S_i^* \geq 0$ and $S_i = 0$ if $S_i^* < 0$. Assuming that $\mu_i \sim N(0, \sigma_\mu^2)$, the probability of being enroled in a selective major is given by:

$$\Pr(S = 1|Z_i) = \Phi(Z_i\gamma; \sigma_\mu^2) \tag{1}$$

Let w_i^* denote individual i's log wage at graduation (in logs of lacs):

$$w_i^* = \alpha_1 S_i + X_i\beta_1 - \epsilon_i$$

where X_i is contained in Z_i. Our data gives interval data for wages, so we define a discrete variable, w_i, which will be equal to category W_k if $\xi_{k-1} \leq w_i^* < \xi_k$. We use three wage groups with known thresholds, ξ_k, at $\xi_0 = -\infty$, $\xi_1 = \log(5)$, $\xi_2 = \log(7)$, and $\xi_3 = \infty$. Assuming that $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, and representing the normal cumulative distribution with mean zero and variance σ^2 by $\Phi(x; \sigma^2)$, the probability of $w_i = W_k$ is given by:

$$\begin{aligned}
\Pr(w_i = W_k|X_i) &= \Pr(\xi_{k-1} \leq w_i^* < \xi_k) \\
&= \Phi(\alpha_1 S_i + X_i\beta_1 - \xi_{k-1}; \sigma_\epsilon^2) - \Phi(\alpha_1 S_i + X_i\beta_1 - \xi_k; \sigma_\epsilon^2)
\end{aligned} \tag{2}$$

Our model can be summarized by the system (1)-(2). In particular, the parameter of interest is α_1: if $\alpha_1 \leq 0$, we cannot reject the mismatch hypothesis for SC/ST students in selective majors. However, we need to take into account the endogenous selection process that determines allocation into IIT majors, summarized by (1). Measuring the wage premium of selective majors as the difference in wages of students in selective and non-selective majors gives a biased picture: upward bias is likely since observable and unobservable personal traits that make a student more likely to get into a selective major also make a student more likely to get higher wages. For example, students with higher JEE scores, measured by first year CGPA, are more likely to be in selective majors *and* earn higher wages. If we ignore the role of JEE score in selection into majors, the beneficial effect of major selectivity on wages is overestimated as selective majors have students with relatively higher scores. If we assume that both ϵ_i and μ_i are uncorrelated random shocks, so that selection into more selective majors is only driven by observables, propensity score matching techniques yield an unbiased effect of S_i on wages. However, if ϵ_i and μ_i contain unobserved traits that increase the probability of being in a more selective major *and* that of getting higher wages, propensity score matching methods yield biased estimates of α_1. In this case, the bias caused by the correlation between ϵ_i and μ_i can be prevented by jointly estimating (1)-(2).

Assuming selection is exclusively driven by observables, we obtain propensity scores (or the probability of a student being enroled in a selective major) separately for general and minority students based on their observable characteristics in X_i as well as additional controls such as JEE coaching expenditures, number of household members, and age of the student. We then include the estimated propensity score as an additional regressor in the wage equation, assuming that enrollment in selective majors is random with respect to wages once we control for X_i and the propensity score.[22]

If we also allow selection into selective majors to be related to students' unobservable characteristics (i.e., ϵ_i and μ_i are correlated through unobservables), the system in (1)-(2) must be jointly estimated. Despite the availability of extensive survey data at the student level, we could not find a good exclusion restriction. Finding an instrument that would only affect the major selectivity equation but not the wage equation was impossible given that the determinants of both equations overlap a great deal. Therefore, identification comes from the normality assumptions on the error terms.[23]

Table 1 summarizes our results (see Table C.3 and C.4 for more details). Column 1 provides a baseline estimate of α_1 when the endogeneity of the major choice is ignored and the wage equation in (2) is estimated including S_i and X_i as regressors. Columns 2 provides estimates based on propensity score matching methods, where the propensity score has been added to equation (2) as an additional control.

Among general students, being enrolled in a more selective major seems to have a positive and significant effect even after controlling for selection on observables. Of course, $\hat{\alpha_1}$ is lower once we control for selection on observables, which is expected if students who enroll in more

[22]See Rosenbaum and Rubin (1983). Unfortunately, we have interval data on wages. If the outcome variable were binary or continuous, we could also implement the traditional matching approach which provides a non-parametric estimator of α_1 without making assumptions on the relationship between wages and X_i (or Z_i).

[23]We also experimented with instruments. For the GE group, the results were very similar to those relying on functional form assumptions. The sample size was too small in the SC/ST group for us to consider using instruments there.

Table 1: Effect of Attending a Selective Major at IIT-Delhi on Wages ($\hat{\alpha_1}$)

	Interval Regression[a]	PS[b]	Joint Estimation[c]
GE	0.185***	0.176***	0.115
	(0.036)	(0.037)	(0.149)
			$\rho = 0.196$
SC/ST	0.055	-0.001	-0.380*
	(0.085)	(0.093)	(0.199)
			$\rho = 0.913$

Source: Survey data from IIT-Delhi's graduating class, 2008.
Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$
[a] Coefficient on observed major choice in an interval regression for wages.
[b] Propensity score included as control in interval regression for wages.
[c] Joint estimation of selective major choice and wages equations, no instruments.

selective majors are also more likely to earn higher wages. Propensity score matching methods show that general students in highly competitive majors earn 17% more than their same-caste peers in less selective majors. Among minority students, this wage premium is not significantly different from zero, suggesting that admission preferences that facilitate admission of SC/ST into selective majors do not increase future wages in this group.

However, these results should not be taken as the final word. As mentioned above, if students with higher *unobserved* ability tend to choose more selective majors, then controlling for selection as we have done so far will overestimate the effect of being in a selective major. The results in the last column of Table 1 show that the estimated correlation in the errors across equations is positive within each group (general and SC/ST) and, as a result, $\hat{\alpha_1}$ falls, becoming insignificant for the GE group and significantly negative for the SC/ST group. Being in a selective major has no effect on the earnings of GE students: the wage premium previously identified for them was only due to unobserved ability differences. This is consistent with the work of Altonji et. al. (2005) in the US on the effect of attending Catholic schools on the probability of graduation. For the SC/ST group, it suggests mismatch. Minority students who enroll in selective majors as a consequence of AA policies obtain lower wages than they would have had if they had chosen a less selective major.[24]

[24] A weakness in our estimates is that the estimated ρs, though positive, are not significantly different from zero

In addition to its effect on wages, being enrolled in a more selective major can also affect minority students' well-being while at IIT. Since they see themselves falling behind general students in the same major they could also be facing higher instantaneous costs of going to college by being more stressed and frustrated than their same-caste peers in less selective majors.[25]

We rely on survey data on academic life experience and hostel life to analyze students' well-being while studying at IIT-Delhi. In particular, we focus on two survey questions. The first one asks the students if they had ever felt stressed, depressed, lonely, or discriminated at the IIT where the possible answers are *Never*, *Occasionally*, *Often* and *Regularly.* We also rely on a question that asks the student if he or she felt that the hostel was like a home away from home.

Figure 10 reveals important levels of discontent and/or social discomfort which seem to be higher among minority students, more so in more selective majors. Panels (a)-(d) show that SC/ST students in more selective majors are more likely to experience some degree of stress, depression, loneliness, or discrimination compared to their counterparts in easier majors. Moreover, minority students in the tougher majors are less likely to feel as comfortable as at home while living in IIT's hostels (see panel (e)). These patterns do not appear among general students, where the distribution of the answers seems to be less affected by major selectivity.

Let variable y_i^* be a latent variable that denotes an aspect of student's well being such as stress, depression, loneliness, discrimination, or feeling comfortable at the IIT hostels:

$$y_i^* = \alpha_2 S_i + X_i \beta_2 - \varepsilon_i$$

Define $y_i = 1$ if $y_i^* \geq 0$ and $y_i = 0$ if $y_i^* < 0$. Assuming that $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$, the probability of feeling stressed (depressed, lonely, discriminated, or comfortable at the hostels) is then:

$$\Pr(y_i = 1 | X_i) = \Phi(\alpha_2 S_i + X_i \beta_2; \sigma_\varepsilon^2) \tag{3}$$

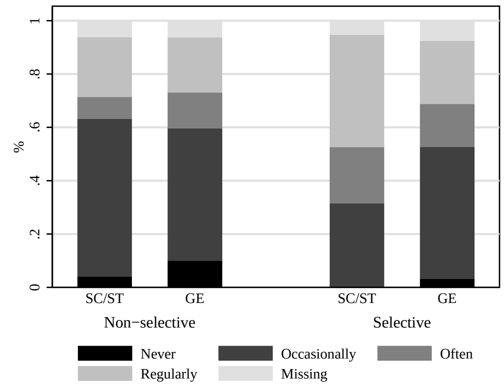
Table 2 reports our estimates of α_2.[26] Again, we control for selection into more selective

as we can see from Table 1 above.

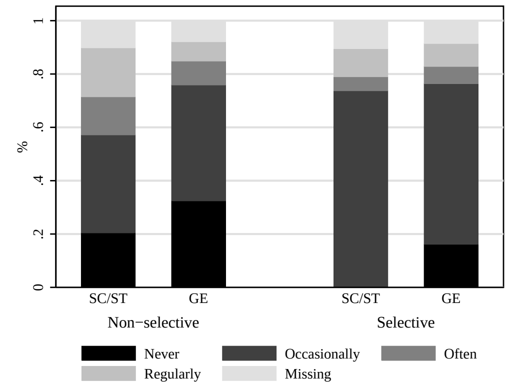
[25]Table C.5 in Appendix C shows that indeed, students whose first year CGPA is below the average performance in their major seem to be the most emotionally affected as the coefficient on the demeaned first year CGPA shows.

[26]We drop individuals with missing observations in each of the well-being variables. We did not include these variables in the imputation process because the missing at random assumption is less defendable in these cases. The problem is small though; missing values in the stress, depression, loneliness, discrimination, or hostel is like home variables only represent between 6% and 10% of the individuals in each caste.

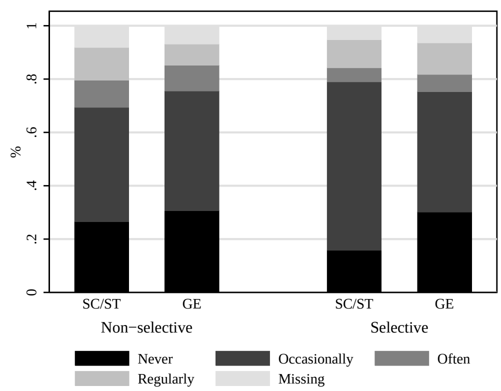
Figure 10: Well-being of the Students While at IIT-Delhi



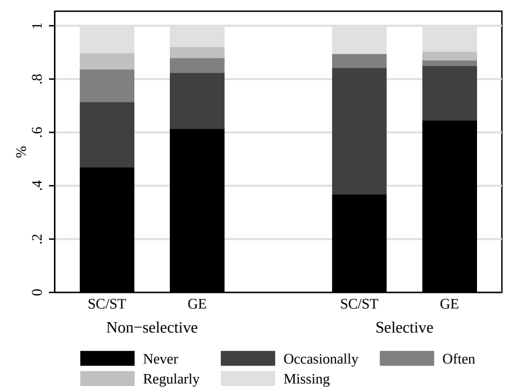
(a) Stress



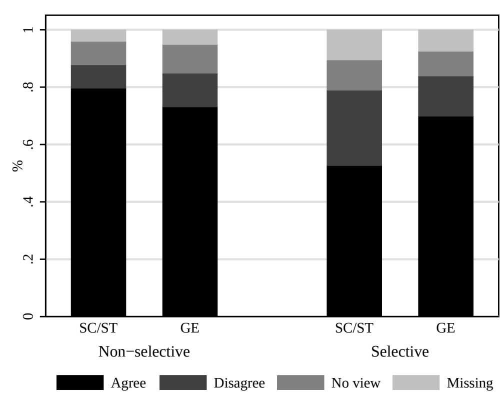
(b) Depression



(c) Loneliness



(d) Discrimination



(e) Hostel feels like home

Source: Survey data from IIT Delhi's graduating class, 2008.

Table 2: Effect of Attending a Selective Major at IIT-Delhi on Emotional and Social Well-Being ($\hat{\alpha_2}$)

	Probit[a]	PS[b]	Joint Estimation[c]
Stress			
GE	0.112*	0.125*	-0.082***
	(0.065)	(0.068)	(0.010)
			ρ=0.85
SC/ST	0.351***	0.320**	0.115**
	(0.120)	(0.137)	(0.046)
			ρ= -0.34
Depression			
GE	0.015	0.019	-0.110***
	(0.052)	(0.053)	(0.014)
			ρ= 0.86***
SC/ST	-0.249	-0.212	-0.013
	(0.152)	(0.162)	(0.353)
			ρ= -0.38
Loneliness			
GE	0.035	0.042	-0.093*
	(0.052)	(0.054)	(0.049)
			$\rho = 0.96$
SC/ST	-0.089	-0.068	-0.061
	(0.141)	(0.161)	(0.109)
			$\rho = 0.32$
Discrimination			
GE	-0.026	-0.026	-0.096**
	(0.043)	(0.045)	(0.049)
			$\rho = 0.68$**
SC/ST	-0.018	-0.018	-0.082
	(0.172)	(0.212)	(0.124)
			$\rho = 0.56$
Hostel feels like home			
GE	0.005	0.008	0.014
	(0.056)	(0.058)	(0.099)
			ρ = -0.09
SC/ST	-0.159	-0.088	-0.103**
	(0.118)	(0.125)	(0.052)
			ρ=0.78

Source: Survey data from IIT-Delhi's graduating class, 2008.
Standard errors in parentheses.
Note: Stress, depression, loneliness, and discrimination dummies are constructed by coding *Often* and *Regularly* as ones. Hostel feels like home dummy is obtained by coding with ones all students who agree with the statement.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$
[a] Coefficient on observed major choice in a probit model.
[b] Propensity score included as control in probit model.
[c] Joint estimation of selective major choice and well-being equations, no instruments.

majors using propensity score matching methods and jointly estimating the system (1)-(3). Once selection into observables and unobservables is taken into account, general students in more selective majors tend to be happier at IIT than their same-race counterparts in less selective majors. The estimated ρs for general students suggest that more able or motivated students who are more likely to choose tougher majors are also more likely to face higher levels of emotional discomfort during their IIT experience. Once this correlation is taken into account, general students in more selective majors are more likely to fit in socially than non-minority students in less selective majors. However, the story is different for SC/ST students. Compared to minority students in less selective majors, SC/ST students in tougher majors are significantly more stressed and feel less comfortable at IIT's hostel facilities. Mismatch is not only generating lower wages, but it is also imposing a higher cost of going to college on them.[27]

In sum, the results suggest that general students do not benefit from being in selective majors; they earn more because they are better. However, they do tend to face relatively lower emotional and social costs of studying at IIT compared to general students in less selective majors. Taking unobservables into account shows that SC/ST students in selective majors earn less than minority students in less selective majors, supporting the mismatch hypothesis. These students also seem to experience what we call *social* mismatch as they also feel more stressed and less comfortable at IIT's hostels when compared to SC/ST in less selective majors.

5.4 Discrimination in the Labor Market

In the previous sub-section, we found that SC/ST students in selective majors earn less *relative* to minority students in easier majors. However, we cannot tell if minority students are better off as a group due to AA in higher education. Lack of an adequate control group or a structural model impede us to compare minority students' wages after graduating from IIT to their wages in the absence of minority preferences.

Since we know that very few students from the minority group enter the JEE's common merit list, it is very likely that without minority preferences most SC/ST students admitted into IIT would have gone to a worse school - that is, if they even made it into college. Given that the Indian labor market rewards degrees from better universities with higher wages, AA beneficiaries

[27]Social mismatch does not seem to affect wages or grades directly.

are expected to be better off as a group. However, even if SC/ST students experience gains from attending IIT, these benefits could be offset by labor market discrimination. If employers with negative stereotypes against minority students are less likely to assign workers from that group to highly rewarded jobs, we could observe discrimination in terms of wages or occupations (Coate and Loury, 1993). Even worse, if employers are aware of the extent of minority quotas, coming from IIT's SC/ST group could become a negative signal in the labor market.

Here we provide some basic evidence on discrimination. We evaluate if, conditional on the admission minority preferences in place, there is discrimination in the workplace. Although we are aware of the limitations of a regression-based approach that controls for observable differences among minority and non-minority students, there is still some informative value in this exercise. Of course, differences in unobservable traits across groups can bias our estimates.[28]

Tables C.3 and C.4 showed that the constant term in the wage equation varies very little by caste, implying that caste-specific mean wages net of observables are very close. Indeed, even after controlling for selection into majors, grades, and type of job we cannot reject the null that the constant in the wage equation is equal across groups. But what if discrimination operates in more subtle ways? We ask below whether SC/ST students seem to get what are seen as less prestigious or upwardly mobile jobs like those in Finance. Of course, this could be (i) because such jobs are riskier than jobs using their core competencies in engineering and SC/ST students are risk averse and choose not to go for such jobs in which case the outcome is benign, or (ii) because employers are giving SC/ST worse jobs.

To test this hypothesis, we measure occupational differences by caste, conditioning on major choice and grades. We estimate an ordered probit model where the outcome variable increases as the job becomes less menial from core and analytical jobs to management/consulting, financial, and non-technical positions.[29] Let j_i^* be a continuous latent variable increasing in job's prestige:

$$j_i^* = \alpha_3 S_i + X_i \beta_3 - \nu_i$$

In the data, we find 5 job types, j_i, where the first category are core jobs and the last one corresponds to non-technical positions. The thresholds between categories, ζ_k, are unknown and

[28] See Charles and Guryan (2011) for more details.

[29] We order them according to the average wages by occupation in our data.

have to be estimated. Again, we assume that $\nu_i \sim N(0, \sigma_\nu^2)$ so that the probability of getting a job in category k is:

$$\begin{aligned} \Pr(j_i = J_k | X_i) &= \Pr(\zeta_{k-1} \leq j_i^* < \zeta_k) \\ &= \Phi(\alpha_3 S_i + X_i \beta_3 - \zeta_{k-1}; \sigma_\nu^2) - \Phi(\alpha_3 S_i + X_i \beta_3 - \zeta_k; \sigma_\nu^2) \end{aligned} \tag{4}$$

We estimate the model described by (4) in the complete sample of students but adding a caste dummy to matrix X_i which is equal to one when the student is from the general group. If the coefficient for caste in vector β_3 turns out to be positive and significant, we argue that there is some evidence in favor of discrimination against SC/ST graduates in terms of occupation.

Table 3 presents the results for discrimination without controlling for the endogeneity of major choice, as well as the results that rely on propensity scores and joint estimation of (4)-(1) to control for selection. When selection into selective majors is ignored, the results from an ordered probit indicate that initial performance, as measured by the first year CGPA, has an important positive effect on the probability of getting better jobs. Being from the general group also seems to have a positive and significant effect on the type of jobs IIT-Delhi graduates get. Even after controlling for selection in observables and unobservables, general students get placed in better occupations relative to comparable SC/ST students but the effect of grades on job type vanishes. It is worth noting that major selectivity and father's education turn out to have a positive and significant effect on job type that persists after controlling for selection in unobservables. In sum, although there are no earnings differentials by caste within occupations, the evidence shows that minority students tend to start their labor market experience in less lucrative occupations which are also less likely to offer professional development opportunities.[30]

6 Conclusion

AA policies always generate divided opinions and provoke intense debate. In the US, for example, evidence showing that AA policies have only benefited richer black students has promoted

[30]Notice that the estimated ρ is negative, which implies that students who have higher shocks in the selection equation also have lower shocks in the type of job equation. In other words, students who prefer more selective majors also prefer worse jobs. This is potentially driven by students in selective majors who prefer to work in academic jobs instead of highly paid jobs in the private sector.

Table 3: Occupational Discrimination

	Ordered[a] Probit	PS[b]	Joint Estimation[c]
Selective Major	0.194	0.170	1.089***
	(0.131)	(0.136)	(0.365)
GE	1.757***	1.883***	1.888***
	(0.208)	(0.280)	(0.202)
Male	0.097	0.072	0.040
	(0.172)	(0.176)	(0.173)
Dual degree/Master	0.132	0.068	-0.009
	(0.126)	(0.158)	(0.139)
CGPA First Year (CGPA1)	0.127**	0.058	-0.021
	(0.055)	(0.115)	(0.083)
Household Income (base: <3 lac)			
$[3-6]$ lac	0.001	-0.030	-0.064
	(0.117)	(0.126)	(0.119)
>6 lac	0.022	-0.028	-0.084
	(0.167)	(0.183)	(0.171)
Father's Education (base: High School or less)			
College or Technical Training	0.252	0.312*	0.356**
	(0.158)	(0.182)	(0.159)
Grad Education	0.183	0.220	0.244
	(0.173)	(0.181)	(0.171)
School Type (base: Public)			
KV/Minority	-0.068	-0.036	-0.003
	(0.243)	(0.248)	(0.243)
Private	0.043	0.042	0.036
	(0.154)	(0.154)	(0.152)
Propensity Score		0.469	
		(0.699)	
Observations	451	451	451
Log likelihood	-621.3	-621.1	-831.1
ρ			-0.535**

Source: Survey data from IIT-Delhi's graduating class, 2008.
Standard errors in parentheses.
$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$
[a] Coefficient on observed major choice in a ordered probit model.
[b] Propensity score included as control in ordered probit model.
[c] Joint estimation of selective major choice and type of job equations, no instruments.

proposals to shift from race-based to economic-based affirmative action. This paper offers evidence that can contribute to the debate on reservation policies in higher education admissions in India by analyzing the real impact of such policies on their intended beneficiaries.

Using detailed data on the 2008 graduating class from the Indian Institute of Technology (IIT) in Delhi and applicant data on students taking the JEE in 2009, this article tries to cast some light on the effects of AA on Indian minorities. Our paper is particularly relevant because it overcomes the limitations of US studies. First, admission criteria are clear and rigid so that performance in the JEE is all that matters to get into an IIT in contrast to the much more nebulous admissions process in the US. Moreover, the use of AA policies in India results in very large differences in admission standards, which implies that our empirical results are unlikely to be confounded by the program being a marginal one. Finally, the strict curricula in higher education Indian institutions minimizes the issue of self selection into easier courses within a major so that grades are a better indicator of performance.

We provide some basic evidence on three issues central to the debate over AA: targeting, catch up, and mismatch. We also offer some basic evidence on discrimination. We find that minority admission preferences seem to be doing a reasonable job targeting poorer populations, though there seems to be little evidence of catch up. In fact, minority students in more selective majors seem to be falling behind, suggesting that mismatch effects might be present. Finally, we find no effect of major selectivity on wages for minority students when only observables are assumed to drive selection into selective majors. However, when the effect of unobservables is taken into account, the results suggest that minority students do *not* benefit from being in selective majors to which they are attracted by the preferences. In fact, SC/ST students in selective majors earn less than minority students in less selective majors, supporting the mismatch hypothesis. Although there are no earnings differentials by caste within occupations, minority students seem to be placed in worse occupations than general students.

As Rothstein and Yoon (2008) point out, the pervasive nature of affirmative action policies eliminates the possibility of finding a control group for the minorities benefited by the policy. Rothstein and Yoon (2009) argue that much of the existing evidence on catch up and mismatch is flawed since differences in the control groups chosen have been an important source of variation across studies. One of our planned extensions is to develop a structural model that can help us overcome a difficulty faced by most empirical studies on the topic: finding a control group.

References

[1] Allison, Paul. 2001. "Missing Data". Thousand Oaks, CA: Sage.

[2] Alon, Sigal, and Marta Tienda. 2007. "Diversity, Opportunity, and the Shifting Meritocracy in Higher Education" , *American Sociological Review*, 72(4) August. 487-511.

[3] _______________. 2005. "Assessing the "Mismatch" Hypothesis: Differences in College Graduation Rates by Institutional Selectivity", *Sociology of Education*, 78: 294-315.

[4] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools", *Journal of Political Economy*, University of Chicago Press, 113(1): 151-184.

[5] Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?", *Econometrica*, Vol. 73, No. 5 (September), 1477-1524.

[6] Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan. 2010. "Affirmative Action in Education: Evidence from Engineering College Admissions in India", *Journal of Public Economics*, 94(1-2): 16-29 .

[7] Blau, Judith, Stephanie Moller, and Lyle Jones. 2004. "Why test? Talent loss and enrollment loss", *Social Science Research*, 33:40934.

[8] Bowen, William and Derek Bok. 1998. "The Shape of the River", Princeton University Press.

[9] Carnevale, Anthony P. and Stephen J. Rose. 2003. "Socioeconomic Status, Race/Ethnicity and Selective College Admissions", *The Century Foundation*, New York.

[10] Chakravarty, Sujoy and E Somanathan. 2008. "Discrimination in an Elite Labour Market? Job Placements at IIM-Ahmedabad", *Economic and Political Weekly*, 43(44): 45-50.

[11] Charles, Kerwin Kofi and Jonathan Guryan. 2011. "Studying Discrimination: Fundamental Challenges and Recent Progress", *The Annual Review of Economics*, 3:479–511.

[12] Chaudhuri, Siladitya and Nivedita Gupta. 2009. "Levels of Living and Poverty Patterns: A District-Wise Analysis for India", *Economic and Political Weekly*, 44(9): 94-110.

[13] Coate, Stephen and Glenn Loury. 1993. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?", *The American Economic Review*, 83(5): 1220-1240.

[14] Kahlenberg, Richard D. (ed). 2004. "America's Untapped Resource: Low-Income Students in Higher Education", Century Foundation Press, New York.

[15] _______________. 1996. "The Remedy: Class, Race, and Affirmative Action", Basic Books, New York.

[16] Kochar, Anjini. 2010. "Affirmative Action through Quotas: The Effect on Learning In India". *Stanford Center for International Development*, Working Paper No. 430.

[17] ________________. 1993. "Affirmative Action in Higher Education", *American Economic Review, American Economic Association*, 83(2), pp. 99-103.

[18] Loury, Linda Datcher and David Garman. 1995. "College Selectivity and Earnings", *Journal of Labor Economics*, University of Chicago Press, 13(2), pp. 289-308.

[19] National Commission for Enterprises in the Unorganised Sector. 2009. "The Challenge of Employment in India An Informal Economy Perspective", Volume 1.

[20] Rothstein, Jesse M. and Albert Yoon. 2009. "Mismatch in Law School", Working paper.

[21] ________________. 2008. "Affirmative Action in Law School Admissions: What Do Racial Preferences Do?", *University of Chicago Law Review* 75(2), Spring, pp. 649-714.

[22] Sander, Richard H. 2004. "A Systemic Analysis of Affirmative Action in American Law Schools", *Stanford Law Review*, 57(2): 367-483.

[23] Van Buuren, S., Brand, J.P.L., Groothuis-Oudshoorn, C.G.M. and Rubin, D.B. 2006. "Fully conditional specification in multivariate imputation", *Journal of Statistical Computation and Simulation* 76: 1046-1064.

A Appendix A: Pattern of missing observations

Table A.1: Summary Statistics

Variable	% Missing (N=453)	Original Dataset Mean	Original Dataset S.D.	Imputed Dataset Mean	Imputed Dataset S.D.
Male	0.0	0.90		0.90	
General Category	0.0	0.85		0.85	
First year CGPA	0.0	7.20	1.17	7.20	1.17
CGPA at graduation	0.0	7.40	1.03	7.40	1.03
Bachelor's degree (Dual/Master's = 0)	0.0	0.79		0.79	
Major					
CS	0.0	0.13		0.13	
ME	0.0	0.20		0.20	
EE	0.0	0.17		0.17	
CH	0.0	0.13		0.13	
BB	0.0	0.05		0.05	
MT	0.0	0.04		0.04	
PH	0.0	0.06		0.06	
TT	0.0	0.09		0.09	
CE	0.0	0.12		0.12	
Family income (lac)	1.8				
<1		0.09		0.10	
1-2		0.17		0.17	
2-3		0.19		0.19	
3-4		0.17		0.17	
4-6		0.16		0.16	
6-9		0.08		0.08	
9-12		0.05		0.05	
>12		0.08		0.08	
Number of household members	2.4	2.22	0.75	2.25	0.82
Enter IIT in first attempt	3.1	0.49		0.49	
Father's occupation	3.5				
Non main earner - retired		0.05		0.08	
Government or public sector		0.54		0.52	
Private sector		0.42		0.40	
Mother's occupation	3.5				
Non main earner - house wife		0.71		0.72	
Government or public sector		0.17		0.17	
Private sector		0.12		0.11	
Father's education	5.3				
Middle school or less		0.04		0.09	
High school		0.10		0.09	
College		0.53		0.50	
Post-graduate		0.34		0.32	
Mother's education	5.3				
Middle school or less		0.11		0.14	

(*continues*)

(*continuation*)

Variable	% Missing (N=453)	Original Dataset Mean	Original Dataset S.D.	Imputed Dataset Mean	Imputed Dataset S.D.
High school		0.16		0.16	
College		0.36		0.35	
Post-graduate		0.36		0.35	
First wage after graduation (lac)	10.0				
No job		0.03		0.09	
3-4		0.07		0.06	
4-5		0.12		0.11	
5-7		0.39		0.35	
7-10		0.19		0.17	
10-15		0.12		0.11	
15-25		0.03		0.03	
>25		0.05		0.07	
Type of school	14.8				
State/Government		0.17		0.17	
KV/Minority		0.05		0.08	
Private		0.78		0.75	

Source: Survey data from IIT Delhi's graduating class, 2008.

B Appendix B: Multiple Random Imputation

Allison (2001) proposes multiple imputation methods as an alternative to maximum likelihood function methods. Like maximum likelihood, multiple imputation estimates are consistent and asymptotically normal and close to being asymptotically efficient. In addition, multiple imputation has two big advantages over maximum likelihood: i) it can be applied to any kind of data or model and ii) the imputation procedure can be implemented using conventional software. Since imputed values are random draws, the major disadvantage of multiple imputation is that it produces different imputed databases every time you use it.

The most widely-used method for multiple imputation is the Markov Chain Monte Carlo algorithm based on linear regression. However, in the case of the IIT data, important complications arise from the fact that some of the missing variables are categorical. The Monte Carlo method presumes that every variable with missing data is normally distributed and that is clearly not the case for categorical variables. An alternative approach is known as "sequential generalized regression" or "multiple imputation for chained equations" (MICE). Instead of assuming a single multivariate model for all the data, this method specifies a separate regression model for each variable which is used to impute missing values. This method is thus flexible and allows us to incorporate logistic, binomial, or multinomial models for categorical variables.[31]

Each model is estimated sequentially using available data, starting with the variable that has the fewest missing data and proceeding to the variable with the most missing data. After each model is estimated, the parameter estimates are used to generate imputed values. Once imputed values have been generated for all the missing data, the sequential imputation process is repeated, except now the imputed values of the previous round are used as predictors for imputing other variables. The main drawback of sequential generalized regressions is that no theory guarantees convergence to the correct distribution for the missing values. However, simulation based evidence in Van Buuren et al. (2006) suggests that the method works well.

The MICE method is implemented in the following way:

1. For each dependent variable to be imputed, choose a model that reflects the type of data.
2. First round: Imputation starts with dependent variable with the fewest missing data and proceeds to dependent variable with the most missing data.
   - Order dependent variables according to amount of missing data from Y_1 to Y_k. Denote variables with complete data values as X.
   - Regress Y_1 on X and obtain $\widehat{\beta}$ and $\widehat{V}(\widehat{\beta})$. Generate imputed values using observed covariates and coefficients drawn from N($\widehat{\beta}$,$\widehat{V}(\widehat{\beta})$).
   - Regress Y_2 on X and Y_1 (including imputed values of Y_1) and obtain imputed values.
   - Continue until all regression models have been estimated.
3. Second and subsequent rounds repeat the process, but each variable is regressed on *all* other variables, using imputed values from previous rounds.
4. Process ends when stable imputed values are reached or after a specified number of rounds.

[31] See Allison (2001).

C Appendix C: Additional Tables and Figures

Table C.1: Average First Year CGPA by Major and Caste

Major	Total	GE	SC/ST
Bachelor's			
Computer Science	8.20	8.56	6.65
Electrical Engineering, Power	7.70	8.03	6.37
Mechanical Engineering, Production	7.39	7.72	5.59
Electrical Engineering	7.32	7.74	5.99
Mechanical Engineering	7.26	7.57	5.50
Civil Engineering	6.94	6.94	
Engineering Physics	6.89	6.90	6.44
Chemical Engineering	6.75	7.19	5.31
Textile Technology	6.69	7.01	5.19
Dual degree			
Computer Science	8.09	8.21	6.79
Electrical Engineering	7.26	7.31	6.50
Chemical Engineering	7.20	7.36	5.25
Biotechnology	6.92	7.08	5.48
Mathematics and Computing (Master's)	6.45	6.54	5.87

Source: Survey data from IIT-Delhi's graduating class, 2008.

Table C.2: First Year CGPA as a Function of Income and Caste

	First Year CGPA
Constant	7.517***
	(0.205)
Selective Major	0.832***
	(0.104)
Male	-0.402***
	(0.148)
Dual degree/Master	-0.314***
	(0.109)
SC/ST	-1.545***
	(0.168)
Household Income (base: < 3 lac)	
$[3-9]$ lac	0.167
	(0.107)
> 9 lac	0.194
	(0.149)
Household Income x Caste	
$[3-9]$ lac x SC/ST	-0.349
	(0.271)
> 9 lac x SC/ST	1.367***
	(0.516)
Father's Education (base: High School or less)	
College or Technical Training	0.053
	(0.130)
Grad Education	-0.052
	(0.143)
School Type (base: Public)	
KV or Minority	-0.141
	(0.194)
Private	0.070
	(0.129)
Observations	451
R-squared	0.383

Source: Survey data from IIT-Delhi's graduating class, 2008.
Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table C.3: Wage Regressions for General Students

	Interval Regression[a]	PS[b]	Joint Estimation [c]
Constant	1.060***	1.205***	1.012***
	(0.133)	(0.196)	(0.166)
Selective Major	0.185***	0.176***	0.115
	(0.036)	(0.037)	(0.149)
Male	-0.032	-0.048	-0.027
	(0.043)	(0.046)	(0.045)
Dual degree/Master	-0.013	-0.041	-0.002
	(0.032)	(0.042)	(0.040)
CGPA First Year	0.043*	0.020	0.056
	(0.025)	(0.035)	(0.036)
CGPA at Graduation	0.033	0.033	0.028
	(0.028)	(0.028)	(0.030)
Household Income (base: <3 lac)			
$[3-6]$ lac	0.055*	0.043	0.060*
	(0.029)	(0.031)	(0.031)
>6 lac	0.129***	0.112**	0.137***
	(0.043)	(0.046)	(0.046)
Father's Education (base: High School or less)			
College or Technical Training	0.045	0.078	0.032
	(0.040)	(0.052)	(0.049)
Grad Education	0.009	0.034	-0.001
	(0.043)	(0.050)	(0.048)
School Type (base: Public)			
KV/Minority	0.160***	0.164***	0.160***
	(0.057)	(0.058)	(0.058)
Private	0.144***	0.136***	0.148***
	(0.039)	(0.040)	(0.040)
Type of job (base: Non-technical)			
Core	0.069*	0.071*	0.068*
	(0.041)	(0.041)	(0.041)
Management/Consulting	0.021	0.024	0.021
	(0.045)	(0.045)	(0.045)
Analytical	0.058	0.057	0.058
	(0.046)	(0.046)	(0.046)
Financial	0.106**	0.106**	0.107**
	(0.046)	(0.046)	(0.046)
Propensity Score		0.181	
		(0.182)	
ρ			0.196
			(0.396)
Observations	383	383	383
σ	0.203	0.203	0.204

Source: Survey data from IIT-Delhi's graduating class, 2008.
Standard errors in parentheses.
$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.
[a] Coefficient on observed major choice in an interval regression for wages.
[b] Propensity score included as control in interval regression for wages.
[c] Joint estimation of selective major choice and wages equations, no instruments.

Table C.4: Wage Regressions for SC/ST Students

	Interval Regression[a]	PS[b]	Joint Estimation [c]
Constant	1.247***	1.459***	0.830*
	(0.370)	(0.392)	(0.497)
Selective Major	0.055	-0.001	-0.380*
	(0.085)	(0.093)	(0.199)
Male	-0.174	-0.092	-0.331*
	(0.136)	(0.148)	(0.182)
Dual degree/Master	-0.155	-0.144	-0.159
	(0.112)	(0.109)	(0.125)
CGPA First Year	0.046	-0.010	0.172
	(0.070)	(0.080)	(0.106)
CGPA at Graduation	0.056	0.049	0.053
	(0.077)	(0.075)	(0.092)
Household Income (base: <3 lac)			
$[3-6]$ lac	-0.102	-0.115	-0.069
	(0.078)	(0.077)	(0.095)
>6 lac	-0.195	-0.223	-0.136
	(0.158)	(0.157)	(0.199)
Father's Education (base: High School or less)			
College or Technical Training	0.001	-0.005	0.004
	(0.074)	(0.073)	(0.093)
Grad Education	-0.001	-0.024	0.037
	(0.092)	(0.092)	(0.116)
School Type (base: Public)			
KV/Minority or Private			
Type of job (base: Non-technical)	0.098	0.132*	0.028
Core	(0.072)	(0.075)	(0.091)
Management/Consulting	0.021	0.029	
	(0.140)	(0.138)	
Analytical	0.014	0.010	
	(0.178)	(0.175)	
Financial	0.079	0.097	
	(0.189)	(0.187)	
Management/Consulting, Analytical, or Financial	-0.037	-0.026	0.010
	(0.213)	(0.208)	(0.094)
Propensity Score		0.297	
		(0.214)	
ρ			0.913
			(0.182)
Observations	68	68	68
σ	0.219	0.215	0.289

Source: Survey data from IIT-Delhi's graduating class, 2008.
Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
[a] Coefficient on observed major choice in an interval regression for wages.
[b] Propensity score included as control in interval regression for wages.
[c] Joint estimation of selective major choice and wages equations, no instruments.

Table C.5: Well-being as a Function of Relative Academic Disadvantage and Controls

	Stressed	Depressed	Lonely	Discriminated	Hostel feels like home
Selective Major	0.005	-0.114	-0.161**	-0.106*	0.003
	(0.089)	(0.075)	(0.071)	(0.060)	(0.077)
Male	-0.072	-0.024	0.012	0.003	-0.115
	(0.081)	(0.070)	(0.067)	(0.054)	(0.076)
Dual degree/Master	0.047	0.064	0.079*	0.038	0.025
	(0.060)	(0.049)	(0.047)	(0.037)	(0.054)
CGPA First Year (CGPA1)	0.120	0.078	0.158**	0.035	-0.061
	(0.082)	(0.068)	(0.066)	(0.053)	(0.072)
GE	0.069	-0.076	0.193	0.020	-0.046
	(0.136)	(0.105)	(0.136)	(0.093)	(0.126)
(CGPA1 - Major's mean CGPA1)	-0.194*	-0.103	-0.296***	-0.095	0.099
	(0.113)	(0.090)	(0.097)	(0.073)	(0.101)
GE x (CGPA1 - Major's mean CGPA1)	-0.012	-0.032	0.052	0.021	-0.052
	(0.091)	(0.070)	(0.083)	(0.059)	(0.080)
Household Income (base: < 3 lac)					
$[3-6]$ lac	-0.016	-0.040	0.026	0.025	-0.020
	(0.053)	(0.043)	(0.042)	(0.034)	(0.047)
> 6 lac	0.046	-0.016	-0.052	0.039	-0.115*
	(0.076)	(0.063)	(0.065)	(0.049)	(0.063)
Father's Education (base: High School or less)					
College or Technical Training	0.080	0.005	-0.024	0.039	0.003
	(0.069)	(0.055)	(0.055)	(0.047)	(0.060)
Grad Education	-0.005	-0.033	0.050	0.043	-0.001
	(0.076)	(0.061)	(0.059)	(0.051)	(0.066)
School Type (base: Public)					
KV/Minority	-0.096	-0.018	-0.174*	0.027	-0.025
	(0.100)	(0.079)	(0.094)	(0.053)	(0.082)
Private	0.029	0.026	-0.031	-0.040	-0.007
	(0.066)	(0.053)	(0.052)	(0.039)	(0.056)
Observations	422	413	420	412	425
Log likelihood	-270.5	-194.5	-189.4	-128.8	-226.0

Source: Survey data from IIT-Delhi's graduating class, 2008.
Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

D Appendix D: 2009 JEE Applicant Data

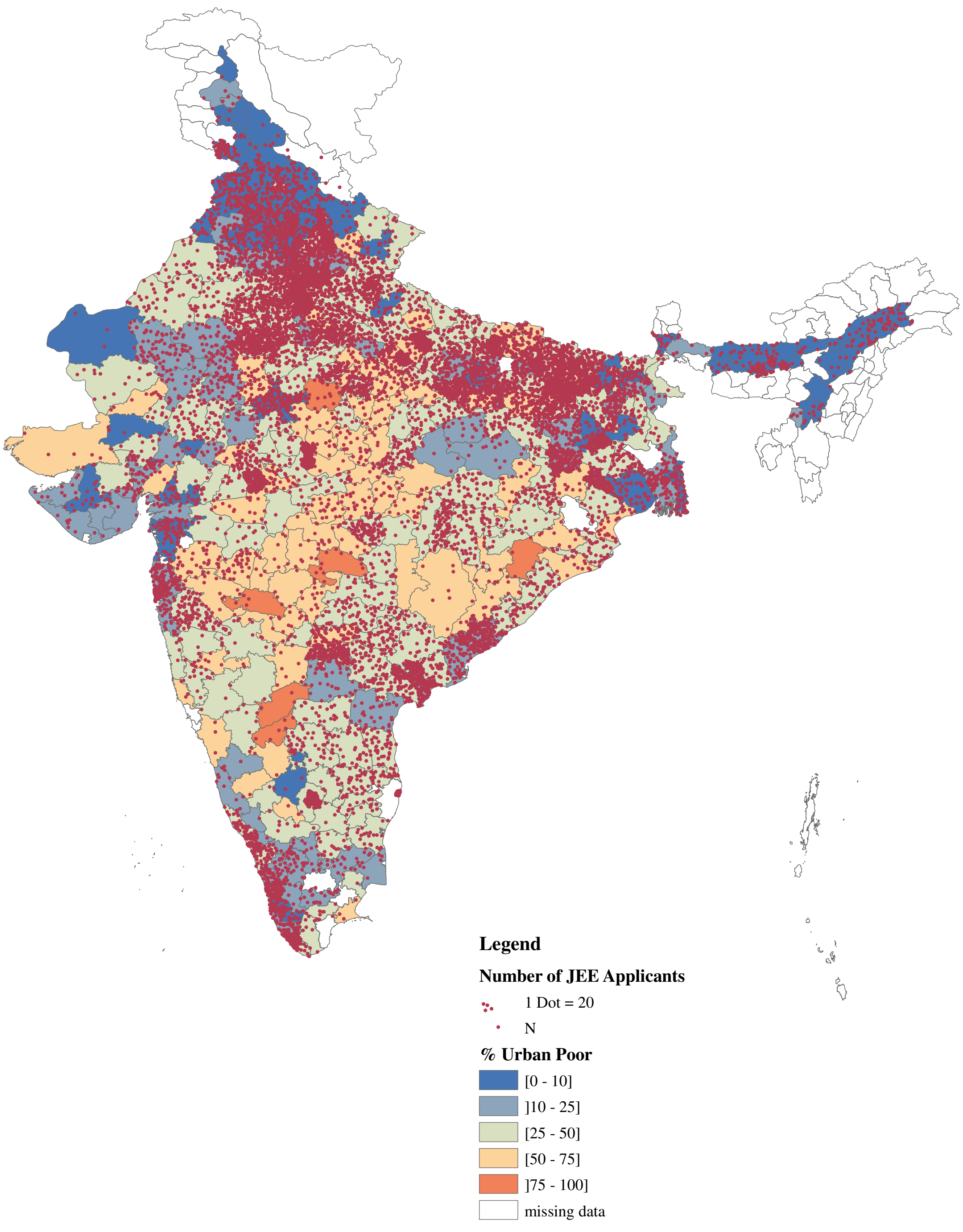
In general, Figure D.1 shows that there are greater concentrations of JEE applicants in north eastern areas, especially in richer districts. Moreover, central regions of the country with higher urban poverty rates are contributing less to the JEE applicant pool. Although this evidence is suggestive, we need to take into account the population "at risk" of taking the exam in each district to check if certain district characteristics are related to a relatively higher proportion of high school graduates taking the entrance exam. We approximate the probability of taking the JEE in a given district as the number of 2009 applicants from a given PIN code divided by the number of high school students enroled in grades 9th through 12th in 2006.[32] Even though the 2009 JEE applicant data does not contain information about students' placement, we code all students who made it into their respective merit list *and* are ranked above the maximum number of seats available for their group as admitted students. We can then proxy the district's probability of getting into college as the number of admitted students divided by the total number of applicants from the corresponding PIN code.

Using the results from non-parametric locally linear regressions in the sample of districts with JEE applicants in 2009, Figure D.2 plots districts' probability of taking the JEE and the probability of getting into college as functions of urban poverty rates, share of rural population and share SC/ST population. The blue lines in panels (a) and (b) show that both the probability of taking the college entrance exam and the probability of getting in do not seem to be affected by the share of minority population in the district. This pattern is particularly interesting if one takes into account that areas with higher concentrations of SC/ST population tend to have lower average performance in the JEE (both in the aggregate and by subject) as shown in panel (a), Figure D.3. In the absence of AA policies, the lower prospects of success in the JEE in areas with higher concentrations of minority population would lead to a lower probability of taking the JEE as well as lower chances of being admitted. The fact that we do not identify a relationship between the district's share of SC/ST population and the probabilities of writing the exam or getting in suggests that admission preferences particularly motivate minority students to take the JEE and facilitate their admission into college.

When we order districts by their poverty rate, panel (a) in Figure D.2 shows a modest decline in the probability of taking the JEE as the percentage of urban poor rises. However, the black line in panel (b) suggests that there is no relationship between the probability of getting in and poverty. This could be due to preferences for SC/ST who tend to be poorer than the GE group. The small differences in the proportion of people getting in may follow from the small differences in average performance among poor and rich districts exhibited in panel (b) in Figure D.3. Finally, the gray lines in panels (a) and (b) in Figure D.2 show that the probability of taking the JEE and the probability of getting into college are both decreasing in the share of rural population, which reflects AA's lack of focus on rural students.

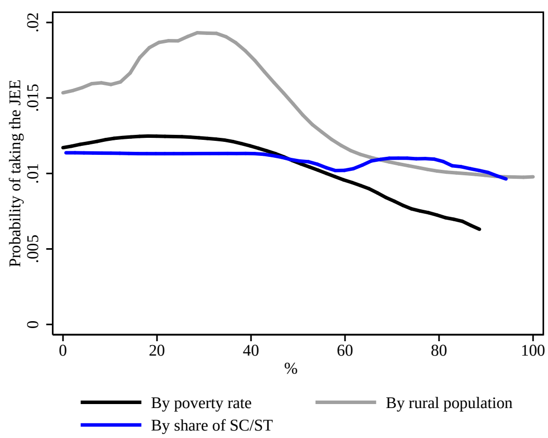
[32]The latter comes from the State Profile 2005-06 prepared by the Indian Ministry of Human Resource and Development. We acknowledge that the total population "at risk"of taking the exam will be overestimated, especially due to higher dropout in the last years of high school, and more so in poorer districts. Unfortunately, data for 12th grade enrolment in 2008 was not publicly available at the district level, only at the state level.

Figure D.1: District-Level Poverty Rate and Number of JEE Applicants

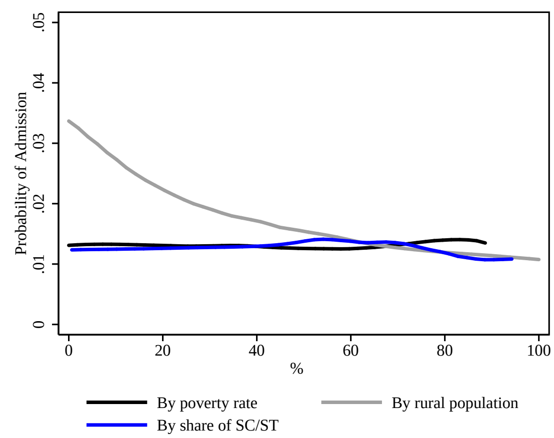


Source: JEE Applicant data, 2009. Poverty rates from and Chakravarty and Somanathan (2008).

Figure D.2: Probability of taking the JEE and probability of admission



(a) Probability of taking JEE

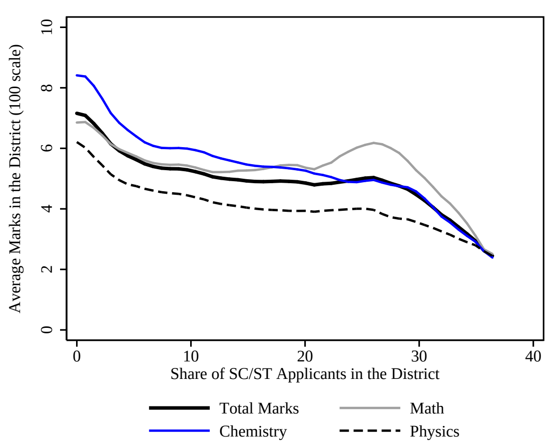


(b) Probability of Admission

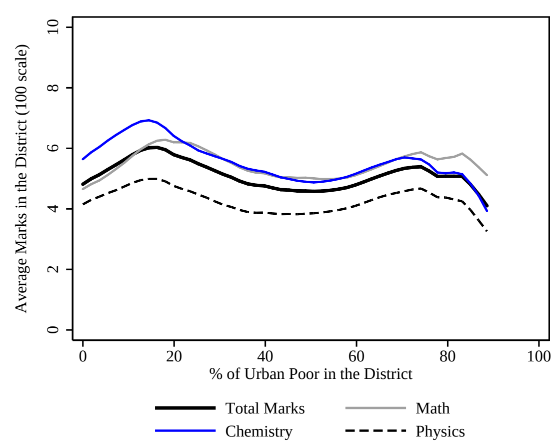
Source: Source: JEE Applicant data, 2009.

Note: Probabilities are locally mean-smoothed using a kernel-weighted local polynomial smoother.

Figure D.3: District Average Total Marks and Marks by Subject in the JEE 2009



(a) SC/ST Share



(b) Poverty

Source: JEE Applicant data, 2009.

Note: Average Marks in panel (a) are locally mean-smoothed using a kernel-weighted local polynomial smoother.