



Munich Personal RePEc Archive

Measuring Inequality of Access to Higher Education in India

Borooah, Vani

University of Ulster

2016

Online at <https://mpra.ub.uni-muenchen.de/75683/>
MPRA Paper No. 75683, posted 21 Dec 2016 17:14 UTC

Higher Education in India: Access, Quality, and Structure

Vani K. Borooh*
University of Ulster

March 2016

Abstract

This paper uses data from the 71st NSS round (Education Survey: January-June 2014) to estimate the probabilities of person in India, between the ages of 18 and 22 years, of currently attending graduate or post-graduate courses in institutions of higher education, conditional on their social and economic status, their gender, their marital status, and their urban or rural location. It then examines inequality by social group in the quality of education received. Using the technique of inequality decomposition it estimates the proportionate contributions of the above factors to inequality in the inter-personal distribution of the probabilities of currently attending higher education. It compares how access to higher education has changed between the 64th NSS (July 2007-June 2008) and the 71st (January-June 2014) rounds of the NSS.

* Corresponding author: University of Ulster, Co. Antrim, Northern Ireland BT37 0QB, United Kingdom
Email: vk.borooh@ulster.ac.uk. I am grateful to Ajaya Kumar Naik for help with the data though, needless to say, I assume full responsibility for the paper's deficiencies.

1. Introduction

There are several studies of inter-group disparities in school enrolment and learning in schools (Borooah and Iyer, 2005; Deolalikar (2010); Borooah and Iyer, 2005; Bhalotra and Zamora, 2010; Borooah, 2012; ASER, 2014). Against this emphasis in the literature on schools, academic interest in issues pertaining to inequalities in the higher education sector in India has been relatively neglected (Deshpande, 2013).¹

The issue of underrepresentation by certain groups in higher education is important for a number of reasons. First, conventional wisdom has it that education is the handmaiden of prosperity. As Chamarbagwala (2006) notes, those in India with a college education have gained the most from its economic growth with the result that, in contemporary India, acquiring a tertiary qualification has “become [the] key to gaining entry to the most dynamic segments of employment” (Mohanty, 2006, p. 3777). Second, if India aspires to be an open and equitable society, then gross imbalances between social groups, in spheres of activity which are important to personal success, like higher education, make a mockery of such aspirations.

Against this background, this paper uses data from the 71st National Sample Survey (NSS) round (Education Survey: January-June 2014) to estimate the probabilities of person in India, between the ages of 18 and 22 years, of currently attending graduate or post-graduate courses (irrespective of subject and irrespective of whether it was a degree/diploma/certificate course) in institutions of higher education. Persons so attending are hereafter referred to as ‘in higher education attendance’ (abbreviated to in HEA) - conditional on their social (caste/religion) and economic status (poor/not poor), their gender, and their location (urban/rural). The age band was determined by the fact that the *gross tertiary attendance ratio* (GTAR) is defined in this paper as the percentage of persons currently

¹ See, however, Basant and Sen (2014) who analysed three rounds of the NSS– 55th (1999-2000), 61st (2004-05), and 66th (2009-10) – to investigate the influence of socio-religious group affiliation on participation in higher education and to examine how this might have changed over time. In a similar vein, Sundaram (2006), using data from the 55th (1999-2000) round of the NSS, sought to examine the issue of fair access to higher education.

in HEA, *irrespective of age*, as a proportion of the *total* population in the five year age band after secondary education.² In this study, this five year band was taken as 18-22 years.³

In particular, using these probabilities, we address two questions: (a) in terms of *access* to higher education, is there a ‘social group’/gender/poverty/location effect to HEA stemming from disadvantage? (b) for those in higher education, are there inter-group disparities in the quality of education received? Having answered these questions, the paper uses the technique of inequality decomposition to estimate the proportionate contributions of the above factors to inter-personal in the likelihood of HEA. A comparison with results from the NSS 71st round (January-June 2014) with those from the 64th (July 2007-June 2008) is woven into the paper’s fabric.

At the outset of this study, it is important to draw attention to the fact that all the results reported in it are based upon *grossing up* the survey data using the observation-specific weights provided by the NSS for each of the surveys.

2. Barriers to Higher Education Attendance: Model and Hypotheses

We estimated a logit equation over the sample of persons in the NSS 71st round *who were between the ages of 18 and 22 years* (and, therefore, of an age when they might be most likely to be in ‘higher education’).⁴ The dependent variable (Y_i) in this logit equation took the value 1 if a person was in HEA and zero if he/she was not.

Underlying the logit equation is a *latent variable* model. The idea behind this model is that a person’s propensity for HEA may be represented by the value of the *latent variable*, H_i , with higher values of H_i representing a greater propensity to be in HEA. One may consider this latent variable to be a linear function of K ‘propensity-determining’ factors whose values for individual i are:

X_{ik} , $k = 1 \dots K$. Consequently,

$$H_i = \sum_{k=1}^K X_{ik} \beta_k + \varepsilon_i \quad (1)$$

² UNESCO (2015).

³ Since there were very few persons (51 in the 71st round) in the sample who were enrolled in higher education but not in HEA there was little or no difference between enrolment and attendance in higher education.

⁴ That is, currently attending a diploma or certificate course at graduate or post-graduate level or a degree level or a post-graduate course.

where: β_k is the coefficient associated with the k^{th} variable. An increase in the value of the k^{th} factor will cause the propensity for HEA to increase if $\beta_k > 0$ and to decrease if $\beta_k < 0$.

Since the values of H_i are unobservable, equation (1) represents a latent regression which, as it stands, cannot be estimated. However, what is observable is a person's *HEA status*. The categorisation of persons in the sample in terms of their HEA status is implicitly based on the values of the latent variable H_i in conjunction with a 'threshold value', δ such that:

$$Y_i = 1, \text{ if } H_i \geq \delta \text{ and } Y_i = 0, \text{ if } H_i < \delta \quad (2)$$

The value of δ may be thought of as the *height* of the 'HEA barrier'.

The crucial hypothesis underlying this study is that different persons are faced with HEA barriers of different heights and that these heights are determined by 'disadvantage'. In turn, we locate four sources of disadvantage: 'social group' disadvantage; 'gender' disadvantage; 'economic' disadvantage; and 'locational' disadvantage.

Social Group Disadvantage

In terms of social groups, the Indian constitution recognised that two groups in India – the Scheduled Castes (SC) and Scheduled Tribes (ST) – have historically suffered disadvantage in terms of access to education and jobs. In recognition of this disadvantage, constitutionally guaranteed affirmative action policies reserve places for members of these groups in publically funded educational institutions, jobs in the public sector, seats in the national parliament, state legislatures, municipality boards and village councils (*panchayats*).⁵

In terms of religion, it is Muslims who bear the brunt of deprivation and exclusion in India. The Sachar Committee (2006) in its report to the government of India quantified and highlighted the backwardness of Indian Muslims. This Report drew attention to a number of areas of disadvantage: *inter alia* the existence of Muslim ghettos stemming from their concern with physical security; low levels of education engendered by the poor quality of education provided by schools in Muslim areas;

⁵ Articles 341 and 342 enumerate a list of castes and tribes entitled to 'reservation' benefits. It is important to emphasise that the Constitution restricted SC status to *Hindu* groups in 'unclean' occupations: their non-Hindu equivalents were not accorded this status and, therefore, could not benefit from reservation policies. However, subsequent extensions were made to this list for *Mazhabi Sikhs* (in 1956) and *neo-Buddhists* (in 1990).

pessimism that education would lead to employment; difficulty in getting credit from banks; the poor quality of public services in Muslim areas.

Another group of persons who were disadvantaged belonged to the 'Other Backward Classes' which was an intermediate group with the SC and the ST 'below', and the upper classes 'above', them. The 1980 report of the "Mandal" Commission recommended that, in addition to the 23 percent of various positions reserved for the SC and ST, a *further* 27 percent be reserved for the OBC.

Against this background, the hypothesis that is being considered under this heading is that certain groups, who have been historically excluded from the educational process, do not participate as fully as they might in higher education either because they do not see the value or the point of higher education or because their learning outcomes at school do not equip them for higher educational courses. In their study of Muslims in Bijnor, Jeffery and Jeffery (1997) argued that many Muslims regarded their relative economic weakness as stemming from their being excluded from jobs due to discriminatory practices in hiring. The belief that their sons would not get jobs then led Muslim parents to devalue the importance of education as an instrument of upward economic mobility.

Gender Disadvantage

Another source of disadvantage in an essentially patriarchal society like India's, is being a woman. Dreze and Kingdon (2001) observed that the decision to enrol a child at school/college may be viewed as a cost-benefit decision in which the present value of the expected flow of benefits from education is compared to the costs that must be incurred in order to secure such benefits. The costs are the direct costs of education (expenditure on books, fees), plus the indirect costs in terms of foregone earnings while the child is in education, while benefits are represented by the opportunities for higher earnings to which education gives rise. In a patriarchal society, in which women, after marriage, leave the parental home, so that responsibility for the care and maintenance of parents devolves on sons, the benefits of educating sons outweigh those of educating daughters.⁶ In terms of HEA, the 71st NSS informs us that 20.8% of women in the 18-22 years age group were in HEA while the corresponding figure for men was 22.6%.

⁶ Jeejeebhoy and Sathar (2001) in their discussion of women's autonomy in India observe that "the cultures of South Asia are gender-stratified characterised by hierarchical relations...in which the patriarch or his relatives have control over family members."

Economic Disadvantage

The third form of disadvantage is a lack of economic resources to pursue a course of education both on account of the direct costs of education and on account of the opportunity costs in terms of forgone earnings. Data from the 71st NSS round (after grossing up using the NSS-provided multipliers) shows that while the mean ‘monthly per capita consumption expenditure’ (MPCE) of households, to which the 18-22 year olds in the sample belonged, was ₹8,537, the mean annual total expenditure on education of those in HEA was ₹30,088. Even those in the lowest quartile of expenditure on higher education spent an average of ₹12,287.

The 71st NSS round informs that the mean household MPCE of 18-22 year olds who were ‘poor’ – defined as those living in households whose MPCE was in the lowest quintile – was ₹3,181 implying that, for such persons, their mean expenditure on higher education (₹12,287) was nearly one-third of their household annual median MPCE (₹38,172). More poignantly, it has been reported that an effect of the recent drought in the Marathwada district of the state of Maharashtra has been to force young people to abandon their plans of higher education because the drought-induced fall in family income no longer made it financially possible.⁷ So, under this hypothesis, several families might view higher education as a luxury good so that that, in straitened circumstances, when their resources are stretched to meet their usual needs, it becomes an expendable item.

Locational Disadvantage

Lastly, the issue of poverty and location in influencing decisions to enter higher education are not unrelated. Over 69% of the 18-22 year olds in the 71st NSS round lived in rural areas yet most the institutions of higher education were located in urban areas.⁸ Consequently, in order to be in HEA, students from rural areas had to travel longer distances than urban students: the 71st NSS round shows

⁷ Kavita Iyer, “In drought hit Marathwada, an early casualty: education”, *Indian Express*, 7 September 2015.

⁸ For example, as of September 2015, the state of Assam had 18 universities, most of them situated in the four towns of Guwahati, Dibrugarh, Jorhat, and Silchar. See the University Grant Commission’s list of universities in India: <http://www.ugc.ac.in/oldpdf/alluniversity.pdf>

that, for persons in HEA, the mean distance travelled from their homes to their educational institutions was 4.1 kilometres for rural residents and 3.7 kilometres for urban residents.⁹

So, all things considered, compared to their urban counterparts, rural students are likely to experience greater difficulty in accessing higher education. In addition, if higher education equips a person for, in the NSS's phrase, "regular salaried and wage employment", then opportunities for employment are likely to be fewer in rural, compared to urban areas and, therefore, demand for higher education is likely to be less among rural residents. Under both considerations – the limited supply of higher educational institutions but also limited demand for higher education - it is likely that the rates of HEA will be lower in rural than in urban areas.¹⁰

The above four factors may be regarded as 'macro' factors – in the sense that they relate to broad categories of respondents - affecting the likelihood of being in HEA. Underlying these macro factors could be 'micro' factors which affect *specific* individuals. Unfortunately, the data set (described in the next section) is silent on many of these micro issues. For example, there is no information on parental education or on school performance. The only personal information – which we use - is on marital status. Here our hypothesis is that while marriage may act as a barrier to pursuing higher education it would affect women more than men.

It is worth pointing out, however, that many of the forces that move young persons towards, or away from, higher education may work through group membership. Peer pressure on, and parental expectations of, 18-22 year olds will depend upon whether, say, they belong to an upper class, or to a SC, household. In the former situation, it is likely that their peers will also be considering higher education and that their (most likely, educated) parents will be encouraging them to do so. In the latter case, lack of familiarity with higher education, difficulties in accessing higher education, and ignorance of its value may mean that choosing higher education goes against the norms of peer behaviour and parental expectation.

⁹ Of course universities have constituent colleges in rural areas but these are likely to be fewer than colleges located in towns. Moreover, the largest growth has been in private universities and these have urban locations and are barred from having constituent colleges.

¹⁰ See Krishna (2013) on the importance of location.

3. The Data

The data for this study are from the 71st Round of the NSS and pertain to the period January-July 2014. The analysis is supplemented by data from the 64th round which pertains to the period July-2007-June 2008. Both the 71st and the 64th and NSS rounds, unlike the more ‘generalist’ rounds, are aimed at providing specific information on education.¹¹

The 71st (and the 64th) NSS rounds provided information about whether the respondents *between the ages of 5 and 29 years* were currently in attendance at a variety of educational levels from primary school upwards. From these levels, we chose as our point of focus those *currently attending graduate or post-graduate* courses, irrespective of subject, with the aim of acquiring a degree, diploma, or certificate through these courses. Persons were said to be in HEA if they were currently attending such (graduate or post-graduate) courses and not in HEA if they were not. Table 1 shows the distribution of persons in HEA by age: 79% of such persons in the 71st round, and 80% in the 64th round, were between 18 and 22 years of age.¹²

¹¹ At the outset of describing the data, it is again important to draw attention to the fact that all the results reported in this study are based upon grossing up the survey data using the observation-specific weights provided by the NSS for each of the surveys.

¹² There may have been persons older than 29 in HEA but these were not recorded by the NSS since it imposed an age ceiling of 29 years in collecting data on HEA,.

Table 1: The Distribution of Numbers in Higher Education Attendance in India, by Age*

Age→	16	17	18	19	20	21	22	23	24	25	26	27	28	29	Total
HEA numbers 71 st round →	498	14,965	48,989	49,269	61,957	37,882	32,154	16,576	10,540	9,447	4,216	2,541	2,487	841	292,364
% of total →	0.2	5.1	16.8	16.9	21.2	13.0	11.0	5.7	3.6	3.2	1.4	0.9	0.9	0.3	100.0
HEA numbers 64 th round →	583	7,333	24,050	23,337	31,434	17,623	13,165	7,413	5,007	2,802	1,587	874	912	269	136,387
% of total →	0.4	5.4	17.6	17.1	23.1	12.9	9.7	5.4	3.7	2.1	1.2	0.6	0.7	0.2	100.0

*16-29 years of age

Source: NSS 71st and 64th rounds, after applying sample weights.

The second item of particular interest to this study was the construction of the social groups with each person in the estimation sample being placed in one, and only one, of these groups. The NSS categorised persons by four social groups (Scheduled Tribes (ST); Scheduled Castes (SC); Other Backward Classes (OBC); and 'Other') and simultaneously by eight religion groups (Hindus; Islam; Christianity; Sikhism; Jainism; Buddhism; Zoroastrianism; 'Other'). Since Jains and Zoroastrians comprised less than 0.25% of the sample they are not separately identified in this study but included in the 'Other' category. The fact that Muslims, too, have their 'backward classes' and 'forward' classes, with a conspicuous lack of inter-marriage between the two groups, meant that it was sensible to separate Muslims into two groups: Muslims from the OBC and non-OBC Muslims.

Combining the NSS 'social group' and 'religion' categories, we subdivided households into the following groups which are used as the basis for analysis in this paper:

1. Scheduled Tribes (ST). These comprised 13.1% of the 65,923 households in the 71st NSS round and 9.5% of the grossed up NSS of 2,484,620 households.
2. Scheduled Castes (SC). These comprised 16% of the 65,923 households in the 71st NSS round and 18.9% of the grossed up NSS of 2,484,620 households. Over 90% of households in this category were Hindu.¹³
3. Non-Muslim Other Backward Classes (NMOBC). These comprised 32.7% of the 65,923 households in the 71st NSS round and 36.1% of the grossed up NSS of 2,484,620 households with 96% of these households being Hindu.
4. Muslim Other Backward Classes (MOBC). These comprised 6.4% of the 65,923 households in the 71st NSS round and 6.7% of the grossed up NSS of 2,484,620 households.¹⁴
5. Muslims who were not from the Other Backward Classes. They are, hereafter, referred to as Muslim upper classes (MUC) comprised 6.2% of the 65,923 households in the 71st NSS round and 5.7% of the grossed up NSS of 2,484,620 households.

¹³ This category also included some Muslim households. Since Muslims from the SC are not entitled to SC reservation benefits these Muslim SC households have been moved to the Muslim OBC category.

¹⁴ Including Muslim SC households (see previous footnote).

6. Non-Muslim upper classes (NMUC). These comprised 25.7% of the 65,923 households in the 71st NSS round and 23.1% of the grossed up NSS of 2,484,620 households: over 90% of the households in this category were Hindu.

Table 2 shows the GTAR values (defined earlier) by social group.¹⁵ This shows that the GTAR increased dramatically in the six years between the 71st and 64th rounds from, in aggregate, 14.2% to 27.6%, with all groups experiencing a considerable rise in their GTAR. Paralleling this increase of 14 percentage points in the GTAR was an increase in the number of universities in India from 436 in 2009 to 733 in 2015 with the largest growth being in the number of private universities. Of the 733 universities in 2015: 46 were central universities; 336 were state universities; 127 were ‘deemed’ universities; and 224 were private universities.¹⁶

The Muslim OBC had the lowest GTAR (14.8%) in the 71st round, followed by a GTAR of 15.2% for the ST, 15.7% for the MUC, 20.4% for the SC, 29.6% for the non-Muslim OBC, and 46.3% for the NMUC. A striking feature of the GTAR for the ST is that its value is very different depending upon whether one is considering Hindu or Christian ST: the GTAR was 28.4% for the Christian ST but only 13.3% for the Hindu ST.

Between the two rounds, the GTAR for the Muslim OBC increased from 7.4% to 14.8%, for the Muslim UC from 7.8% to 15.7%, for the ST from 6% to 15.2%, for the SC from 9.5% to 20.4%, for the non-Muslim OBC from 13.2% to 29.6%, and for the non-Muslim UC from 27.4 % to 46.3%.

¹⁵ Strictly speaking GTAR should include in the numerator all persons in HEA. The NSS data, however, only provides information about persons up to the age of 29 years in HEA.

¹⁶ See the University Grant Commission’s list of universities in India: <http://www.ugc.ac.in/oldpdf/alluniversity.pdf>. ‘Central’ and ‘State’ universities were funded by, respectively, the central government and the state governments; ‘deemed’ universities were those Institutes which the University Grants commission regarded as the equivalent of universities and which, therefore, were allowed to award degrees; private universities were approved by the UGC but were not allowed to have off-campus colleges.

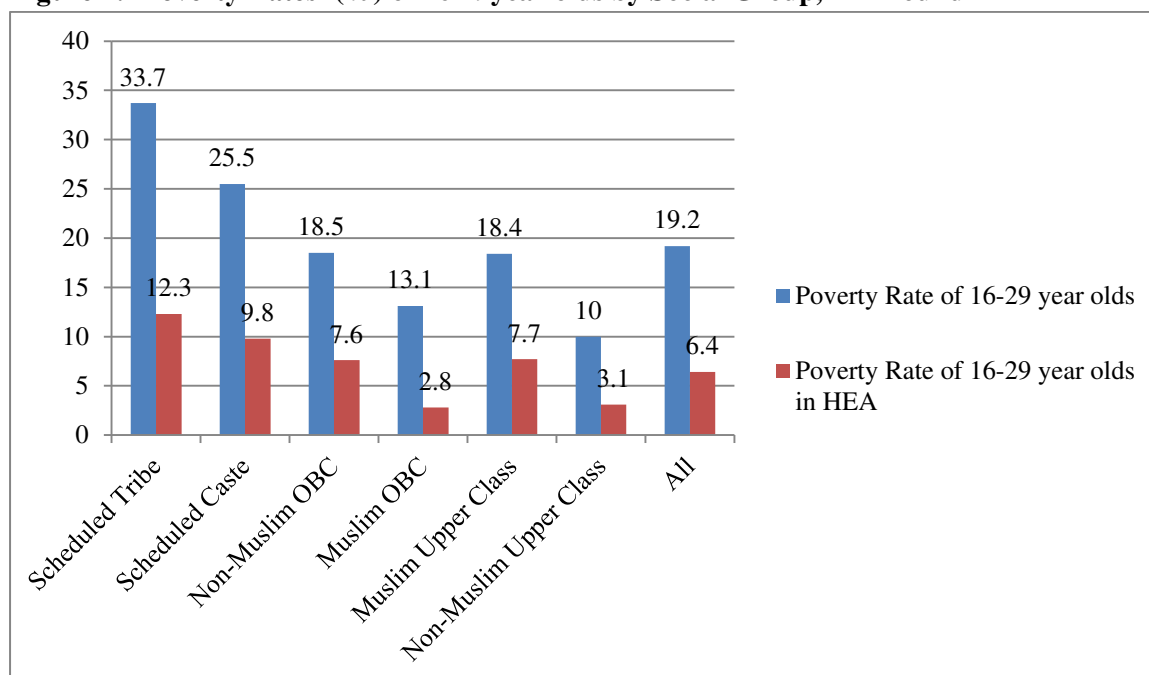
Table 2: Gross Tertiary Attendance Rate in India, by Social Group: 71st and 64th rounds

71 st round							
	Scheduled Tribe	Scheduled Caste (excl. Muslims)	Non-Muslim OBC	Muslim OBC (incl. SC Muslims)	Muslim Upper Class	Non-Muslim Upper Class	Total
Numbers in HEA 16-29 years	16,217	43,009	108,262	13,240	11,134	100,503	292,364
Total aged 18-22	106,062	211,179	365,769	89,299	71,141	216,951	1,060,401
GTAR %	15.2	20.4	29.6	14.8	15.7	46.3	27.6
64 th round							
	Scheduled Tribe	Scheduled Caste	Non-Muslim OBC	Muslim OBC	Muslim Upper Class	Hindu Upper Caste	Total
Numbers in HEA 16-29 years	4,632	18,231	45,265	4,463	6,082	57,715	136,387
Total aged 18-22	77,006	192,795	343,425	60,025	77,735	210,364	961,350
GTAR %	6.0	9.5	13.2	7.4	7.8	27.4	14.2

Source: NSS 71st and 64th rounds, after applying sample weights

The final feature relating to organising the data is an economic measure of deprivation. In two seminal papers, Basu (2001, 2006) proposed a *quintile axiom*, according to which “we should focus attention on the per-capita income of the poorest 20% of the population (‘quintile income’) and the growth rate of the per-capita income of the poorest 20% (‘quintile growth’) (Basu, 2001, p. 66). Using this axiom, we constructed quintiles of household MPCE over *all* the households in the 71st round and all the households in the 64th NSS round; following that, we defined a person as being ‘poor’ if his/her household’s MPCE was in the bottom 20% of the distribution of MPCE. *Ipso facto* a person was not poor (‘non-poor’) if his/her household’s MPCE was in the upper 80% of the distribution. The poverty rates (that is, the proportion of person in the group who were living in poor households) of persons in the estimation sample in the different social groups are shown in Figure 1 for two groups: (i) persons who were between the ages of 16-29 years and (ii) persons who were between the ages of 16-29 years *and* in HEA. As Figure 1 makes clear the poverty rates of the first group were considerably higher than that of the second group.

Figure 1: ‘Poverty Rates’ (%) of 16-29year olds by Social Group, 71st Round



Defined as the percentage of persons in relevant group living in households in MPCE in the lowest decile

Source: Own Calculations from NSS 71st round, after applying sample weights

4. Estimation Results

The previous section referred to four sources of overlapping disadvantage - ‘social group’ disadvantage; ‘gender’ disadvantage; ‘economic’ disadvantage; and ‘locational’ disadvantage – in terms of participating in higher education. In the context of this study, a natural question to ask was whether the effect of the *social group* of persons, on their probabilities of HEA, varied according to their: (i) gender; (ii) poverty status; (iii) location (rural/urban), with controls being imposed for the state of residence. A further question was whether marital status affected women’s chances of being in higher education more than it did men’s.

In practical terms, the interdependency between these four factors can be modelled through *interaction effects*. These effects are used to examine whether the effect of a specific variable (say social group) on the outcome probability varies according to values of another variable (say, gender).¹⁷ An Appendix to the paper details how these interaction effects are incorporated into the specification of the HEA equation and how the equation’s coefficients should be interpreted. The coefficients of the HEA equation (detailed as equation (3) of the Appendix) were estimated over the

¹⁷ For example, does being male or female affect the probabilities of being in HEA differently for SC and NMUC? In terms of being in HEA, do persons from different groups respond differently to: belonging to poor households; to living in rural locations?

sample of 36,692 persons in the NSS 71st round *who were between the ages of 18 and 22 years (inclusive) with each observation weighted by its NSS-provided weight.*¹⁸ Following the advice contained in Long and Freese (2014), the results from the estimated equation are presented in Table 3 in the form of *predicted probabilities* from the estimated logit coefficients and not in terms of the estimates themselves. This is because the logit estimates themselves do not have a natural interpretation – they exist mainly as a basis for computing more meaningful statistics and, in this case, these are the predicted probabilities.¹⁹

Consequently, the logit estimates,²⁰ were used to derive the *average likelihood* of HEA (*currently attending a higher education course*) for persons aged 18-22 years from the six different social groups – ST, SC, NMOBC, MOBC, MUC, and NMUC. These probabilities, shown in column 2 of Table 3, are the predicted probabilities of being in HEA of 18-22 year olds from the different social/gender/poverty/location groups.

¹⁸ Of the 36,692 persons, aged 18-22, in the *estimation sample*: 13,010 persons (35.5%) were in HEA; 16,286 persons (44.4%) were women; 4,009 persons (10.9%) were ‘poor’, in the sense that their household’s MPCE was in the lowest quintile; and 20,585 persons (56.1%) lived in rural areas. In terms of their social group, 13.3% were from the ST, 16.0% were from the SC, 31.8% were non-Muslim OBC, 7.6% were Muslim OBC, 6.8% were from the Muslim upper classes, and 24.6% were from the non-Muslim upper classes.

¹⁹ Two points should be emphasised in respect of the probabilities shown in Table 3: (i) in studying the relationship between social group and HEA, the effects of other variables, like gender, poverty status, and sector and state of residence were controlled for and, consequently, the predicted probabilities will differ from the sample proportions; (ii) in computing these probabilities shown in Table 3 *all* the interaction effects – in this case, the interactions of gender, poverty status, sector of residence and social group as well as the separate interaction of gender, marital status, and social group, as set out in equation (3) of the Appendix – were taken into account.

²⁰ Which, for reasons of space, are not shown but may be obtained on request from the author.

Table 3: Predicted and Marginal Probabilities of Currently Being in Higher Education Attendance by Persons Aged 18-22 Years *

1	2	3	4	5	6
Conditioning Variable	Probability	Marginal Probability	SE	z value	Pr> z
All Respondents: 36,692					
Scheduled Tribe [4,871]	0.168	-0.166	0.002	-91.4	0.00
Scheduled Caste [5,853]	0.169	-0.165	0.001	-126.7	0.00
Non-Muslim Other Backward Classes [11,674]	0.231	-0.102	0.001	-82.4	0.00
Muslim Other Backward Classes [2,796]	0.108	-0.225	0.001	-160.4	0.00
Muslim Upper Classes [2,490]	0.111	-0.222	0.002	-145.4	0.00
Non-Muslim Upper Classes [9,008]	0.334				
All Male Respondents: 20,406					
Scheduled Tribe [2,542]	0.185	-0.149	0.003	-59.5	0.00
Scheduled Caste [3,241]	0.179	-0.155	0.002	-86.3	0.00
Non-Muslim Other Backward Classes [6,549]	0.241	-0.093	0.002	-55.7	0.00
Muslim Other Backward Classes [1,539]	0.130	-0.204	0.002	-103.1	0.00
Muslim Upper Classes [1,373]	0.114	-0.220	0.002	-106.8	0.00
Non-Muslim Upper Classes [5,162]	0.334				
All Female Respondents: 16,286					
Scheduled Tribe [2,329]	0.149	-0.184	0.002	-74.2	0.00
Scheduled Caste [2,612]	0.156	-0.177	0.002	-95.0	0.00
Non-Muslim Other Backward Classes [5,125]	0.220	-0.113	0.002	-64.1	0.00
Muslim Other Backward Classes [1,257]	0.084	-0.249	0.002	-128.1	0.00
Muslim Upper Classes [1,117]	0.108	-0.225	0.002	-101.4	0.00
Non-Muslim Upper Classes [3,846]	0.333				
All non-Poor Respondents: 32,683					
Scheduled Tribe [4,016]	0.186	-0.189	0.002	-91.8	0.00
Scheduled Caste [4,949]	0.188	-0.187	0.001	-125.6	0.00
Non-Muslim Other Backward Classes [10,454]	0.254	-0.121	0.001	-86.9	0.00
Muslim Other Backward Classes [2,536]	0.126	-0.249	0.002	-152.5	0.00
Muslim Upper Classes [2,222]	0.126	-0.249	0.002	-141.3	0.00
Non-Muslim Upper Classes [8,506]	0.375				
All Poor Respondents: 4,009					
Scheduled Tribe [855]	0.073	-0.049	0.003	-14.4	0.00
Scheduled Caste [904]	0.066	-0.055	0.003	-21.0	0.00
Non-Muslim Other Backward Classes [1,220]	0.115	-0.006	0.003	-2.1	0.04
Muslim Other Backward Classes [260]	0.027	-0.094	0.003	-33.2	0.00
Muslim Upper Classes [268]	0.035	-0.086	0.003	-28.8	0.00
Non-Muslim Upper Classes [502]	0.121				
All Rural Respondents: 20,585					
Scheduled Tribe [3,443]	0.125	-0.160	0.002	-88.5	0.00
Scheduled Caste [3,717]	0.152	-0.133	0.002	-86.0	0.00
Non-Muslim Other Backward Classes [6,853]	0.195	-0.090	0.001	-60.3	0.00
Muslim Other Backward Classes [1,252]	0.105	-0.180	0.002	-100.8	0.00
Muslim Upper Classes [1,240]	0.084	-0.201	0.002	-108.5	0.00
Non-Muslim Upper Classes [4,080]	0.285				
All Urban Respondents: 16,107					
Scheduled Tribe [1,428]	0.256	-0.173	0.004	-43.1	0.00
Scheduled Caste [2,136]	0.201	-0.229	0.002	-96.4	0.00
Non-Muslim Other Backward Classes [4,821]	0.304	-0.125	0.002	-56.9	0.00
Muslim Other Backward Classes [1,544]	0.118	-0.312	0.002	-136.4	0.00
Muslim Upper Classes [1,250]	0.166	-0.263	0.003	-96.3	0.00
Non-Muslim Upper Classes [4,928]	0.429				

* Results pertain to a logit model estimated over 36,692 persons (see equation (3) of the Appendix for its specification) with each observation weighted by its NSS-provided weight.

Source: Own calculations from NSS 71th Round (January - July 2014)

The first panel (labelled: ‘all respondents’) of Table 3 shows, for example, that the probabilities of HEA of ST and NMUC persons were, respectively, 16.8% and 33.4%. The ST probability was computed by assuming that *all* the 36,692 persons in the sample were ST, with the values of the other variables unchanged at their observed values. Applying the logit estimates to this ‘synthetic’ case, the probability of HEA of persons from the ST was estimated as 16.8%. Similarly, the NMUC probability of HEA was computed by assuming that *all* the 36,692 persons in the sample were NMUC, with the values of the other variables unchanged at their observed values. Applying the logit estimates to this ‘synthetic’ case yielded the estimated probability of HEA for persons who were NMUC as 33.4%. Since the *only* difference between the two synthetic cases was that in the first all the 36,692 persons in the sample were regarded as ST and, in the second, they were all regarded as NMUC the difference between the two predicted probabilities (16.8% and 33.4%) was *entirely* the result of a difference between being ST and NMUC.

The *marginal probability* associated with a variable refers to the *change* in the outcome probability consequent upon a unit change in the value of the variable, *the values of the other variables remaining unchanged*. For discrete variables (as, indeed, are all the variables reported above), a unit change in the value of a variable refers to a move from the *reference* category to the category in question, the values of the other variables remaining unchanged.²¹ Dividing these marginal probabilities (in column 3 of Tables 3) by their corresponding standard errors (column 4), yields the z-value associated with these marginal probabilities (column 5).²²

The results in Table 3 show four main factors which affected the predicted probabilities of HEA of 18-22 year olds: social group; gender; poverty; and rural/urban location. In terms of social group, there was a clear hierarchy, with the predicted probability of HEA being highest for 18-22 year olds from the non-Muslim upper classes (33.4%), followed by the non-Muslim OBC (23.1%), and

²¹ So, the *marginal probability* associated with ST persons is defined as the *difference* between ST and NMUC (the reference category) persons in their predicted probabilities of HEA. For the first panel (labelled: all respondents) of Table 3, this marginal probability was 16.8-33.4 = -16.6 percentage points (pp) which is shown in column 3 of Table 3 as -0.166.

²² For ST persons $z=91.4$ and, as the p -value in column 6 of Table 3 suggests, this marginal probability was significantly different from zero: the (average) probability of HEA of persons, aged 18-22 years, from the ST (16.8%) was *significantly* lower than that of their NMUC counterparts (33.4%).

followed by the Scheduled Castes and the Scheduled Tribes (16.8% and 16.9%, respectively) with Muslims bringing up the rear (11.1% for the MUC and 10.8% for the MOBC). For every social group these probabilities were significantly lower than that for the reference group of the NMUC.²³

As Table 4 shows, in terms of gender, the 2.1 point gap in the predicted probability of men and women, aged 18-22 years, being in HEA (men 22.7%, women 20.6%) was significantly different from zero and a similar gap was also significant for the social groups *except* for the MUC and NMUC for which there was very little difference in the predicted probabilities of men and women, aged 18-22 years, being in HEA. In terms of poverty and location, the poor in every social group were significantly less likely to be in HEA than their non-poor counterparts and rural persons in every social group were significantly less likely to be in HEA than their urban counterparts. These gaps are detailed in Table 4. In the context of the predicted probabilities of being in HEA, it is striking that while the NMUC had a non-existent gender gap it had the largest poverty gap (Table 3: non-poor 37.5%; poor 12.1%) and the largest location gap (Table 3: urban 42.9%; rural 28.5%) of all the six social groups.

**Table 4: Poverty and Location Gaps in HEA, by Social Group
(Percentage Points)**

	Gender Gap	Poverty Gap	Location Gap
All Respondents: 36,692	2.1*	15.1*	9.7*
Scheduled Tribe	3.6*	11.3*	13.1*
Scheduled Caste	2.3*	12.3*	4.9*
Non-Muslim Other Backward Classes	2.1*	13.8*	10.9*
Muslim Other Backward Classes	4.7*	9.9*	1.3*
Muslim Upper Classes	0.6	9.1*	8.2*
Non-Muslim Upper Castes	0.1	25.4*	14.4*

* Significant at 5% level

Source: Own calculations from NSS 71th Round (January - July 2014).

The results in terms of gender were nuanced by whether the persons were married or not. Being married, imposed a constraint upon HEA: the predicted probability of HEA, of persons aged 18-22 years, was lower for married than for 'never married' persons (17.1% compared to 23.4%). In the estimation sample comprising 36,692 persons aged 18-22 years, 29% of women and 6.7% of men were married, the remainder being 'never married'. Consequently, for persons in the 18-22 age group, marriage imposed a greater constraint upon women than it did upon men.

²³ The probabilities of HEA were not significantly different between the ST and the SC but they were significantly different between the SC and the MUC and between the MUC and the MOBC.

This constraint was compounded by the fact that, as Table 5 shows, whether married or never married, the predicted probability of being in HEA was greater for men than for women. For example, as Table 5 shows, the predicted probability of married persons being in HEA was 17.1% for men and 16.6% for women while the predicted probability of never married persons being in HEA was 24.1% for men and 22.1% for women.

Table 5: The Probabilities of Men and Women, Aged 18-22, Being in HEA, by Social Group and Marital Status

	All		Men		Women	
	Married	Never Married	Married	Never Married	Married	Never Married
All Respondents:	0.171	0.234	0.190	0.241	0.166	0.221
Scheduled Tribe	0.129	0.182	0.158	0.195	0.121	0.159
Scheduled Caste	0.139	0.180	0.160	0.187	0.133	0.165
Non-Muslim Other Backward Classes	0.195	0.245	0.213	0.252	0.189	0.232
Muslim Other Backward Classes	0.081	0.118	0.112	0.137	0.073	0.088
Muslim Upper Classes [2,484]	0.090	0.119	0.095	0.121	0.088	0.116
Non-Muslim Upper Castes [8,998]	0.290	0.350	0.292	0.350	0.289	0.349

Source: Own calculations from NSS 71th Round (January - July 2014).

5. The Quality of Higher Education

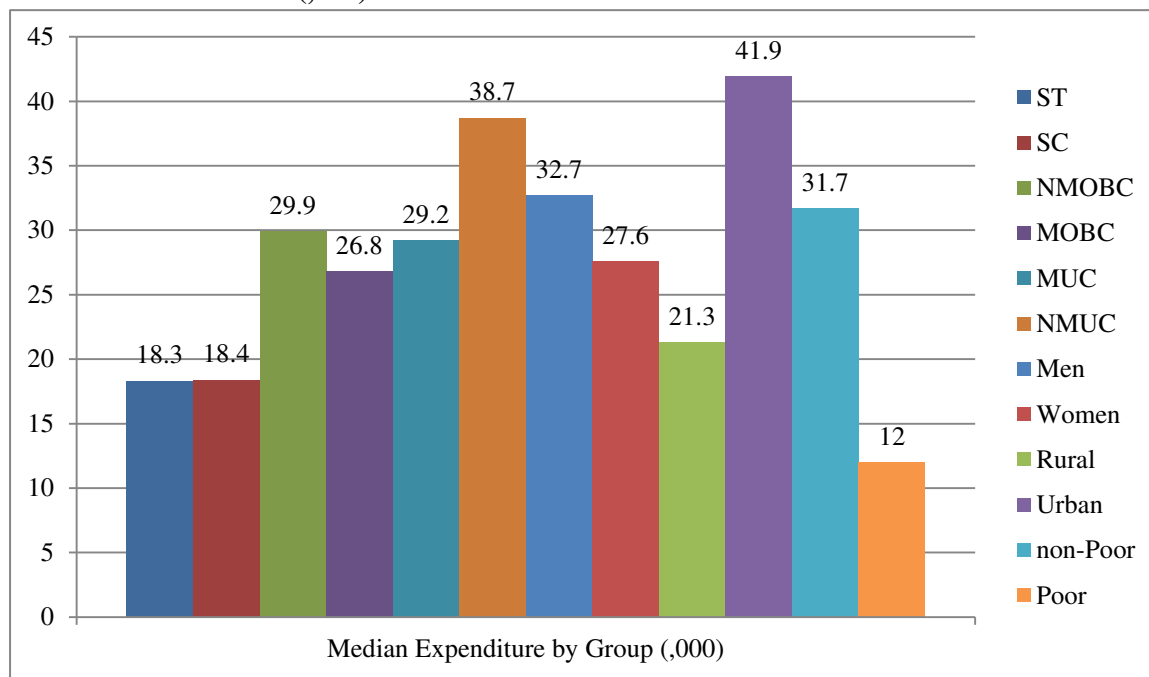
The preceding material focused on the likelihood of being in HEA. But this begs the question of the *quality* of higher education that students received. Both the 71st NSS and the 64th NSS rounds provide information on items that could reasonably be considered as *proxies* for ‘quality’. One such item is the *total* expenditure on education by persons in HEA who were aged 18-22 years. For example, a clue to differences in the quality of higher education received by women and men is provided by comparing the mean amounts that the two groups spent on such education: the mean total expenditure of women and men in HEA was, respectively, ₹27,613 and ₹32,710. So, while gender disadvantage in terms of attendance might be small, there was clear gender disparity in terms of expenditure on higher education by those in HEA.²⁴

Similarly, there was a clear hierarchy of expenditure on higher education by those in HEA from the different social groups. As Figure 2 shows, for persons in HEA, those from the NMUC spent the most (mean: ₹38,677), followed by the NMOBC (₹29,897) while those from the SC, the ST, MOBC, and MUC (respectively, ₹18,335, ₹18,406, ₹26,751, ₹29,164) spent the least. Mean expenditure on higher education by those in the rural sector was considerably lower than that spent by

²⁴ Note that we are comparing *all* men and women in HEA and, by the constraint of the NSS, the upper limit for this is 29 years.

urban residents (₹21,289 versus ₹41,927) and mean expenditure on higher education by those in poor households was considerably lower than that by those in non-poor households (₹12,010 versus ₹32,656).

Figure 2: Median Expenditure on Higher Education by those aged 18-22 years and in Higher Education Attendance (,000)



Source: Own Calculations from NSS 71th Round (January - July 2014), after applying sample weights

English as a Medium of Instruction

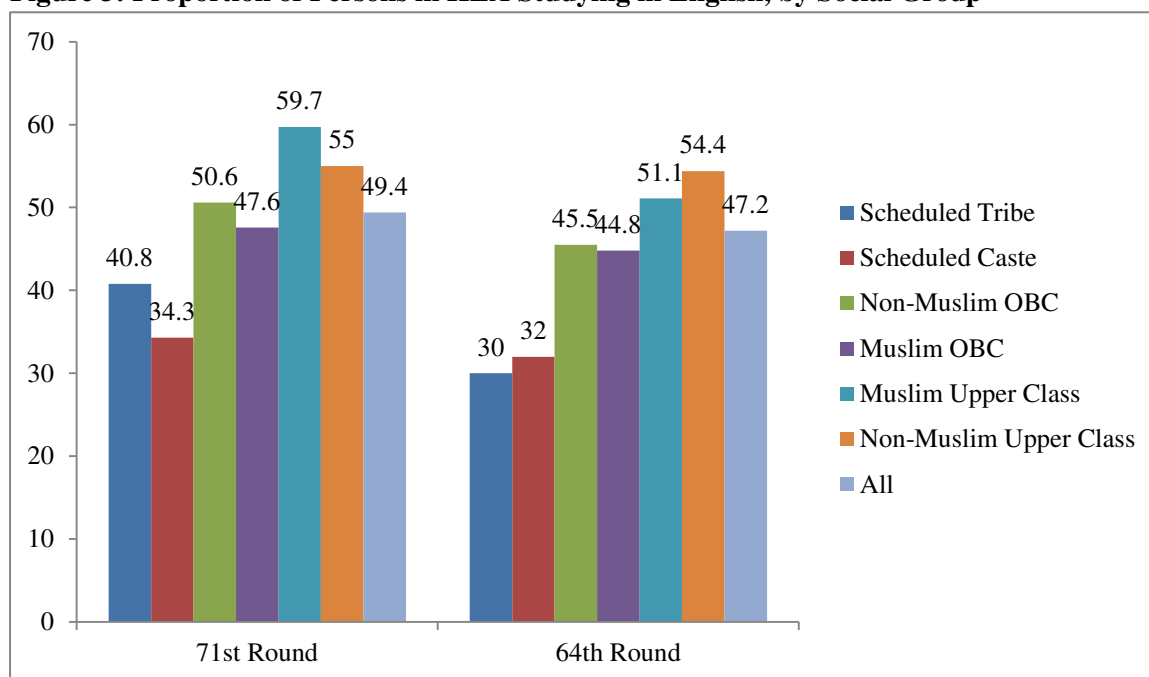
Another proxy of quality is the medium of instruction (MoI) in higher education. Without prejudice, one could regard higher education courses with English as the MoI (hereafter, ‘studying in English’) as offering better employment prospects (and, therefore, of higher ‘quality’) than courses delivered in other languages. Figure 3 shows, for the 71st and 64th rounds, the proportion of persons in the different social groups in HEA, who were studying a course with English as the MoI.

In the 64th round (2007-08), 47% of all persons in HEA, were studying in English; by the 71st round, this proportion had risen to 49% and this rise in the popularity of studying in English was particularly experienced by persons from the ST: the proportion of ST persons in HEA studying in English rose from 30% in the 64th Round to nearly 41% in the 71st Round.

Persons from the ST divide into two groups: Hindu ST and Christian ST (respectively, 86% and 9% of the grossed up NSS 71st round 1,091,429 ST persons). However, when it came to the

16,216 ST persons in HEA, 75% were Hindu and 17% persons were Christian. Of the latter group (Christian ST), 82% were studying in English (up from 78% in the 64th Round) while of the former group (Hindu ST) only 30% were studying in English (up from 21.7% in the 64th Round). The four point rise in the proportion of Christian STs in HEA studying in English (78% to 82%), and the eight point rise in the proportion of Hindu STs in HEA studying in English (22% to 30%), between the 64th and the 71st Round, largely explains the rise, from 30% to 41% in the aggregate proportion of ST persons in HEA studying in English. In addition, between the two Rounds, the proportion of Christians, in the total of ST persons in HEA, increased from 14% to 17% while the proportion of Hindus, in the total of ST persons in HEA, decreased from 82% to 75%.

Figure 3: Proportion of Persons in HEA Studying in English, by Social Group



Source: Own Calculations from NSS 71th Round (January - July 2014), after applying sample weights

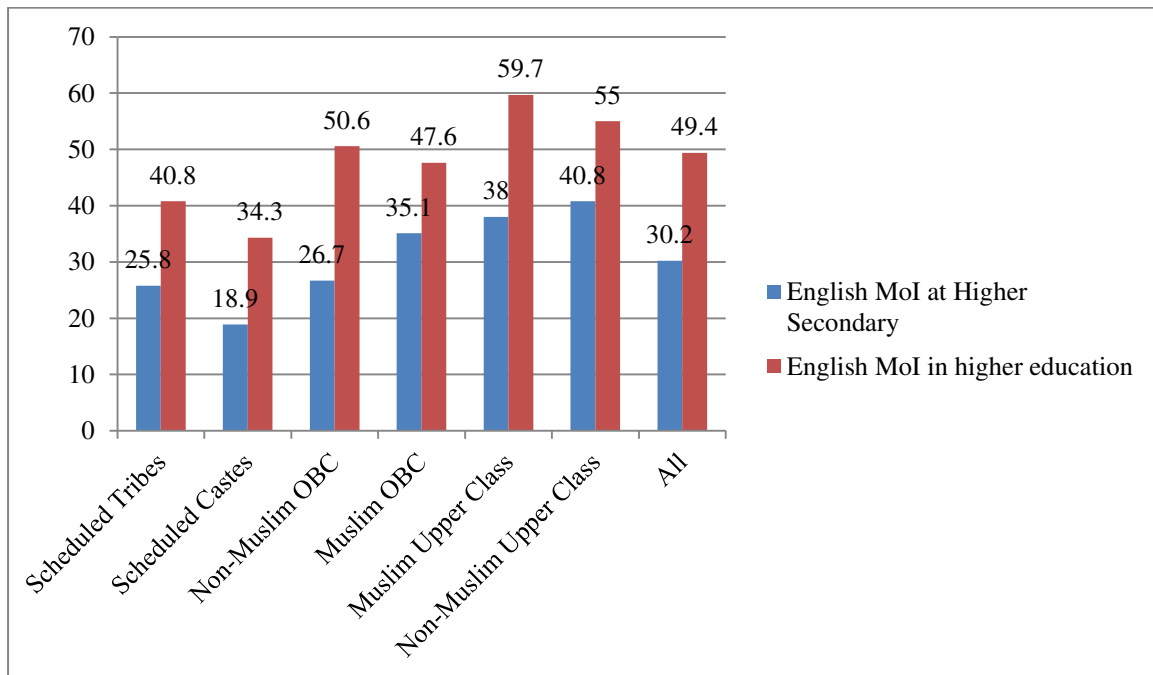
Coping with English as the Medium of Instruction in Higher Education

The fact that, as Figure 3 shows, nearly half of those in HEA in 2014 were taking courses which were delivered in English raises the question of how well they coped with the language since for many students higher education provided their first experience of studying in English, their medium of instruction prior to HEA being in another language.

This failure to cope with English was highlighted anecdotally when, in July 2015, the Indian Institute of Technology at Roorkee failed 72 students after their first year of studies which, in turn, was supposed to entail their automatic expulsion from the Institute. Of these 72 students, 90% were from the ‘reserved’ categories (that is, groups for whom a certain proportion of places were reserved under affirmative action policies): Scheduled Tribes, Scheduled Castes, and Other Backward Classes. Explaining this failure one of the students said: “English is our big problem. We are from Hindi-medium schools and then we come to the campus and realise it is all high-level English. We see students speaking English, asking questions in English and we can do none of that. Our confidence drains away” (Vishnu, 2015).

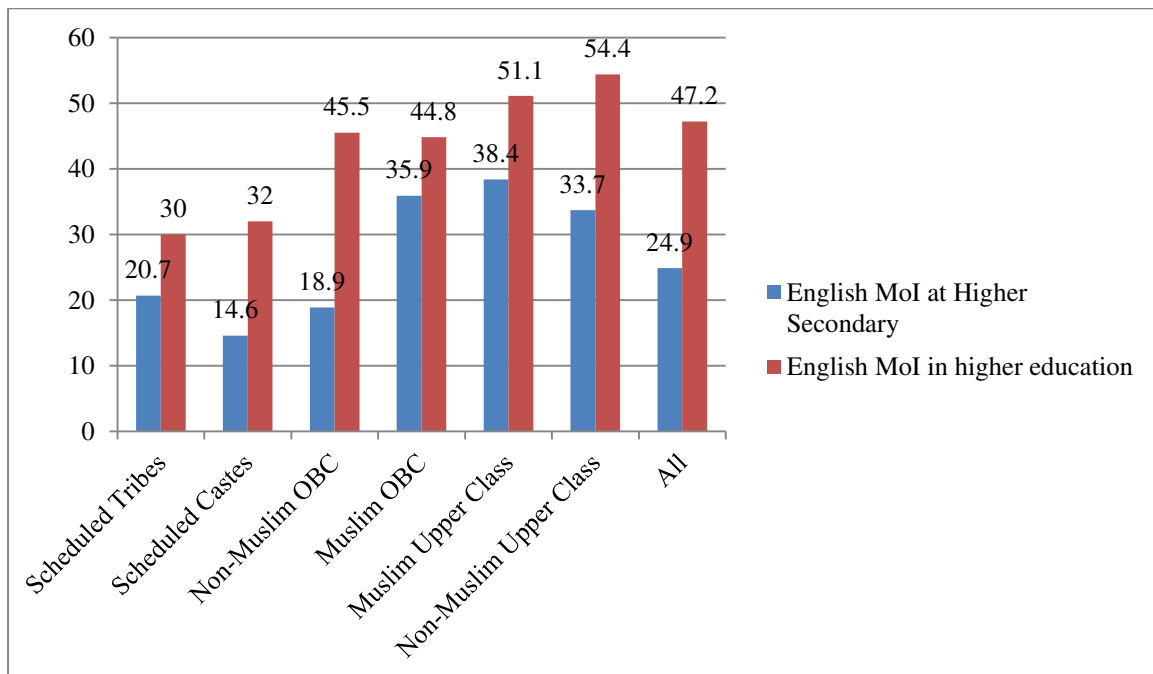
These problems are likely to be exacerbated as an increasing proportion of students, who had done their schooling in Hindi or a regional language, choose to do their higher education in English. Figures 4 and 5 compare by social group, for, respectively, the 71st and 64th Rounds, the proportion of persons studying in English at Higher Secondary – which is the jumping off point for higher education – with the proportion of persons studying in English in higher education. For each of the six groups there is a considerable gap between the proportions of students studying in English at these two levels but the gap is largest for the Scheduled Castes suggesting the problems of coping with English as a medium of instruction in higher education may not have diminished much between the 64th and 71st Rounds. The proportion of persons studying for higher secondary in English has increased from 25% to 30% but the proportion of persons in higher education studying in English has also increased from 47% to 49%.

Figure 4: Proportion of Persons studying in English at Higher Secondary and in Higher Education: NSS 71st Round



Source: Own Calculations from NSS 71th Round (January - July 2014), after applying sample weights

Figure 5: Proportion of Persons studying in English at Higher Secondary and in Higher Education: NSS 64th round



Source: Own Calculations from 64th Round (January - July 2008), after applying sample weights

6. What, and Where, do Young Persons in Higher Education Study

The NSS 71st Round also provides information on the broad subject categories in which persons, aged 18-22 years, in HEA are enrolled and also the type of institutions – government, private aided, private unaided - in which they are enrolled. Table 6 shows that, in 2014, the largest proportion (39%) of persons in HEA were enrolled on Humanities courses (down from 43% in 2008), followed by 17% in Commerce, 16% in Science, and 15% in Engineering (up from, respectively, 16%, 12%, and 10% in 2008), and 3% in IT-type courses (down from 9% in 2008). So, in summary, the shift in subject popularity between 2008 and 2014 has been the reduction in the popularity of Humanities and IT and the increased popularity of Science and Engineering. The proportion of persons in Medicine and in Business did not change greatly between 2008 and 2014. As Table 6 also shows, the groups most likely to study Humanities were the ST and the SC, women, the poor, and those living in rural areas while the groups most likely to study Engineering were the NMOBC and the NMUC, men, the non-poor, and those living in urban areas.

Table 6: Courses of Study in Higher Education by Social Group, Gender, Poverty, and Location NSS 71st Round*

	Percentage of Students in the Group Studying the Subject								
	Humanities	Science	Commerce	Medicine	Engineering	Management	IT	Other	Total
Scheduled Tribe	49.5	14.1	18.4	1.3	6.8	0.8	2.6	6.5	100
Scheduled Caste	52.7	12.9	14.9	1.5	8.7	1.5	2.2	5.7	100
Non-Muslim OBC	36.5	19.7	14.7	2.4	16.7	2.1	2.7	5.2	100
Muslim OBC	35.7	19.5	16.4	2.0	13.1	3.5	3.2	6.7	100
Muslim Upper Class	47.7	13.3	12.9	4.4	10.8	2.5	1.7	6.7	100
Non-Muslim Upper Class	38.7	15.9	17.0	2.4	14.5	2.5	3.3	5.8	100
Men	35.0	15.8	17.2	1.5	19.0	2.5	3.4	5.6	100
Women	43.3	16.0	16.7	3.5	8.9	2.5	3.2	5.9	100
Poor	56.4	16.6	13.1	0.5	5.2	1.2	1.3	5.7	100
Non-Poor	37.5	15.9	17.2	2.6	15.1	2.5	3.4	5.8	100
Rural	50.2	17.1	13.3	2.0	8.5	1.3	2.5	5.3	100
Urban	24.3	14.5	21.7	3.0	22.0	3.9	4.4	6.4	100
Total	38.7	15.9	17.0	2.4	14.5	2.5	3.3	5.8	100

* Aged 18-22 years

Source: Own Calculations from the NSS 71th Round (January - July 2014), after applying sample weights

Table 7 shows the types of higher educational institutions (HEI) at which persons, aged 18-22 years, in HEA were enrolled. In 2014, 41% were in government HEI (down from 49% in 2008), 26% were in private-aided HEI (down from 32% in 2008), and 33% were in private-unaided HEI (up from 19% in 2008). So, the biggest change between 2008 and 2014, in the type of HEI attended, has been that while in 2008 only one in five persons, aged 18-22 years, in HEA was enrolled in a private-unaided HEI, by 2014 this number had risen to one in three.

Table 7 also shows that government HEI were particularly popular with ST and SC students with nearly 60% of ST students (down from 63% in 2008) and 51% of SC students (down from 54% in 2008), aged between 18-22 years and in HEA, enrolled in government HEI. They were also popular with poor students and rural students with 58% of poor students (up from 56% in 2008) and 46% of rural students (down from 50% in 2008), aged between 18-22 years and in HEA, enrolled in government HEI.

On the other hand, private unaided HEI were particular popular with NMOBC and NMUC students with 39% of NMOBC students (up from 22% in 2008) and 32% of NMUC students (up from 20% in 2008), aged between 18-22 years and in HEA, enrolled in private unaided HEI. They were also popular with non-poor students and urban students with 34% of non-poor students (up from 20% in 2008) and 35% of urban students (down from 22% in 2008), aged between 18-22 years and in HEA, enrolled in private unaided HEI.

Table 7: Type of Higher Education Institution Attended, by Social Group, Gender, Poverty, and Location, NSS 71st Round*

	Government	Private Aided	Private Unaided	Total
Scheduled Tribe	59.5	21.3	19.3	100
Scheduled Caste	50.6	21.7	27.7	100
Non-Muslim OBC	35.9	24.9	39.3	100
Muslim OBC	43.6	28.1	28.3	100
Muslim Upper Class	55.7	22.1	22.2	100
Non-Muslim Upper Class	37.7	30.4	31.9	100
Men	42.4	24.0	33.6	100
Women	39.4	28.8	31.9	100
Poor	57.8	23.0	19.2	100
Non-Poor	39.9	26.4	33.7	100
Rural	45.7	23.5	30.8	100
Urban	35.2	29.5	35.3	100
Total	41.1	26.2	32.8	100

* Aged 18-22 years

Source: Own Calculations from the NSS 71th Round (January - July 2014), after applying sample weights

As Table 8 shows, in 2014, 77% of the student body, aged 18-22 years, in private unaided HEI were drawn from just two groups: 44% from the NMOBC and 33% from the NMUC. In 2008, 82% of the student body, aged 18-22 years, in private unaided HEI were drawn from these two groups: 38% from the NMOBC and 44% from the NMUC. Private aided HEI were also largely the preserve of the NMOBC and NMUC: in 2014, 75% of the student body, aged 18-22 years, in private aided HEI were drawn from these two groups: 35% from the NMOBC and 40% from the NMUC. In 2008, 76% of the student body, aged 18-22 years, in private unaided HEI were drawn from these two groups: 39% from the NMOBC and 37% from the NMUC.

Table 8: The Social Group Composition of Higher Educational Institutions, 71st and 64th Rounds*

	71 st Round: 2014			64 th Round: 2008		
	Government	Private Aided	Private Unaided	Government	Private Aided	Private Unaided
Scheduled Tribe	8.0	4.5	3.2	3.9	2.2	2.3
Scheduled Caste	17.9	12.1	12.3	14.8	14.4	7.3
Non-Muslim OBC	32.5	35.3	44.5	29.7	37.1	38.3
Muslim OBC	5.2	5.3	4.2	3.2	3.3	4.9
Muslim Upper Class	5.0	3.1	2.5	4.4	4.4	3.6
Non-Muslim Upper Class	31.5	39.8	33.3	44.0	38.6	43.7
Total	100	100	100	100	100	100

* Aged 18-22 years

Source: Own Calculations from the NSS 71th Round (January - July 2014), after applying sample weights

6. How Much of Inequality in Higher Education Attendance Can We Explain?

The analysis of the preceding section highlighted three factors which affected the likelihood of a person being in HEA: gender, poverty status, and sector. In turn, each of these factors interacted with the person's social group (Scheduled Tribe, Scheduled Caste, non-Muslim OBC, Muslim OBC, Muslim upper class, and Hindu upper castes) to produce the observed likelihood of HEA. The issue that is analysed in this section, using the tools of inequality decomposition, is the relative contribution of these factors to inter-personal inequality in the likelihood of 18-22 year olds being in HEA.

The estimated HEA logit equation, specified in equation (3), predicts, for each of the 36,692 persons in the estimation sample, the probability of HEA, conditional upon the relevant values of the determining variables (social group, gender, poverty status, sector, and state) for the person. Armed with a knowledge of these individual probabilities, one can estimate how much of the overall inequality in these probabilities can be explained by a particular factor. For example, how much of the inequality in the 36,692 probabilities of HEA can be accounted for by differences in: social group, gender, poverty status, and sector?

This section provides an answer to this question, using the methodology of *inequality decomposition*. This attempts to decompose overall inequality into its constituent parts: between-group inequality and within-group inequality. When the decomposition is *additive*, overall inequality can be written as the *sum* of within group and between group inequality:

$$\begin{array}{ccccccc} I & = & A & + & B \\ \text{overall inequality} & & \text{within group inequality} & & \text{between group inequality} \end{array}$$

When inequality is additively decomposed then one can say that the basis on which the individuals were subdivided (say, gender) contributed $[(B/I) \times 100]\%$ to overall inequality, the remaining inequality, $[(A/I) \times 100]\%$, being due to inequality *within* the groups. So, inequality decomposition provides a way of analysing the extent to which inter-personal inequality (in this case, in the HEA probabilities) is 'explained' by a set of factors (in this case, gender, sector, poverty, social group).²⁵ However, in order to decompose inequality *additively*, inequality has to be measured in a

²⁵ If, indeed, inequality can be 'additively decomposed' then, as Cowell and Jenkins (1995) have shown, the proportionate contribution of the between-group component (**B**) to overall inequality is the income inequality

very specific way. Only inequality indices which belong to the family of *Generalised Entropy Indices* are additively decomposable (Shorrocks, 1980) and one of these indices is Theil's (1967) Mean Logarithmic Deviation (MLD) Index (defined in equation (5) of the Appendix) which is used in the analysis.

The value of the MLD computed over the predicted probabilities of HEA of the 36,692 persons in the estimation sample was 0.203. When the sample was divided by the six social groups - ST, SC, HOBC, MOBC, MUC, and HUC- so that the number of groups was six, the within group (term **A** above) and the between group (term **B** above) contributions to the overall inequality of 0.203 were, respectively, 0.137 and 0.066. In other words, the division of the sample by social groups explained 33% of overall inequality in the probabilities of HEA.

When the sample was divided by the six social groups *and* by gender, so that the number of groups was 12, the within group and the between group contributions to the overall inequality of 0.203 were, respectively, 0.133 and 0.07. In other words, social groups *in conjunction with gender* explained 34.5% of overall inequality in the probabilities of HEA. The implication is that the addition of gender contributed 1.5 percentage points to the explanatory power of the inequality decomposition.

When the sample was divided by the six social groups, by gender, *and* by sector, so that the number of groups was 24, the within group and the between group contributions to the overall inequality of 0.203 were, respectively, 0.099 and 0.104. In other words, social groups *in conjunction with gender and sector* explained 51.2% of overall inequality in the probabilities of HEA. The implication is that the addition of sector, over and above social group and gender, contributed 16.7 percentage points to the explanatory power of the inequality decomposition.

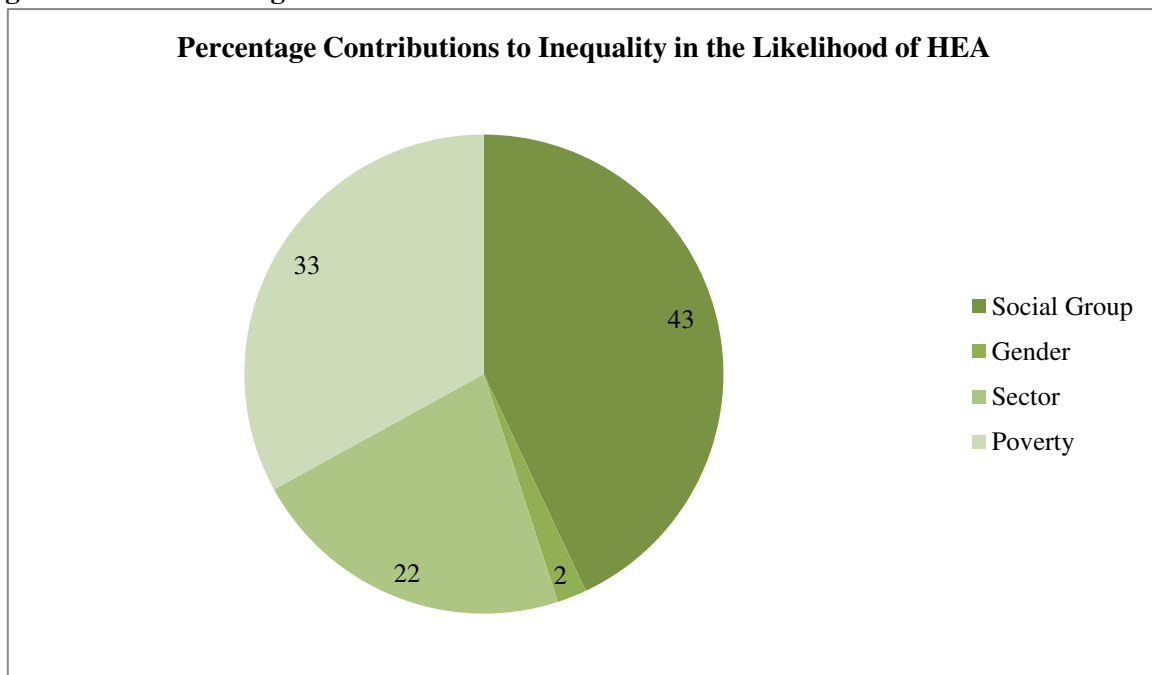
Lastly, when the sample was divided by the six social groups, by gender, by sector, *and* by poverty status (poor/non-poor), so that the number of groups was 48, the within group and the between group contributions to the overall inequality of 0.203 were, respectively, 0.048 and 0.155. In other words, social groups *in conjunction with gender, sector and poverty status* explained 76.4% of overall inequality in the probabilities of HEA. The implication is that the addition of poverty status -

literature's analogue of the R^2 statistic used in regression analysis: the size of this contribution is a measure of the amount of inequality that can be 'explained' by the factor (or factors) used to subdivide the sample

over and above social group, gender, and sector -contributed 25.2 percentage points to the explanatory power of the inequality decomposition.

So, in summary, over three-fourths of inequality in the distribution of the probabilities of 18-22 year olds being in HEA can be explained by just four factors: social group; gender; sector; and poverty status. Of this *explained* part, as Figure 5 shows: 43% (33/76.4) could be explained by social group; 2% (1.5/76.4) could be explained by gender; 22% (16.7/76.4) could be explained by sector; and 33% (25.2/76.4) could be explained by being poor.

Figure 5: Percentage Contributions to Inequality in the Distribution of Probabilities of Persons Aged 18-22 Years Being in HEA



Inequality in Expenditure on Higher Education

The preceding analysis focused on access to HEA. One can, however, also focus on the quality of education by examining inequality in the distribution between persons in HEA in their total expenditure on higher education (referred to earlier in Figure 2) and decomposing this inequality along the lines analysed above. The value of the MLD computed over the total expenditure on higher education (hereafter, HE expenditure) of the 36,692 persons in the estimation sample was 0.693.

When the sample was divided by the six social groups - ST, SC, HOBC, MOBC, MUC, and HUC— so that the number of groups was six, the within group (term **A** above) and the between group (term **B** above) contributions to the overall inequality of 0.693 were, respectively, 0.668 and 0.025. In other

words, the division of the sample by social groups explained 3.6% of overall inequality in HE expenditure.

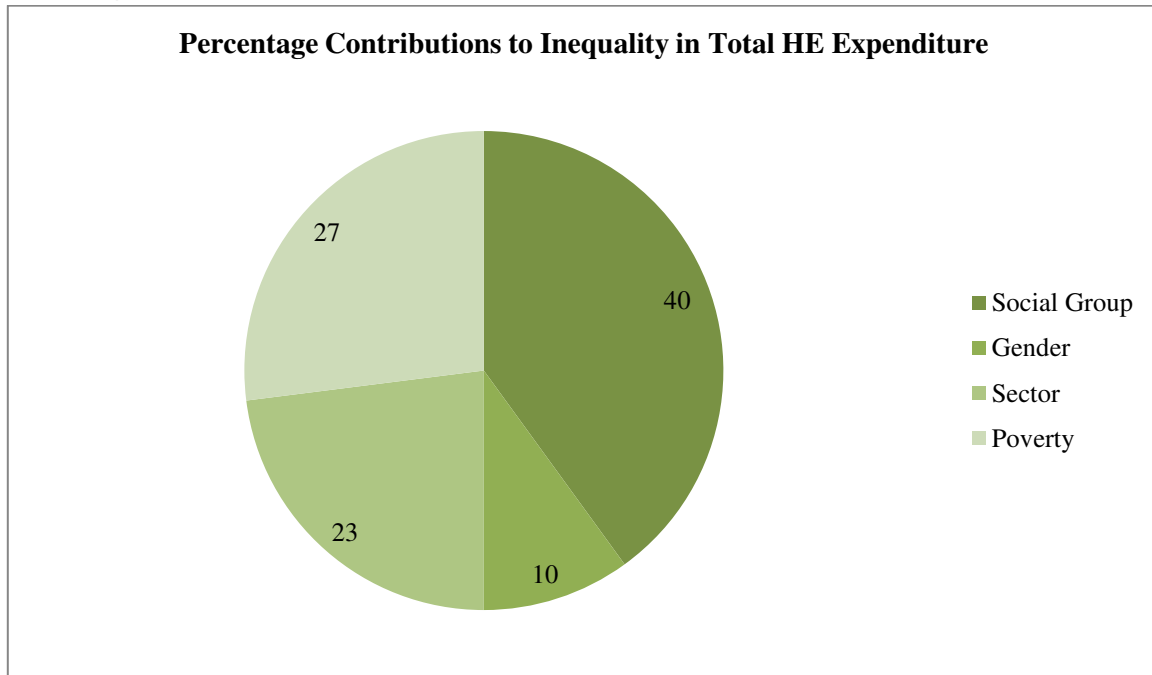
When the sample was divided by the six social groups *and* by gender, so that the number of groups was 12, the within group and the between group contributions to the overall inequality of 0.693 were, respectively, 0.662 and 0.031. In other words, social groups *in conjunction with gender* explained 4.5% of overall inequality in HE expenditure. The implication is that the addition of gender contributed 0.9 percentage points to the explanatory power of the inequality decomposition.

When the sample was divided by the six social groups, by gender, *and* by sector, so that the number of groups was 24, the within group and the between group contributions to the overall inequality of 0.693 were, respectively, 0.646 and 0.046. In other words, social groups *in conjunction with gender and sector* explained 6.6% of overall inequality in HE expenditure. The implication is that the addition of sector, over and above social group and gender, contributed 2.1 percentage points to the explanatory power of the inequality decomposition.

Lastly, when the sample was divided by the six social groups, by gender, by sector, *and* by poverty status (poor/non-poor), so that the number of groups was 48, the within group and the between group contributions to the overall inequality of 0.693 were, respectively, 0.630 and 0.063. In other words, social groups *in conjunction with gender, sector and poverty status* explained 9.1% of overall inequality in HE expenditure. The implication is that the addition of poverty status - over and above social group, gender, and sector - contributed 2.5 percentage points to the explanatory power of the inequality decomposition.

So, in summary, just under one-tenth of inequality in the distribution of HE expenditure between 18-22 year olds in HEA can be explained by four factors: social group; gender; sector; and poverty status. Of this *explained* part, as Figure 6 shows: 40% (3.6/9.1) could be explained by social group; 10% (0.9/9.1) could be explained by gender; 23% (2.1/9.1) could be explained by sector; and 27% (2.5/9.1) could be explained by being poor.

Figure 6: Percentage Contributions to Inequality in the Distribution of Total HE Expenditure of Persons Aged 18-22 Years who are in HEA



7. Measuring ‘Fair Access’

The most usual concept of ‘unfair access’ by a group to a particular ‘facility’ is that there is disproportionality between its representation in the population and in the facility. However, when there are many groups, the relevant question is how to merge these group disproportionalities into a *single* measure of *access inequality*. Ideally such a measure should satisfy the "Pigou-Dalton condition" which, applied to the present study, requires that an increase in numbers of deprived persons, *at the expense of an equal reduction in the number of non-deprived persons*, would reduce access inequality.²⁶

Suppose that the (estimation) sample of 36,692 persons, aged 18-22 years, is divided into M mutually exclusive and collectively exhaustive groups with N_m ($m=1\dots M$) persons in each group such that N_m and H_m are the numbers of pupils from *each* group in, respectively the ‘reference’ population

and in higher education. Then $N = \sum_{m=1}^M N_m$ and $H = \sum_{m=1}^M H_m$ are, respectively, the total numbers of

persons in the reference population and in higher education.

²⁶ In the language of inequality analysis this transfer *from* an "access-rich" group *to* an "access-poor" group constitutes a progressive transfer and, by virtue of this, is inequality reducing.

One way of measuring inequality in a variable is by *the natural logarithm of the ratio of the arithmetic mean of the variable to its geometric mean*.²⁷ As Bourguignon (1979) demonstrates, such a measure satisfies *inter alia* the Pigou-Dalton condition (discussed above). This idea translates very naturally, from its usual application to income inequality, to measuring the degree of inequality associated with educational outcomes in which people in different population groups meet with different degrees of success in securing a ‘desirable outcome’.

In this study, persons from the M groups meet with different degrees of success in terms of accessing higher education. The variable of interest is the *access rate* to higher education of persons from group m - defined as the proportion of persons from that group who are in HEA - and it is inequality in the distribution of this rate between the M groups that is sought to be measured. This inequality is referred to, hereafter, as “access inequality”.

If one takes the six social groups used in this study - ST, SC, NMOBC, MOBC, MUC, and NMUC - then Table 9 shows firstly, group shares among those, aged 18-22 years, who were in HEA and, secondly, among those who were in the 18-22 year age bracket, irrespective of whether they were in HEA.

Table 9: Shares of Social Groups in Higher Education Attendance

Social Group	NSS 71 st Round		NSS 64 th Round	
	Share in HEA of 18-22 year olds	Share in 18-22 year olds	Share in HEA of 18-22 year olds	Share in 18-22 year olds
Scheduled Tribe	5.5	10.0	3.0	8.0
Scheduled Caste	14.5	19.9	13.2	20.1
Non-Muslim OBC	37.2	34.5	33.7	35.7
Muslim OBC	4.9	8.4	3.6	6.2
Muslim Upper Caste	3.6	6.7	4.3	8.1
Non-Muslim Upper Caste	34.3	20.5	42.3	21.9
Total	100 [230,252]	100 [1,060,401]	100 [109,608]	100 [961,350]

The values of inequality (defined by J in equation (9) of the Appendix), calculated using the shares in Table 6, were 12.4 in the 64th round and 7.8 in the 71st round. So, in the seven years between the 64th round (July 2007-June 2008) and the 71st round (January-June 2014) access inequality *fell* by 37% from its 64th round value. This is, without doubt, an achievement for Indian higher education

²⁷ See Bourguignon (1979) and Theil (1967).

policy and it was obtained by increasing the share in HEA of all the groups at the expense of the Non-Muslim upper classes whose share in HEA was reduced from 41.3% in 2008 to 34.3% in 2014.

9. Conclusions

This paper began by examining access inequality to higher education in India in terms of estimating the likelihood that persons, aged 18-22 years, would be currently attending higher education (in HEA) courses, *conditional* on their social group, gender, poverty status, location, and marital status while allowing for *interaction* between these elements. Using the method of inequality decomposition the paper then computed the proportionate contribution of these factors to inter-personal inequality in the probabilities of 18-22 year olds in India being in HEA with the largest contributors being social group and poverty, followed by location, with a very small contribution by gender. The computation of access inequality – that is, aggregating group proportions in HEA and in the 18-22 year old population to arrive at a scalar measure of inequality – revealed a dramatic fall in access inequality between the 64th and the 71st NSS rounds. This fact, combined with a sharp rise in the Gross Tertiary Enrolment Rate between the two rounds, would suggest that in the past seven years there has been considerable achievement in higher education in India.

The problem that remains is one of ensuring quality in higher education. ‘Quality’ is an amorphous concept and, in this paper, it was proxied, firstly, by total expenditure on education and, secondly, by the proportion of students studying courses with English as the medium of instruction. As regards expenditure, it was shown that there was considerable difference between the social groups in the amounts that their members, who were in HEA, spent on their education: the median expenditure by those from the non-Muslim upper classes who were in HEA was ₹10,000-₹15,000 higher than that incurred by persons from the other groups.

The proportion of those in HEA who were studying in English rose from 51% of persons in HEA in the 64th round to 67% of persons in HEA in the 71st round. This might suggest a sharp rise in education quality but for a niggling doubt about the ability of persons studying in English to cope with instruction delivered in English. The basis for this doubt is that while 67% of persons in HEA in 2014 were studying in English, only 44% of persons were studying in English at the Higher Secondary level. Thus it is reasonable to suppose that a significant proportion of persons in HEA were, for the

first time in their life, encountering instruction delivered in English a fact which would possibly (though not necessarily) adversely affect their learning outcomes.

These anxieties are fanned by reservation policies which admit as students, to institutions of higher education, persons from 'reserved' categories who would not have won admission in open competition. Vishnu (2015) reports that in the academic year 2014-15 the elite Indian Institute of Technologies admitted 2,029 students from the SC and 856 students from the ST of whom, only 432 and 80, respectively would have secured admission in open competition based on examination performance. Thus for every 'academically reservation-unassisted' SC and ST student admitted to an IIT, 3.7 'academically reservation-assisted' SC students and 8.5 'academically reservation-assisted' SC students also secured a place.²⁸ The latter type of students struggle not only because they got in through 'reservation' rather than through the merit of examination results but also because in terms of their background and training they are ill-equipped to cope with academically challenging courses delivered in an unfamiliar language. The fact that they struggle and often fail then has repercussions for the quality of education offered: the 72 students, referred to in section 5, who should have been expelled from IIT Rorkee were, in fact, readmitted and given a second chance.

All these considerations call for rethinking the policy of setting aside a certain proportion of seats in higher education for persons from 'reserved categories'. Is a reservation-unassisted/reservation-assisted ratio in the IITs of 1:3.7 for the SC, rising to 1:8.5 for the ST, damaging these institutions by admitting students who are academically unprepared and weak? Would these institutions be strengthened if they had a student base which was stronger, chosen more on academic merit and less on social background? These are questions which deserve serious consideration.

Conversely, if, in the cause of correcting historical wrongs, it was thought legitimate to admit significant numbers from the reserved categories, the majority of whom would not have been admitted on academic merit, should educational institutions do more to redress these weaknesses? Many of the IITs have remedial classes and mentoring arrangements; but, what of other institutions? A serious

²⁸ The terms 'reservation-unassisted' and 'reservation-assisted' are used here *only* in the context of examination performance.

flaw in India's approach to redressing 'historical wrongs' that it relies solely on outcomes (reservation based quotas) but leaves people to sink or swim after their entry into an institution has been secured. In the process some, with great effort - as detailed in Singh (2013) - learn to swim but many drown. An effective and intelligent educational policy should be based, instead, on a more caring introduction to the deep waters of academic study.

References

- ASER (2014), Annual Status of Education Report (Rural) 2014), New Delhi: ASER Centre.
- Basant, R. and Sen, G. (2014), "Access to Higher Education in India: an exploration of its antecedents", *Economic and Political Weekly*, vol. XLIX, pp. 38-45.
- Basu, K. (2001), "On the Goals of Development", in G.M. Meier and J.E. Stiglitz (eds), *Frontiers of Development Economics: the Future in Perspective*, New York: Oxford University Press (pp. 61-86).
- Basu, K. (2006), "Globalisation, Poverty, and Inequality: What is the Relationship? What can be Done?", *World Development*, **35**: 1361-1373.
- Bhalotra, S. and Zamora, B. (2010), "Social Divisions in Education in India", in R. Basant and A. Shariff, *Handbook of Muslims in India: Empirical and Policy Perspectives*, New Delhi: Oxford University Press, pp. 165-198.
- Borooah, V.K. and Iyer, S. (2005), "Vidya, Veda, and Varna: The Influence of Religion and Caste on Education in Rural India", *Journal of Development Studies*, vol. 41, pp. 1369-1404.
- Borooah, V.K. (2012), "Social Identity and Educational Attainment: The Role of Caste and Religion in Explaining Differences between Children in India", *Journal of Development Studies*, vol. 48, number 7, pp. 887-903.
- Bourguignon, F. (1979), "Decomposable Income Inequality Measures", *Econometrica*, vol. 47, pp. 901-20.
- Chamarbagwala, R. (2006), "Economic Liberalisation and Wage Inequality in India", *World Development*, vol. 34, pp. 1997-2015.
- Cowell, F.A. and Jenkins, S.P. (1995), 'How Much Inequality Can We Explain? A Methodology and an Application to the United States', *Economic Journal*, vol. 105, pp. 421-30.
- Deshpande, S. (2013), "Introduction: Beyond Inclusion", in S. Deshpande and U. Zacharias (edited), *Beyond Inclusion: the Practice of Equal Access in Indian Higher Education*, New Delhi Routledge, pp. 1-12.
- Dreze, J. and Kingdon G.G. (2001), 'School Participation in Rural India', *Review of Development Economics*, vol. 5 (1), pp. 1-24.

Jeffery, R. and Jeffery, P. (1997), *Population, Gender and Politics*, Cambridge: Cambridge University Press.

Jeejeebhoy, S.J. and Sathar, Z.A. (2001), Women's Autonomy in India and Pakistan: The Role of Religion and Region, *Population and Development Review*, **27**, pp. 687-712.

Krishna, A. (2013), "Making it in India: Examining Social Mobility in Three Walks of Life", *Economic and Political Weekly*, vol. XLVIII, pp. 38-49.

Long, J.S. and Freese, J. (2014), *Regression Models for Categorical Dependent Variables using Stata*, Stata Press: College Station, Tx.

Mohanty, M. (2006), "Social Inequality, Labour Market Dynamics, and Reservation", *Economic and Political Weekly*, vol. XLI, p. 3777-89.

Sachar Committee Report (2006), *The Social and Economic Status of the Muslim Community in India*, New Delhi: Government of India (Cabinet Secretariat).

Shorrocks, A.F. (1980), 'A Class of Additively Decomposable Measures', *Econometrica*, vol. 50, pp. 613-25.

Singh, A.K. (2013), "Defying the Odds: the Triumphs and Tragedies of Dalit and Adivasi Students in Higher Education in India", in S. Deshpande and U. Zacharias (edited), *Beyond Inclusion: the Practice of Equal Access in Indian Higher Education*, New Delhi Routledge, pp. 174-204.

Sundaram, K. (2006), "On Backwardness and Fair Access to Higher Education: results from the NSS 55th Round Surveys, 1999-2000", *Economic and Political Weekly*, vol. XLIII, p.5173-5182.

Theil, H. (1967), *Economics and Information Theory*, Amsterdam: North-Holland.

UNESCO (2015), *Education for All 200-2015: Achievements and Challenges*, UNESCO Publishing: Paris.

Vishnu, U. (2015), "They get leg-up at JEE but hard landing on campus", *Indian Express*, 6 August.

Appendix
Interpretation of the Interaction Terms in Equation (3)

By virtue of the interactions characteristic, the estimated model can be written as:

$$\begin{aligned}
 F[\Pr(Y_i = 1)] = & \beta_0 + \beta_1 \times FEM_i + \beta_2 \times POOR_i + \beta_3 \times SEC_i \\
 & + \beta_4 \times FEM_i \times POOR_i + \beta_5 \times FEM_i \times SEC_i + \beta_6 \times POOR_i \times SEC_i + \beta_7 \times FEM_i \times POOR_i \times SEC_i \\
 & + \beta_8 \times X_i + \beta_9 \times X_i \times FEM_i + \beta_{10} \times X_i \times POOR_i + \beta_{11} \times X_i \times SEC_i \\
 & + \beta_{12} \times X_i \times FEM_i \times POOR_i + \beta_{13} \times X_i \times FEM_i \times SEC_i + \beta_{14} \times X_i \times POOR_i \times SEC_i \\
 & + \beta_{15} \times X_i \times FEM_i \times POOR_i \times SEC_i + \beta_{16} \times X_i \times FEM_i \times MAR_i + \beta_{17} \times STATE_i + \varepsilon_i
 \end{aligned} \tag{3}$$

Where:

- a) $Pr(Y_i=1)$ is the probability of the person in HEA and $F(.)$ is a non-linear function of this probability to reflect the fact that a non-linear (logit) model is being estimated.
- b) $FEM_i=1$, if person i is female, $FEM_i=0$, if person i is male.
- c) $POOR_i=1$, if the person i 's household's MPCE is in the lowest quintile, $POOR_i=0$, if it is not.
- d) $MAR_i=1$, if the person is married; $=0$ if never married.
- e) X_i is a social group control variable. For example, $X_i=1$ if person i is from the SC, $X_i=0$ if person i is not from the SC.
- f) $STATE_i$ is the state in which the person lived (35 states).

The interpretation of the coefficients of equation (3) is as follows:

1. The coefficient β_0 is an intercept term which operates *regardless* of the gender, or poverty status, or sector of the person. So, β_0 , on its own, relates to a person who is: male ($FEM_i=0$); not poor ($POOR_i=0$); lives in a rural area ($SEC_i=0$); and is not SC ($X_i=0$).
2. The coefficients β_1 , β_2 , and β_3 measure the *additional* effects of being, respectively, female, poor, and living in an urban area.
3. The coefficient β_4 measures the *additional* effect of being *both* female and poor; the coefficient β_5 measures the *additional* effect of being *both* female and living in a urban area; the coefficient β_6 measures the *additional* effect of being *both* female and living in a urban area; and the coefficient β_7 measures the *additional* effect of being female, poor, and living in a urban area.

4. The coefficient β_8 measures the effect of the control variable (for example, being SC) regardless of the gender, poverty status or sector of the person. So, β_0 , on its own, relates to a person who is: male ($FEM_i=0$); not poor ($POOR_i=0$); lives in a rural area ($SEC_i=0$); and is SC ($X_i=1$).
5. The coefficients β_9 measures the *additional* effect of being SC when the person is female
6. The coefficients β_{10} measures the *additional* effect of being SC when the person is poor.
7. The coefficients β_{11} measures the *additional* effect of being SC when the person is in the urban sector.
8. The coefficients β_{12} measures the *additional* effect of being SC when the person is female and poor.
9. The coefficients β_{13} measures the *additional* effect of being SC when the person is female and in the urban sector.
10. The coefficients β_{14} measures the *additional* effect of being SC when the person is poor and in the urban sector.
11. The coefficients β_{15} measures the *additional* effect of being SC when the person is female, poor and in the urban sector.
12. The coefficient β_{16} controlled for the state of residence.

Table A1: Interpretation of the coefficients of equation (3)

Attributes	Coefficient
$X_i=0$ (Non-SC)	
Male, non-poor, rural	β_0
<i>Additional Effect from being:</i>	
Female	β_1
Poor	β_2
Urban	β_3
<i>Additional Effect from being:</i>	
Female & poor	β_4
Female & urban	β_5
Poor & urban	β_6
Female, poor, & urban	β_7
$X_i=1$ (SC)	
Male, non-poor, rural	β_8
<i>Additional Effect from being:</i>	
Female	β_9
Poor	β_{10}
Urban	β_{11}
<i>Additional Effect from being:</i>	
Female & poor	β_{12}
Female & urban	β_{13}

Poor & urban	β_{14}
Female, poor, & urban	β_{15}

Inequality Decomposition

Suppose that the sample of $N=36,692$ persons, aged 18-22 years, is divided into M mutually exclusive and collectively exhaustive groups with N_m ($m=1 \dots M$) persons in each group. Let $\mathbf{p} = \{p_i\}$ and $\mathbf{p}_m = \{p_i\}$ represent the vector of (estimated) probabilities of HEA of, respectively, all the persons in sample ($i=1 \dots N$) and the persons in group m . Then an inequality index $I(\mathbf{p}; N)$ defined over this vector is said to be additively decomposable if:

$$I(\mathbf{p}; N) = \sum_{m=1}^M I(\mathbf{p}_m; N_m) w_m + \mathbf{B} = \mathbf{A} + \mathbf{B} \quad (4)$$

where: $I(\mathbf{p}; N)$ represents the *overall* level of inequality; $I(\mathbf{p}_m; N_m)$ represents the level of inequality within group m ; \mathbf{A} – expressed as the weighted sum of the inequality in each group, w_m being the weights – and \mathbf{B} represent, respectively, the *within-group* and the *between-group* contribution to overall inequality.

Only inequality indices which belong to the family of *Generalised Entropy Indices* are additively decomposable (Shorrocks, 1980). These indices are defined by a parameter θ and, when $\theta=0$, the weights are the population shares of the different groups (that is, $w_j = N_j / N$); since the weights sum to unity, the within-group contribution \mathbf{A} of equation (4) is a weighted average of the inequality levels within the groups. When $\theta=0$, the inequality index takes the form:

$$I(\mathbf{p}; N) = \left(\sum_{i=1}^N \log(p_i / \bar{p}) \right) / N \quad (5)$$

where: $\bar{p} = \sum_{i=1}^N p_i / N$ is the mean probability over the entire sample. The inequality index defined in equation (5) is known as the Theil's (1967) Mean Logarithmic Deviation (MLD) and, because of its attractive features in terms of the interpretation of the weights, it was the one used in this study to decompose inequality in the likelihood of HEA.

Unfair Access

The success rate of group m (denoted e_m) is $e_m = H_m / N_m$, $0 \leq e_m \leq 1$. Then the arithmetic and geometric means of e_m are, respectively:

$$\bar{e} = \sum_{m=1}^M e_m n_m \quad \text{and} \quad \hat{e} = \prod_{m=1}^M (e_m)^{n_m} \quad \text{where} \quad n_m = N_m / N, \quad \sum_{m=1}^M n_m = 1 \quad (6)$$

so that the measure of access inequality is:

$$J = \log(\bar{e} / \hat{e}) = \log(\bar{e}) - \sum_{m=1}^M n_m \log(e_m) \quad (7)$$

Now from the definition of e_m :

$$e_m = H_m / N_m = (H_m / N_m)(N / H)(H / N) = (H_m / H)(N / N_m)(H / N) = (h_m / n_m)\bar{e} \quad (8)$$

where: $h_m = H_m / H$ and $n_m = N_m / N$ are, respectively, group m 's share of higher education attendees and of the population. Employing equation (8) in equation (7) yields:

$$J = \log(\bar{e} / \hat{e}) = \log(\bar{e}) - \sum_{m=1}^M n_m \log(e_m) = \log(\bar{e}) - \sum_{m=1}^M n_m \log\left[\frac{h_m}{n_m}\bar{e}\right] = \sum_{m=1}^M n_m \log\left[\frac{n_m}{h_m}\right] \quad (9)$$

From equation (9), inequality is minimised when $J=0$. This occurs when $n_m = h_m$, that is when each group's share in the 'population' (n_m) is equal to its share in higher education attendees (h_m). Otherwise, $J > 0$.

The inequality measure, J , of equation 6, has along the lines suggested by Bourguignon (1979), an appealing interpretation. If social welfare is the sum of identical and concave group utility functions whose arguments are e_m then social welfare is maximised when e_m - the success rate of a group - is the same for every group. If the utility functions are of the logarithmic form (that is, $U(e_m) = \log(e_m)$), then J represents the distance between *maximum* level of social welfare ($\log(\bar{e})$)

and the *actual* level of social welfare ($\sum_{m=1}^M n_m \log(e_m)$): *social welfare is maximised when access inequality is minimised!*