

Under-Representation of Disadvantaged Classes in Colleges: What Do the Data Tell Us?

Rana Hasan and Aashish Mehta
Economics and Research Department
Asian Development Bank*

Analysis of NSS data reveal the following: (i) Scheduled Castes and Tribes, Other Backward Classes and Muslims are seriously under-represented in India's colleges relative to their population shares. (ii) This can be mostly explained by their low higher-secondary school completion rates. Thus, the primary distortions creating unequal representation in college lie at lower rungs of the education ladder. Attention to the quality of basic education, not college reservation, would therefore be the economically "first-best" response to the problem. (iii) Controlling for higher secondary completion, economic status is a better predictor of college attendance than social identity in urban India. Programs to encourage equitable access to urban colleges could therefore consider targeting on the basis of economic status rather than identity. (iv) Compared to their 15% and 7.5% reservations, Schedule Castes and Tribes comprise only 10.2% and 3.9% of the college availing population. This draws attention to the implementation of existing reservations. – Overall, these results emphasize the unequal representation of groups in college, and urge policy makers to seriously consider ways of making the basic education system better serve marginalized groups. They also highlight the importance of selecting policy instruments based on a clinical analysis of the data.

1. Introduction

The Government of India recently decided to reserve 27% of positions in institutions of higher education managed by the central government for Other Backward Classes (OBC). This has prompted a fierce debate on the role of reservations in addressing inequities in educational attainment across social groups. Many proponents of the decision believe that caste-based discrimination denies educational opportunities to the disadvantaged and that reservations are a useful tool in addressing this. Opponents, however, see things very differently. Even if they accept the existence of discrimination, they point to a variety of flaws with reservation policies including: (i) a dilution of academic standards as some meritorious students are displaced by reservations; (ii) an unhealthy rise in caste consciousness among students; (iii) the capture of reserved seats by the elite amongst disadvantaged castes -- elites who may be better off, at least economically, than low income groups among higher castes; and (iv) serious implementation problems, particularly in verifying the caste of those seeking to avail of quotas.

An interesting but disturbing feature of the various debates is the limited and weak database upon which positions have been taken. For example, arguments that reservations for OBCs would seriously dilute academic standards seem to often rely on anecdotal evidence. Similarly, the 52% share of the OBCs in India's population, cited by the 1979 Mandal Commission and used extensively in shaping both debate and policy on reservation, has drawn upon caste data

-
- * This paper represents the views of the authors and does not necessarily represent those of the Asian Development Bank, its Board of Governors, or the countries that they represent.
 - * We are grateful to Paolo Hernando for his assistance with cleaning the NSS data. Shiela Camingue and Susan Torres provided valuable research assistance. We were greatly aided by discussions with Ifzal Ali, Linda de la Fuente, Frank Harrigan, Dalisay Maligalig, Vivek Rao, Vaskar Saha, Janmejay Singh, and Ajay Tandon. All errors are our own.

from the 1931 Census. Concerns regarding the accuracy of this share have been raised by various analysts (see, for example, the discussion in Ghosh 1997) and along with the very different estimate derived from 1999-2000 National Sample Survey (NSS) data, point to the phenomenon of data lagging policy decisions.

In this note, we take what we believe to be a small but important step in confronting reservation related issues with hard data. In particular, we use data from the employment-unemployment survey of the 55th round of the National Sample Survey (NSS) canvassed in 1999-2000 to examine whether certain social classes and religious groups are underrepresented in higher education.¹ We find that relative to the national average, scheduled castes and Muslims are significantly underrepresented in higher education. Of course, this does not automatically imply that discrimination - defined as the rejection of a college application based on the applicant's social or religious identity - is at work. Under-representation may also exist for a variety of other factors unrelated to pure discrimination, including economic ones (for example, the inability to afford college education) and cultural ones (for example, a preference for a religious education over a secular education or a preference for occupations which do not require a secular education). Another possibility is that members of disadvantaged groups are less likely to possess the prerequisite education to enter college.

Although the NSS data do not allow us to distinguish between all the different factors that might lie behind under-representation, they do allow us to dig deeper into the issues in a way that brings some clarity to the current debates. We focus on three questions: First – are particular social or religious groups under-represented in college relative to the overall population? Second – are they under-represented relative to the slice of the population that is minimally qualified to enter college? Third – can inter-group disparities in economic status explain under-representation in college, relative to the minimally qualified population? Our findings suggest strongly that under-representation is not, at heart, a function of the college admissions process, as it mostly originates at lower rungs on the educational ladder. Further, the low economic status of disadvantaged groups explains much of their under-representation in college relative to the minimally qualified population. Finally, even correcting for qualification and income, there is still evidence, although it is substantially weaker, of under-representation.

Of course, whether reservation is the right policy instrument at any level of the educational system – either for addressing inter-group variations in basic educational attainment, or for dealing with the non-economic factors responsible for the under-representation in higher education -- is another matter entirely, and beyond the scope of this paper. In what follows, we first describe our analysis of under-representation of particular groups in higher education, including the role of economic status. This is followed by a discussion of the implications of our findings for the current debate on education-related reservations.

2. Analysis

Under-Representation in Higher Education

As noted above, we draw upon the employment-unemployment survey of the 55th round of the NSS. We limit our sample to only those persons between the ages of 17 and 30, inclusive, as persons reported to be enrolled in graduate level studies or higher fall exactly in this age range. In addition, we further limit our sample to males. Educational attainment, community and

¹ Because caste may not be the only basis for discrimination and/or under-representation of communities in terms of educational attainments, we carry out our analysis by religious groups along with caste/class groupings.

gender are deeply intertwined variables in India for several reasons, some of which may have little to do with government policy or discrimination. Focusing on males allows us to avoid the added complications that gender related issues bring to the interpretation of our statistical results. In the following discussion all statistical results are calculated and presented for the rural and urban population separately.²

The NSS permits respondents to classify themselves as belonging to one of four social groups³, according to the definitions utilized by the state in which the household resides. For ease of exposition, we allocated reported religions to three groups: Hindu, Muslim and Other. The non-responsive categories capture those who did not supply their caste/class or religious identity. .

Table I: Distributions of Social and Religious Groups in Relevant Sub-populations

	Distribution of Males, Aged 17-30 (Percent of sub-population)		Distribution of Minimally Qualified Males, Aged 17-30 (Percent of sub-population)		Distribution of College Availing Male Sub-population, (Percent of sub-population)	
	(1) Urban	(2) Rural	(3) Urban	(4) Rural	(5) Urban	(6) Rural
<i>Class/Caste:</i>						
Scheduled Tribes	3.63	10.49	2.69	6.89	2.45	5.72
Scheduled Castes	14.68	21.56	7.89	14.72	6.93	14.30
Other Backward Classes	31.70	36.63	24.55	31.78	23.03	28.49
Forward Classes	49.98	31.32	64.87	46.61	67.59	51.49
Non-responsive	0.00	0.00	0.00	0.00	0.00	0.00
	100.00	100.00	100.00	100.00	100.00	100.00
<i>Religion</i>						
Hindu	76.97	84.56	83.05	88.38	84.57	88.41
Muslim	16.62	10.03	8.25	6.06	6.90	6.89
Other	6.40	5.41	8.68	5.56	8.50	4.69
Non-responsive	0.01	0.00	0.02	0.00	0.02	0.00
	100.00	100.00	100.00	100.00	100.00	100.00
Sample size:	32,129	46,015	10,237	7,062	6,537	3,572
Imputed sub-population size:	33,233,529	82,642,251	9,984,536	10,612,781	6,341,213	5,125,667

The first two columns of Table I present the estimated distributions of the urban and rural sub-populations across social (class/caste) and religious groupings. Because sample weights are utilized throughout this paper⁴, two sample sizes are provided – the actual number of sample observations, and the number of persons in the population this sample represents, given the sampling scheme. The data show that the existing centrally required quotas for scheduled castes (SCs) – 15% of college seats, and scheduled Tribes (STs) – 7.5%, slightly exceed their shares in the urban male population of this age-group, but are substantially less than their shares in the rural population. Meanwhile, the 27% reservation proposed for OBCs is less than

² Persons may be allocated to the rural sector if their household is based (and therefore sampled) in a rural area, even if they are currently residing in an urban area (and vice versa).

³ The four social groups are: scheduled tribe, scheduled caste, other backward class, and other. We have used the term "forward class" in lieu of "other". Social and religious groups are orthogonal, in the sense that households can, and do, identify themselves as belonging to any of these castes irrespective of their religious grouping.

⁴ The sample weights were constructed in accordance with the instructions accompanying the NSS unit level data and verified against the implied population estimates reported in *Employment and Unemployment in India, 1999-2000: Key Results* (NSSO 2000).

their population share in either environ. The 52% share of OBCs in the general population cited by the Mandal commission report, and others since, is not consistent with the NSS data.⁵ These columns also show that the SC, ST and OBC population is disproportionately rural, while Muslims are more likely to reside in urban areas.

Many definitions of “under-representation” are possible. Under the simplest and most common definition, a community is said to be underrepresented if its share of the population availing of college is less than its share in the overall population. Another definition would consider a community under-represented if its share of the population availing of college is less than its share in the population that is actually “qualified” to enter college. Definitions of qualification may vary. We adopt the least controversial of these – that a student is minimally qualified to enter college only if he has completed higher secondary (HS) school. This definition is reasonable because completing a course of higher secondary education and passing the associated board examinations are just about universally required to enter college⁶ in India.

Columns (3) and (4) of Table I provides the distribution across caste and religious groups of the collage-age male population that have obtained the minimal qualification for entry into college – a higher secondary education. Comparison with columns (1) and (2) provides compelling evidence that SCs and Muslims are substantially under-represented amongst HS graduates relative to their population shares. There is also evidence of under-representation of OBCs and STs amongst HS graduates. Conversely, Hindus and forward classes (FCs) are substantially over-represented.

We present the HS completion rates by community in columns (1) and (2) of Table II. These figures are implicit in the comparison of columns (1) and (2) to columns (3) and (4) in Table I⁷, but are useful to consider explicitly. Two features of these completion rates are noteworthy. First, they are low - averaging 30% in urban areas, and 12.8% in rural areas, for a combined national average of only 17.8%. (UIS, 2005, which uses different sources to our own, arrives at a slightly higher completion rate of 22% amongst Indian males, which compares poorly to results from China - 33%, Thailand - 44%, Indonesia – 41%, and the Philippines – 56%). There is certainly anecdotal evidence that this situation is improving in line with rising demand for skills, and this issue should be revisited as soon as the 2004 NSS data are released. Second, HS completion rates vary tremendously across communities⁸, with SCs and Muslims around *half*⁹ as likely as forward classes and Hindus to be minimally qualified to enter college. The completion rates amongst STs are roughly three-fourths the national average. The OBC rate of HS completion in urban (rural) areas is 77% (87%) of the national average.

⁵ The share of OBCs in the overall population nationwide is 35.8% according to the 1999 employment-unemployment NSS data. A weighted average (using the imputed sub-population sizes in Table I puts the share of OBCs in the population aged 17-30 at 35.2%).

⁶ The definition of “graduate or post-graduate” education (which we use interchangeably with “college”) adopted by the NSS, excludes diploma courses, some of which may not require higher secondary education.

⁷ For example, from Table I, out of 33.2 million urban males in our age-group, roughly 10.0 million are qualified, yielding the 30.04% figure in Table II, column (1). Similarly, from table I, there are (14.68% x 33.2) million urban SC males, of whom (7.89% x 10) million are qualified. Thus the HS completion rate for urban SCs is 16.14%.

⁸ Four Pearson chi-squared tests of the null that completion rates are the same across groups were rejected with p-values of zero. There are four tests because we have two cuts of the population (by religion and social group) and two sub-populations (urban and rural).

⁹ Compare the Muslim HS completion rate of 16.14% with the national average of 30.04%.

Table II: Higher Secondary Completion Rates Amongst College-Age Persons, by Social and Religious Group

	Percent of Males Aged 17-30 completing HS, by group		Percent of Male HS graduates Aged 17-30 availing of college	
	(1) Urban	(2) Rural	(1) Urban	(2) Rural
<i>Class/Caste:</i>				
Scheduled Tribes	22.32	8.44	57.80	40.09
Scheduled Castes	16.14	8.77	55.82	46.91
Other Backward Classes	23.26	11.14	59.82	43.31
Forward Classes	38.99	19.11	66.17	53.35
Non-responsive	1.33	0.00	50.00	n.a.
<i>Whole sub-population average</i>	<i>30.04</i>	<i>12.84</i>	<i>63.51</i>	<i>48.30</i>
<i>Religion</i>				
Hindu	32.42	13.42	64.68	48.32
Muslim	14.92	7.75	53.13	54.94
Other	40.76	13.20	62.17	40.77
Non-responsive	46.90	19.32	71.16	n.a.
<i>Whole sub-population average</i>	<i>30.04</i>	<i>12.84</i>	<i>63.51</i>	<i>48.30</i>
Sample size:	32,129	46,015	10,237	7,062
Imputed sub-population size:	33,233,529	82,642,251	9,984,536	10,612,781

With this background in hand, columns (5) and (6) of Table I shows how males who are availing of or have recently availed of college, are distributed across caste and religious groups. Specifically, we look at the rural and urban sub-samples of males between the ages of 17 and 30 who are either currently enrolled in college courses, or who report having completed a college course. Columns (3) and (4) of Table II present the percentages of HS graduates from each group that avail of college. These can be calculated from the figures in Table I.¹⁰

These columns throw up five sets of results. First, from Table I, we find that amongst the college age population 19.1%¹¹ of urban (6.2% of rural) males avail of college. However, columns (3) and (4) of Table II show that of those actually minimally qualified, a rather high 63.5% urban (48.3% of qualified rural) males enrol in college.

Second, despite reservations of 15% and 7.5% of college seats respectively, STs and SCs make up only 10.2% and 3.9% of the national college attending population. This suggests that any analysis of the impact of reservations will need to consider carefully the implementation arrangements.

Third, taking a population weighted average of rural and urban shares, OBCs comprise 25.5% of the national male college-availing population. This compares with 27% of college seats that the government has proposed reserving for OBCs. Assuming, for illustration's sake, that the gender composition of OBC college attendees is unaffected by reservation, it follows that *if* the 25.5% figure is representative of all colleges that would be brought under the policy, the fraction

¹⁰ These figures are implicit in the comparison of columns (3) and (4) with (5) and (6) of table 1. See footnote 7.

¹¹ From the sub-population sizes in columns 1 and 5 of Table I, (6.34 million / 33.23 million) = 19.1%

of minimally qualified OBC entering college would only rise from 50.2% to 53.0%.^{12, 13} However, it should be noted that some colleges may currently be fielding more than 27% OBC candidates and others less. As a consequence, while these calculations are crudely indicative of the magnitude of the impact of the policy on the college attending OBC population, the outcome of the policy cannot be precisely determined from these data.

Fourth, comparisons of the caste and religious breakdowns of the college-availing population (Table I, columns 5 and 6) with those of the college-age population uncorrected for minimal qualification (Table I, columns 1 and 2), reveal high under-representation amongst SCs, STs and Muslims, and significant under-representation amongst OBCs.¹⁴ Urban persons of other religions are over-represented, while their rural counterparts are not. Filling the arithmetic gaps left, the large Hindus and FC populations are mildly over-represented in college.

Fifth, most (though not all¹⁵), of the evidence presented above of distorted representation of social groups in the college-availing population (columns 5 and 6), disappears when it is compared to the minimally qualified population (columns 3 and 4). While SCs, STs, OBCs, Other Religions and urban Muslims are under-represented in college compared to the minimally qualified population, the margins of under-representation under this comparison are much, much smaller than before. For example, SCs make up 14.7% of the urban group, but only 6.9% of the college-availing. However, they also make up only 7.9% of the minimally qualified group. Thus the share discrepancy falls from 7.8% to 1% when the minimally qualified group is considered as the appropriate comparator.

Columns (3) and (4) of Table II present this last result from a different angle. They show that the proportion of HS graduates availing of college does not vary dramatically across social groups. For example, 55.8% of SC urban male HS graduates avail of college – a figure that, while still indicative of under-representation, is not too far off the 63.5% of all urban male HS graduates who do. In other words, once qualified, persons of various social groups have comparable proclivities to avail of college.

The above comparisons therefore show that evidence of disproportionate representation of social and religious groups in the college-availing population is largely (though not totally), eliminated when the comparator population is limited to those completing higher secondary education. This is serious and important evidence that under-representation of disadvantaged groups in college relative to their population shares has much more to do with factors in operation at the level of higher secondary school and below, than with the process of college admission. Disadvantaged groups are simply far less likely to receive the prerequisite education.

¹² From Table 1, it is straightforward to calculate that nationwide there are roughly 5.82 million college-age, minimally qualified male OBCs, of whom 2.92 million (or 50.15%) are college availing. Raising the share of OBCs to 27%, given the total college-availing male population of 11.47 million, would imply 3.10 million OBCs in college (or 53.016% of qualified, male, college-aged OBCs).

¹³ This assumes that the 27% reservation for OBC students would include all those OBC students who would have otherwise secured an educational seat to the general category (i.e., on the basis of open competition) but choose instead to apply through the reserved category in order to maximize their likelihood of securing a seat.

¹⁴ Pearson Chi-squared tests of the null - that propensity of the college-age population to avail of college is the same across groups - were rejected with p-values of zero in the urban and rural samples and across religious and social groups.

¹⁵ Again, Pearson Chi-Squared tests indicate that inter-group variation in propensity to attend college is statistically significant, even after restricting the sub-sample to males, aged 17-30, that are minimally qualified to attend.

Accounting for Economic Status

The comparisons just described can be further refined by considering the role of economic status. We take the mean per-capita expenditure level (MPCE) of the person's household¹⁶ to capture economic status. Table III presents the mean and standard deviation of MPCE by caste and religious group, for the minimally qualified sub-population.^{17,18} The results show significant levels of inter-caste economic disparity. For example, MPCE among minimally qualified urban SCs is only 78% of the national average, while among similar FCs it is 108% of the average. Muslims and SCs are the least economically well-off groups.

Table III: MPCE of Social and Religious Groups in the Minimally Qualified (Rs. /Person per Month)

	Urban		Rural	
	Mean	Std. Dev.	Mean	Std. Dev.
<i>Class/Caste:</i>				
Scheduled Tribes	1010.26	672.02	678.64	473.78
Scheduled Castes	852.83	610.40	552.82	373.76
Other Backward Classes	955.87	651.27	579.53	375.18
Forward Classes	1185.00	815.57	710.96	445.19
Non-responsive	1283.50	632.59	n.a.	n.a.
<i>Religion</i>				
Hindu	1084.42	715.16	635.65	418.70
Muslim	905.49	634.11	605.86	330.01
Other	1410.43	1187.78	814.68	512.94
Non-responsive	1070.54	61.20	n.a.	n.a.
<i>Sub-population Average</i>	<i>1097.92</i>	<i>769.45</i>	<i>643.6536</i>	<i>421.704</i>
Overall Sample size:	10,206		7,032	
Imputed sub-population size:	9,964,432		10,591,522	

The fact that the groups which are under-represented in college are also poorer suggests that what few distortions remain in the representation of caste and religious groups in college (once obviously unqualified persons are dropped) might be explained by these income disparities. To examine this, we restrict the sample to the minimally qualified population, and estimate a probit model on whether the respondent is college-availing. The explanatory variables considered are caste and religious identity, as well as MPCE and MPCE-squared. FC Hindus are the base case. Table IV presents the results, with coefficients that are statistically significant at the 5% level presented in bold.¹⁹

¹⁶ Readers should note that the MPCE data used here, obtained from block nine of the employment-unemployment survey, are not directly comparable with MPCE data from the Round 55 consumer expenditure survey. This is because in contrast with earlier rounds, the employment-unemployment and consumer expenditure surveys were administered to entirely different households, and the questions concerning expenditure were posed differently.

¹⁷ The sample has been trimmed. Observations reporting MPCE in the highest and lowest 0.1 percentile were dropped.

¹⁸ The reader is reminded that the restriction to the minimally qualified population means that these figures are not representative of the distribution of economic status in the overall population. The restriction eliminates less educated, typically poorer, households. As indicated above, more of these come from lower castes or are Muslims.

¹⁹ We used sample weights, and standard errors corrected for the stratified and clustered sampling design. The NSS stratification scheme involves a complex mix of state, district and rural/urban identifiers, as well as population. For simplicity we assume stratification by state only for estimating standard errors.

Table IV: Probit Results: Determination of College Attendance amongst Minimally Qualified Males, Aged 17-30.

<i>Independent Variables</i>	(1) Urban		(2) Rural	
	<i>Coeff.</i>	<i>Std. Err.</i>	<i>Coeff.</i>	<i>Std. Err.</i>
MPCE (in thousands of Rs.)	0.5850	0.0795	0.4524	0.1626
MPCE-squared (in 000's of Rs. squared)	-0.0648	0.0165	-0.1431	0.0538
Scheduled Tribe	-0.1448	0.1064	-0.2909	0.1219
Scheduled Caste	-0.1644	0.0699	-0.1098	0.0707
Other Backward Classes	-0.1079	0.0565	-0.2247	0.0544
Muslim	-0.2499	0.0702	0.1085	0.0910
Other Religion	-0.1341	0.0732	-0.2329	0.0973
Constant	-0.0907	0.0696	-0.1358	0.0953
Number of Observations	10,206		7,032	
Imputed sub-population size	9,964,432		10,591,522	

The regression coefficients on MPCE and MPCE-squared indicate, as expected that over the expenditure range actually observed, higher economic status increases the propensity to avail of college. However, the regression results indicate that even after correcting for economic status, and restricting the sub-population whose enrolment rate is considered to the minimally qualified, FC Hindus are more likely than any other group to avail of college.²⁰

The regression coefficients therefore show that *both* economic status and social identity play a role. Policy makers wishing to promote equity in the probability of college admission might, therefore choose to target their policy instruments (of which there are many available, in addition to reservations, most obviously – scholarships and preparatory schools), on the basis of social group, or economic status. It would therefore be useful to know whether low economic status or membership of a disadvantaged social/religious group is a better predictor that a student is at risk of not enrolling. The answer is not obvious from the regression results presented.

To investigate this question further, we tested two restrictions on the above probit regression. First, against the unrestricted model presented in Table IV, we restricted MPCE and MPCE-squared to have no effect on enrolment. This restriction is associated, in the urban sample, with a test statistic of $F_{2,4074} = 45.86$. Second, again against the unrestricted model in Table IV, we test the null that college attendance is invariant to the five identifiers of social and religious group. The resulting test statistic in the urban sample is $F_{5,4071} = 4.53$. Both restrictions carry p-values of zero.

This procedure can be viewed as a non-nested test of two theories of college enrolment (Kennedy – 2003, pp. 100-101). The first holds that only identity is relevant and economic status is irrelevant; the second that identity is irrelevant and only economic status matters. The test statistics imply that neither theory suffices. Both identity and economic status are relevant.

This said, the much larger test statistic associated with the restriction on economic status does indicate that economics are a better predictor of college enrolment than identity in the urban

²⁰ The only (statistically insignificant) exceptions are rural Muslims, who are more likely to avail of college than rural Hindus.

population.²¹ Interestingly, in the rural population, the F statistics on the two restrictions are: $F_{2,5966}=4.23$ (economic status is irrelevant), and $F_{5,5963}=6.11$ (identity is irrelevant). In contrast to the situation in urban areas, these results provide no indication that economic status is a better predictor of college enrolment than identity in rural areas.

These results on the relative importance of economics and identity will bear further investigation for two reasons: (i) they could be sensitive to the measures of economic status and choice of social groupings; and (ii) the results of the comparison could be “local” – i.e. economic status could be a better predictor for some groups than for others.

3. Discussion and Concluding Remarks

To recapitulate our main findings: First, SCs and Muslims are massively underrepresented in college, relative to their shares in the wider college-age population. OBCs and STs are also under-represented, although less so. There is therefore an urgent prima-facie case for working on improving the representation of backward castes and Muslims in college.

Second, most of this distorted representation is explained by the abysmal higher-secondary completion rates of these groups. This finding corroborates those of reports like the Probe Team’s, emphasizing the importance of improving access to and the quality of public schools. Reservation of college seats for OBCs is an instrument that will marginally improve access²² to college for only the 14.3% of the OBC population that completes higher secondary school. Given this, and the expense of college education relative to schooling, it follows that amelioration of distorted representation in college might well be better pursued through concerted and courageous reforms of the *pre-college* education system. Such an approach, if successful, would improve the upward mobility of most OBCs, both by improving representation in college, and by serving the vast majority of OBCs who do not pursue college education.

Third, our probit analyses show that group identity is statistically important for predicting college enrolment, even once we control for economic status (by introducing MPCE as an explanatory variable) and higher secondary completion (by focusing on a sample of the minimally qualified). As noted in the introduction, this does not imply discrimination. Further, in urban areas, economic status is a better predictor of college enrolment than identity, implying that much of the remaining under-representation of social groups relative to the minimally qualified population is explained by the lower economic status of these groups. This suggests that it would be useful to consider whether policies to boost educational equality in urban areas might be targeted on the basis of economic status, rather than social group. On the other hand, in rural areas, there is no evidence that economic status predicts the risks of non-enrolment better than group identity. These results suggest the importance of considering different targeting mechanisms in urban and rural environments, and of paying due attention to the empirics of rural-urban migration by students.

Fourth, SCs and STs make up a smaller share of the college population than their quota dictates.

²¹ This comparison of F-tests is intuitively analogous to comparing the R^2 statistics for each of the restricted regressions to see which fits the data best. However, because probit regression with survey weights involves a *pseudo* log-likelihood, R^2 statistics cannot be computed.

²² See footnote 8 for assumptions leading to this conclusion.

In light of this evidence, it is useful to re-examine the arguments of Thorat (2005), who points out that discrimination is a form of economic inefficiency because it results in poor allocations of human resources. As such, he argues that reservation of college seats or jobs corrects the *primary distortion* (respectively, discrimination in college admissions or hiring practices) and should unambiguously increase economic efficiency. Our results are direct evidence that it is incorrect to apply this argument to college admissions in India. Because the vast majority of under-representation in college is explained by lower higher-secondary graduation rates, reservation of college seats cannot be seen as fixing the *primary distortion*. The primary distortion(s) occur at lower rungs on the educational ladder. The bulk of the inefficiency is associated with the *pre-college* education system. There is therefore no obvious reason to believe that fixing the under-representation problems at the college level will enhance economic efficiency.²³

Policy therefore needs to focus on tackling the distortions that prevent children from accessing and completing lower levels of education. And to do this, it is essential that we have a good understanding of the precise nature of the distortions, and where exactly they apply in terms of both the level of education (for example, primary versus secondary schooling) as well as location (for example, rural versus urban India). More concretely, efforts at data collection like that made for the PROBE report, which paid careful attention to the factors which influence both the supply and demand for primary education, need to be undertaken in more parts of the country and for higher levels of education. Such efforts are urgently required in order to design appropriate policy interventions for addressing inefficiencies and inequities in the education system.

References

Ghosh, Partha (1997): *Positive Discrimination in India: a Political Analysis*. Ethnic Studies Report, volume XV, number 2, July.

Kennedy, Peter (2003): *A Guide to Econometrics*. (pub. Blackwell Publishing: Massachusetts).

National Sample Survey Organisation (2000): *Employment and Unemployment in India, 1999-2000: Key Results*. Ministry of Statistics and Programme Implementation, Government of India.

The Probe Team (1999): *The Public Report on Basic Education in India* (pub. Oxford University Press: New Delhi)

Thorat, Sukhadeo (2005): *Reservation and Efficiency: Myth and Reality*. Economic and Political Weekly, February 26, pages 808- 810.

UNESCO Institute for Statistics (2005): *Global Education Digest 2005: Comparing Education Statistics Across the World* (pub. UIS: Montreal, Canada).

²³ Others with whom we have shared these results have pointed out that the patterns we discern in the data could, in theory, be driven by discrimination at college entry, as this would dissuade disadvantaged groups from completing higher secondary education. After all, 63% of urban higher-secondary graduates go on to attend college, suggesting that higher secondary education is pursued specifically in order to secure entry into college. However, there are at least two good reasons to believe that this theory is not empirically relevant to India. First, the PROBE team has shown that dropout rates in basic education are driven largely by the poor quality of education provided and its cost. Second, in further cuts of the NSS data (not included for reasons of brevity), we find that solid under-representation at the primary level, which becomes progressively worse at higher levels of education. Secondary school completion rates look almost as distorted as higher secondary completion rates. This suggests that evening the odds different groups obtaining entry into college will not do much to alter higher secondary completion rates.