



Swansea University  
Prifysgol Abertawe



## Cronfa - Swansea University Open Access Repository

---

This is an author produced version of a paper published in:  
*World Development*

Cronfa URL for this paper:  
<http://cronfa.swan.ac.uk/Record/cronfa37945>

---

### Paper:

Gupta, P., Mallick, S. & Mishra, T. (2018). Does social identity matter in individual alienation? Household-level evidence in post-reform India. *World Development*, 104, 154-172.  
<http://dx.doi.org/10.1016/j.worlddev.2017.11.007>

Released under the terms of a Creative Commons Attribution Non-Commercial No Derivatives License (CC-BY-NC-ND).

---

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder.

Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>

# Does social identity matter in individual alienation? Household-level evidence in post-reform India

## Abstract

Does consumption distance as a measure of individual alienation reveal the effect of social identity? Using the central idea of Akerlof's social distance theory, individual distance is calculated from their own group mean (median) consumption and then we examine whether individuals from different social groups - caste and religion - are alienated across the distance distribution. Using India's household-level microdata on consumption expenditure covering three major survey rounds since the inception of the reform period, we find a non-unique pattern in the distance distribution where the marginalised and minority group households tend to be alienated across the distribution. However, among them, the households with higher educational attainment become more integrated as reflected in the interaction effect of education. These results are robust even after controlling for the endogeneity of education. Given this significant group difference in consumption, we undertake a group-level comparison by creating a counterfactual group through exchanging the characteristics of the privileged group to the marginalised group (or Hindus to non-Hindus), and find that the privileged group still consumes more than the counterfactual marginalised group, explaining around 77% of the estimated average consumption gap at the median quantile in 2011-12 (or 59% for Hindus versus Non-Hindus). This suggests other inherent identity-specific social factors as possible contributors to within-group alienation (relative to a better-off category) that can only be minimised through promoting education for marginalised (or minority religion) groups.

---

**Key Words:** Consumption distance; Social identity; Social distance;  
Counter-factual distribution; India.

**JEL Classifications:** O1; R2; I3; D1

“Our inequality materializes our upper class, vulgarizes our middle class, brutalizes our lower class”.

- *Matthew Arnold (English Essayist (1822-1838))*

“This civilization is the impact of the world’s consumption behavior”. Toba Beta, *Betelgeuse Incident: Insiden Bait Al-Jauza*

- *Carol Graham, The Economist, Feb 1st, 2010*

## 1 Introduction

Whether income growth alone is sufficient to improve economic well-being in the long run remains a disputed issue<sup>1</sup>, especially in developing countries like India where the economy has experienced a high growth rate of 8-9% in the last decade compared to 5% in the 1990s (Dev, 2013). At the same time, the official statistics and academic research have shown that the increasing trend in income growth has contributed to higher economic inequality both at the aggregate level and among different social identities (Motiram and Sharma, 2011; the Government of India, 2013). Research on this view using data from developing countries also shows inequality limiting social mobility (Andrews & Leigh, 2009; Corak, 2013). This fact, in turn, suggests that if there is lack of social mobility among certain groups, this may then increase social distance<sup>2</sup> between groups. It also signals eroding social cohesion that could force groups to continuously engage in conflicts. Dutta et al. (2014), for instance, established that ‘societies with low levels of inequality show no conflict; groups engage in conflict only when inequality exceeds a certain threshold level’. Arguing in a similar vein, albeit different from the inequality literature, evidence suggests that the lack of interaction between two groups could lead them to higher social alienation (Esteban and Ray, 1999; Esteban et al., 2012), while Jayadev and Reddy (2011) illustrate different types of group-based inequality to capture the degree of segregation, clustering, and polarization between groups that may contribute to social unrest and violence.

Under this backdrop of alienation of certain groups in a society, in this paper we use the theoretical idea of ‘social distance’ as conceptualised in Akerlof (1997) and construct consumption distance (as an indirect measure of social alienation) from a societal benchmark and demonstrate to what extent higher/lower individual consumption distance is associated with their social identities - castes<sup>3</sup> and religions<sup>4</sup>. The term ‘social distance’ refers to how an agent’s

---

<sup>1</sup>Banerjee and Duflo (2007), Basu and Mallick (2008), Kumar et al. (2009); and Mallick (2014)

<sup>2</sup>The concept of social distance was developed first by sociologist Bogardus (1933) which measures the degree of closeness or acceptance that one feels toward other groups. In the recent years, the notion of social distance in economics is defined as the degree of reciprocity that exists within a social interaction.

<sup>3</sup>In India, caste is classified into 4 sub-groups, namely - Scheduled Castes (SC), Scheduled Tribes (ST), Other Backward Class (OBC) - categorized together as marginalised group - and High Caste (HC) or Privileged group. The constitutional term for SC is Dalits; the term refers to their special status under the constitution as a group entitled to special protection and to stop positive discrimination through reservations. This group accounts for around 16% of total population. The constitutional term ST is Adivasis; this term refers to their special status under the constitution as a group entitled to special protection and any positive discrimination through reservations. This group accounts for around 13% of total population. The third group Other Backward Class (OBC) has not been subject to the same discrimination of untouchables in India, but they are considered socially and economically deprived like SC and ST after the Mandal Commission report, and the NSS has started collecting data for this group since 1999-2000 survey. This group accounts for around 39% of total population. The privileged group (HC) has complete privilege over other three groups and this group accounts for around 32% of total population.

<sup>4</sup>In the case of religion, we strictly categorize it into two groups - Hindus and Muslims. According to Population Census, Hindus account for around 88% and Muslims account for around 12% of total population of India.

sense of belonging and aloofness from a desirable social outcome can determine the welfare of an individual. It is not unusual to observe that consumption pattern of some social groups does not dynamically co-move with the general rise in income (or consumption) at the aggregate level. Although social identity is a multidimensional concept, a person's identity is determined at birth in terms of a religion and a particular sub-group. The household surveys conducted by the Government of India identify a single identity to each household which makes our identification strategy credible. In this regard, we raise two core questions. First, whether individuals from some social groups are alienated more relative to other groups, or whether they are alienated less when they get educated? Second, what would have happened to consumption differences between groups if one exchanges their household characteristics? More specifically, this research is focused on consumption behaviour of two social identities - castes and religions, but the results can be generalised to similar socio-economic settings in other countries.

The paper uses quantile regression to answer the first question in assessing the consumption distance across the distribution, and counterfactual decomposition method to answer the second question. We use household-level data on consumption expenditure from National Sample Survey (NSS) for three rounds: 1993-94, 2004-05 and 2011-12 and performs the following tasks. First, we construct a measure of consumption distance at household level using individual consumption expenditure, and investigate whether a unique pattern (higher or lower consumption distance) can be discerned across social identities, controlling for demographic and other household characteristics, including educational attainment levels, while dealing with the endogeneity of education. Second, we also undertake a group-level comparison after exchanging group characteristics in a counterfactual sense to study whether the conditional average consumption gap between two groups tends to get wider across the distribution or rises over time.

Our results reveal that household consumption pattern is indeed influenced by individual identities, while controlling for the effects of education, age, and household type. From the distribution of raw consumption data, we find that there are multiple clusters (or equilibria) indicating the existence of both poverty and growth traps. This is in line with the wider literature that within-country income inequality has been on the rise, despite the reduction in between-country inequality in the recent decades. Using household-level consumption, we have focused on alienation as a form of individual distance from a group mean for different social groups within a country in order to contribute to this issue of growing within-country inequality.

We find that the consumption expenditure, when transformed in terms of distance as a deviation from the group mean consumption, is non-unique across the distribution and among social groups. The differences in consumption expenditure between groups have not changed significantly over time for those households who are at the lower ladder of the consumption expenditure distribution. However, households with education are found to have higher per capita consumption relative to their group mean. In this context, we examine for the first time in the NSS dataset the problem of endogeneity, using the number of schools and school enrollment at the district level as exogenous instruments for educational attainment. Even with this endogeneity correction, both marginalised social groups and minority religious group remain alienated across the distribution, relative to their respective counterparts.

In particular, we find that (i) on average, the consumption distance within a social group is

getting wider across the distribution and rising over time, implying alienation of the marginalised and minority group households, even after 25 years of economic reforms; (ii) the consumption distance of both marginalised and non-Hindus groups, without and with controlling for endogeneity, however remains similar at the bottom end of the distance distribution, while it gets wider at the top end, (iii) After exchanging the group characteristics among both social identities - castes and religions, the estimated consumption gap across the distribution is observed to rise over time in both unmatched and matched samples. In light of our findings, we conclude that the evidence of rising consumption distance across the entire distribution, over time, and across social groups could signal heightened social alienation and thus possible group conflicts. Despite resounding success in India's growth in the recent decades, the rising 'consumption distance' across social divides appears to reflect continued within-country economic inequality - the solution of which will require more inclusive policy decisions including promoting education.

The rest of the paper is structured as follows. Section 2 presents an overview of the literature. Section 3 provides data description, outlines strategies for the construction of consumption distance, and presents preliminary results. In Section 4 we present our main results (without and with accounting for endogeneity concerns). Section 4.3 discusses the counterfactual decomposition results and the robustness exercise. Finally, section 5 provides a critical summary of the main results and their possible policy relevance.

## 2 A brief review of the Literature

The literature has emphasised to what extent social affinities (caste and religious affiliation) govern different channels of economic development. In this section, we look at some closely related studies which have observed how caste or religious identity matters in economic development. Caste affiliation and religious barrier are not new in India and deeply rooted since independence in the Indian society. Despite various efforts by the government since independence, it has been observed that marginalized groups had low occupational mobility and were engaged mostly as agricultural labourers or in self-employed agricultural work (Gang, Sen and Yun, 2012), which implies significant social inequality between high and low castes. Using caste affiliation as a key factor, many studies as cited earlier have tried to explain its impact on development including market insurance, bank credit, poverty, and income inequality. However, what is less known in the literature is that to what extent an individual's social identity might have had a differential effect on their consumption pattern relative to the average level of consumption across the distribution, following two decades of policy reforms in India?

One possible reason for the social hierarchy is to look at the historical legacy between high caste and low castes, where high caste individuals do not allow people from the lower castes to share the same resources. As a consequence of this, lower castes always had low mobility in the labour market and thus reported higher poverty level and income inequality (Desai and Dubey, 2012). The second possible reason is through the development channel where low-caste individuals in India are observed to have benefited less from recent growth process compared to their counterpart in the high caste. Both explanations have been discussed intensively in the literature with respect to social inequality, the wage gap, and caste-based discrimination. For instance, Kijima (2006) uses NSS data for the periods 1983-1999 and has reported that the

disparities of living standards among SC/ST (marginalized groups) compared to non-SC/ST (privileged group) still remain very high. He finds that due to social differences between non-SC/ST and SC/ST, SC/ST have less access to public goods and thus possess lower human and physical capital. Munshi and Rosenzweig (2006) observed that social identities are linked with low paid jobs among SC/ST. This suggests that the disparities of living standards between SC/ST and non-SC/ST households are largely attributable to different returns to schooling.

In order to explain the existence of the large disparity between high and low castes, some recent studies have linked this to the concept of conspicuous consumption behaviour, where economically disadvantaged groups try to lower the consumption distance by spending more on leisure/luxury items to gain higher social status. For instance, Khamis et al. (2012) use India Human Development Survey (IHDS) data to investigate if spending on consumption items can be characterized by social identities (defined by caste and religious affiliations). The authors used those consumption items which people are quick to observe and which they associate with higher income so that such items confer status on those who consume them (conspicuous consumption). They observe that marginalized groups spend nine percent more on visible consumption than high caste, while Muslims in Indian society spend eleven percent less, after controlling for differences in permanent income and demographic composition of households.

Understanding social exclusion based on group or club formation has suggested that India has become more socially disintegrated in the recent decades and is moving towards being a less cohesive society. For instance, Bandyopadhyay (2012) using data on income across Indian states suggests that India was more cohesive before reform in the sixties and started diverging after the seventies and the divergence is observed to be more acute in the nineties. She finds that after the nineties, India has been divided into two societal clubs. In a similar spirit, Mitra and Ray (2014), using the concept of group convergence and dataset on Hindu-Muslim violence between 1950 and 1995 for three consecutive rounds from National Sample Surveys (NSS): 1983, 1987-88 and 1993-94, find that Hindu prosperity reduces fatalities from conflicts between religious groups. For instance, they observe that a 1% increase in Hindu per-capita expenditure is predicted to decrease casualties by anywhere between 3-7%, while the same increase in Muslim per-capita expenditure increases casualties by 3-5%.

Despite the recent high economic growth, India's consumption inequality across caste and religions has been creating a buzz in the academic and policy circles. The rise in within-country inequality raises concerns about a decline in opportunities for the deprived sections of the population in India. This research thus fills a gap in this recent literature on social inequality, with implications for lowering social alienation via human capital formation.

### **3 Data and Preliminary observations**

#### **3.1 Data**

In this section, we describe the data and present some preliminary observations leading to the main empirical results in the following sections. We use data on monthly per-capita consumption expenditure (in real terms) at household level. In India, the NSS collects this data across different geographical regions, demographics, and among social groups. These data are collected every

five years and have been used as an important source to estimate poverty/inequality over time in many studies. For our investigation, we use data from three quinquennial rounds - 50th (period 1993-94), 61st (period 2004-05) and 68th (period 2011-12). These surveys cover the whole country, except some remote villages. In each survey period, the same sample villages/blocks have been allotted for the survey with a view to ensuring uniform spread. A stratified multi-stage sample design was adopted in each period. In the rural sector, each district is generally counted as a separate stratum (populous districts are split into two or more strata), whereas, in the urban sector, strata are formed within the NSS region based on population size of cities. From the 61st round, the stratification method has slightly changed. For each district, the rural and urban strata are further divided into two sub-strata.

The questionnaire records demographic details such as age, sex, marital status, educational attainment level, religion, social groups and information about occupation and industry. It also collects monthly total household expenditure on food and non-food items over 30 days and 365 days recall periods. In our analysis, we have considered age, years of schooling, castes, religions and household type.

Moreover, in this study we distinguish households' monthly per capita consumption expenditure based on 30-days recall period as an indicator of welfare measure in various sub-groups: (i) by rural-urban, (ii) by different educational attainment level, (iii) by social classes, and (iv) by different age cohorts. In Table 1, we present the detailed summary statistics of the average distribution of monthly per capita consumption expenditure across different demographic characteristics for each selected survey period at the mean level and at two interquartile ranges. The individual consumption expenditure for all rounds is expressed in real terms using state-level poverty lines that differ for rural and urban areas. Following Lahiri et al. (2012), we express all consumption expenditures in real terms of 1993 rural and urban Maharashtra poverty lines. Such transformation helps cross-state comparisons in real terms with a specific benchmark. In order to understand consumption expenditure differential among different sub-groups and over time, we use head of the households age 15 and above in order to avoid questions on intra-household distribution, such as the use of consumption expenditure per adult equivalent or equivalence scales<sup>5</sup>. We begin our empirical investigation by undertaking a preliminary data analysis.

### 3.2 Preliminary observations

- *Observation 1: Inter-quartile differences in consumption-expenditure*

In Table 1, some general trends of consumption expenditure for different social categories are presented. We observe that the mean consumption at the national level and across different categories has increased approximately about one-third in 2011 compared to 1993.

However, consumption expenditure at inter-quartile ranges displays varied patterns with

---

<sup>5</sup>The rationale behind not using the equal division of consumption assumption within the same household in the current analysis is that it rejects the assumption that the consumption allocation does not vary with the source of income in the household. A large theoretical and empirical literature (see Browning et al 1994; Lundberg, Pollak and Wales, 1997; Manser and Brown, 1980; McElroy and Homey, 1981; Chiappori, 1988, 1992; Browning and Chiappori, 1998; Chiappori, Fortin and Lacroix, 2002; Donni, 2007) on this issue have also pointed out that if the source of income within household is not consistent, then use of any measure of equivalence scales will produce an inaccurate trend in consumption behaviour over time. Regardless of this, we have used the square root scaling to check the robustness of our baseline results and the estimated results are presented in Table 7 as Model 2.

households at 25th quartile consuming nearly half of the mean consumption expenditure, whereas at 75th quartile these differences are observed approximately 33 percent higher than mean consumption level and over time. To confirm the differences in consumption level at two inter-quartiles, a mean test is also performed for all groups.

Table 1: Consumption (*real terms*) distribution across different social identities at mean and two inter-quartile ranges

	1993			2004			2011		
	Mean	25th	75th	Mean	25th	75th	Mean	25th	75th
All India	451.46 (1.799)	230.97 (0.513)	499.14 (1.897)	498.73 (1.478)	249.55 (0.473)	569.70 (1.714)	627.93 (2.194)	313.71 (0.665)	712.50 (2.137)
<b>Caste</b>									
SC	326.43 (2.481)	197.57 (1.006)	369.48 (2.496)	439.01 (3.197)	231.71 (1.059)	491.45 (3.306)	502.01 (4.991)	275.29 (1.215)	554.30 (4.278)
ST	333.27 (3.528)	198.34 (0.998)	369.75 (2.846)	402.86 (3.619)	221.35 (1.477)	455.30 (3.234)	483.20 (3.793)	267.86 (1.386)	560.60 (5.001)
OBC	-	-	-	464.17 (2.228)	242.99 (0.737)	527.91 (2.436)	587.65 (3.094)	310.48 (1.003)	661.86 (3.273)
HC	500.65 (2.349)	250.60 (0.626)	560.08 (2.501)	598.70 (2.898)	285.69 (1.033)	705.54 (3.605)	792.66 (4.712)	381.43 (1.512)	919.36 (5.391)
<b>Religions</b>									
Hindu	452.80 (2.031)	230.54 (0.571)	502.55 (2.009)	502.30 (1.683)	249.28 (0.547)	576.23 (1.946)	631.66 (2.546)	313.11 (0.749)	716.04 (2.527)
Islam	394.86 (3.725)	225.27 (1.473)	440.67 (4.258)	450.12 (3.761)	240.95 (1.225)	509.70 (3.759)	558.27 (4.758)	306.20 (1.600)	636.08 (4.692)
<i>N</i>	<i>69152</i>	<i>69152</i>	<i>69152</i>	<i>109320</i>	<i>109320</i>	<i>109320</i>	<i>89623</i>	<i>89623</i>	<i>89623</i>

Note: Consumption expenditure is defined in real terms.

- *Observation 2: Density distribution of consumption expenditure*

We first plot the unconditional distributional pattern of consumption expenditure and look at the shift in consumption behaviour at the aggregate level and over time. This relates to our research question whether an increase in consumption expenditure is directly linked to the recent concern on increase (decrease) in social alienation, to the possibilities of articulated rebellion and revolt, and to the existence of social unrest in various parts of the society. The distributional pattern of consumption distance is therefore important to be analysed.

Traditionally, any shift in consumption (income) level to the right of the distribution is always considered as a positive sign of growth, although the distribution could be skewed at either end indicating clustering or polarization. We use consumption distance as an indirect measure of social alienation and examine the increase (or decrease) in social alienation during the last two decades of growth in India. We assume that individuals at both ends of the consumption distribution plotted in Figure 1 represent two different and isolated groups of the same society. In other words, we assume that the households which are extremely on the left of the consumption expenditure distribution are poor <sup>6</sup> and therefore

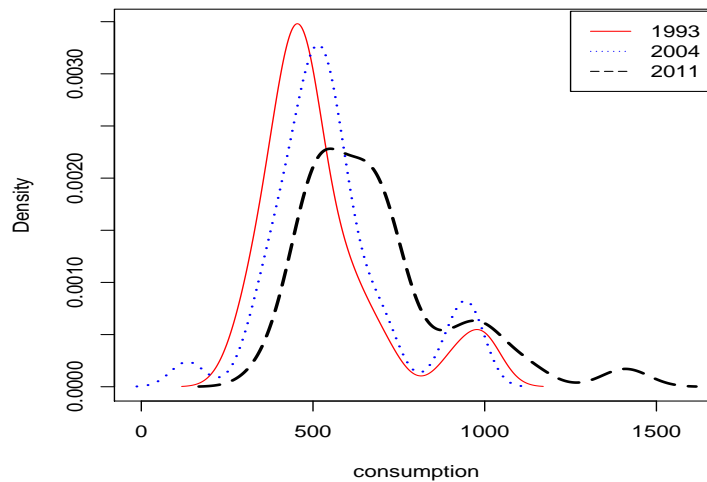
<sup>6</sup>We consider those individuals poor who consume less than the official poverty line



they have less access to economic resources. On the contrary, households on the extreme right end of the same distribution are considered as non-poor (as their consumption level is above the mean consumption) and thus they have full access to available resources. In other words, those households on the right of the distribution tend to be more integrated with the societal norm, but among them, some groups are still alienated (consume less) relative to their counterpart at different points of the distribution.

In Figure 1 the varied pattern of consumption distribution for the periods 1993, 2004, and 2011 confirms our contention and clearly shows clusters at both ends. A prolonged right tail implies that a very large number of households is consuming above the mean consumption expenditure. As expected, we also find the higher mean (see Table 1) for the year 2011 compared to both 1993 and 2004, indicating a real improvement in consumption expenditure reflecting a favourable impact of policy reforms in the last two decades. The extreme right mode in the same distribution can be thought of as a ‘*growth trap*’, where some individuals have full access to economic resources, enjoying all benefits from economic development.

Figure 1: Average distribution of consumption expenditure over time



In order to confirm the normality of the distribution, the Jarque-Bera normality test is used to examine whether the sample data have skewness and kurtosis matching a normal distribution. The skewness and kurtosis for three time points<sup>7</sup> indicate that all distributions are significantly different from a normal distribution.

- *Observation 3: Consumption gap among social identities*

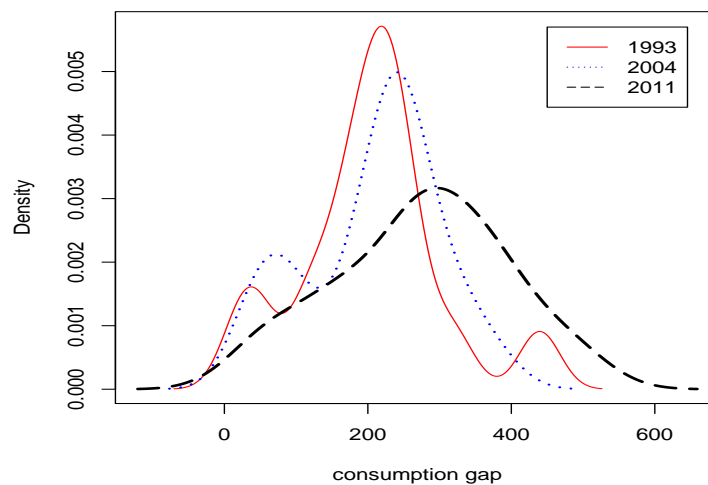
The observed clusters over time in observation 2 reflect that some individuals are consuming above the mean consumption level. A similar pattern is also observed for food vs non-food and cereal vs non-cereal distribution<sup>8</sup>. However, it is not examined yet how these

<sup>7</sup>Skewness (SK) and Kurtosis(KU) are as follows: 1.47 and 4.67 (SK and KU, 1993); 0.68 and 3.76 (SK and KU, 2004); and 1.38 and 4.94 (SK and KU, 2011).

<sup>8</sup>These results are available upon request from the authors.

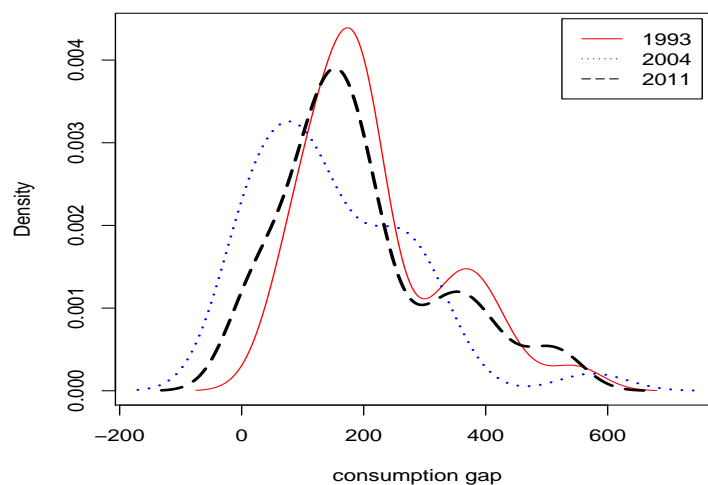
differences in consumption level are characterized between two social identities. To confirm this, we have further looked at the consumption expenditure gap between both social identities - caste and religion, and plotted them as in Figures 2 and 3. The consumption gap between two groups is calculated based on their average differences in consumption level over time and in each state of India. The visual inspection of these graphs suggests that both religious groups and castes are divided into many clusters across the distribution.

Figure 2: Average consumption gap between Hindus and non-Hindus



Source: Authors' calculation

Figure 3: Average consumption gap between Privileged and Marginalized groups



Source: Authors' calculation

The observed clusters at both ends of the distribution in observations 2 and 3 suggest two different but contemporary issues in development economics. First, each mode in

the distribution reflects the possibility of ‘*poverty trap*’ and ‘*growth trap*’ at both ends. Second, this figure relates to a contemporary issue in econometric theory which says that if the dependent variable of the regression (in our case, consumption expenditure and normalized consumption expenditure) is multi-modal, then the estimation of parameters assuming unimodality of the dependent variable would provide biased estimates. Moreover, the information content at other parts of the distribution is disregarded or averaged out and this would affect our inference on consumption behaviour across social identities. The finding of multi-modality is not unusual and has been consistently reported in extant empirical research on employment, income distribution, and consumption behaviour. Surprisingly, only a few studies have adopted alternative estimation methods, such as quantile regression, to address the issue of multiple modes in the distribution.

### 3.3 Empirical strategy for measuring consumption distance

In order to examine differential patterns in consumption distribution among different social identities, we use quantile regression method, and to describe consumption distance, we adopt a method of standardization in the following subsections.

#### 3.3.1 Quantile regression method

The unconditional pattern of consumption expenditure in the density plots and our preliminary investigation of consumption expenditure over time and across social identities have provided enough evidence as to why one should examine its effect across the distribution of consumption expenditure rather than concentrating at the mean. A key aspect of understanding the effect of socio-economic characteristics is to employ an estimation method that can be informative in addressing heterogeneity, whilst also robustly presenting parameter estimates at various points of the distribution of the dependent variable.

To understand these differential spending patterns among different social identities, we use Koenker and Bassett (1978) quantile regression approach. With this approach, one can estimate the change in the parameter distribution at specific quantiles so that the response variable produced by a one unit change in the predictor variable can be gauged at both higher and lower quantiles of the distribution<sup>9</sup>. This can be specified as:

$$Q_{\theta}(Y|X) = X\beta_{\theta} + \epsilon \tag{1}$$

where  $X$  is a set of standard control variables such as age, household size, education level, religion and caste.  $\theta$  represents quantiles to be estimated between  $(0, 1)$ .

One of the most interesting advantages provided by quantile regression is its semi-parametric nature that avoids assumptions about parametric distribution of regression errors and makes quantile regression especially useful when dealing with heteroskedastic errors.

---

<sup>9</sup>The superiority of this approach in relation to mean-based approach can be gauged by its statistical properties. If the response of the distribution to certain endogenous shocks (in our case, consumption distance) is true for a certain quantile, say at 10% and if it remains consistent throughout the entire range of quantiles, one can say that the estimated parameters represent a true relationship.

### 3.3.2 Standardisation of consumption expenditure

Understanding the differential patterns of consumption expenditure across the distribution is a key objective of this paper. It is not only important but also necessary to understand to what extent an individual's own consumption is away from the average of their own group. If this difference rises over time, then one can say that this particular society is facing the problem of social alienation. In particular, if there is no consumption differential within a group, then each individual should have the same level of consumption and hence there will be no sign of consumption distance or alienation for a given identity. The households on the right of the distribution, consuming more than the mean level of consumption, could be more integrated with the societal norm, but they could only be alienated or experience relative inequality if we compare them to their counterparts in a privileged group.

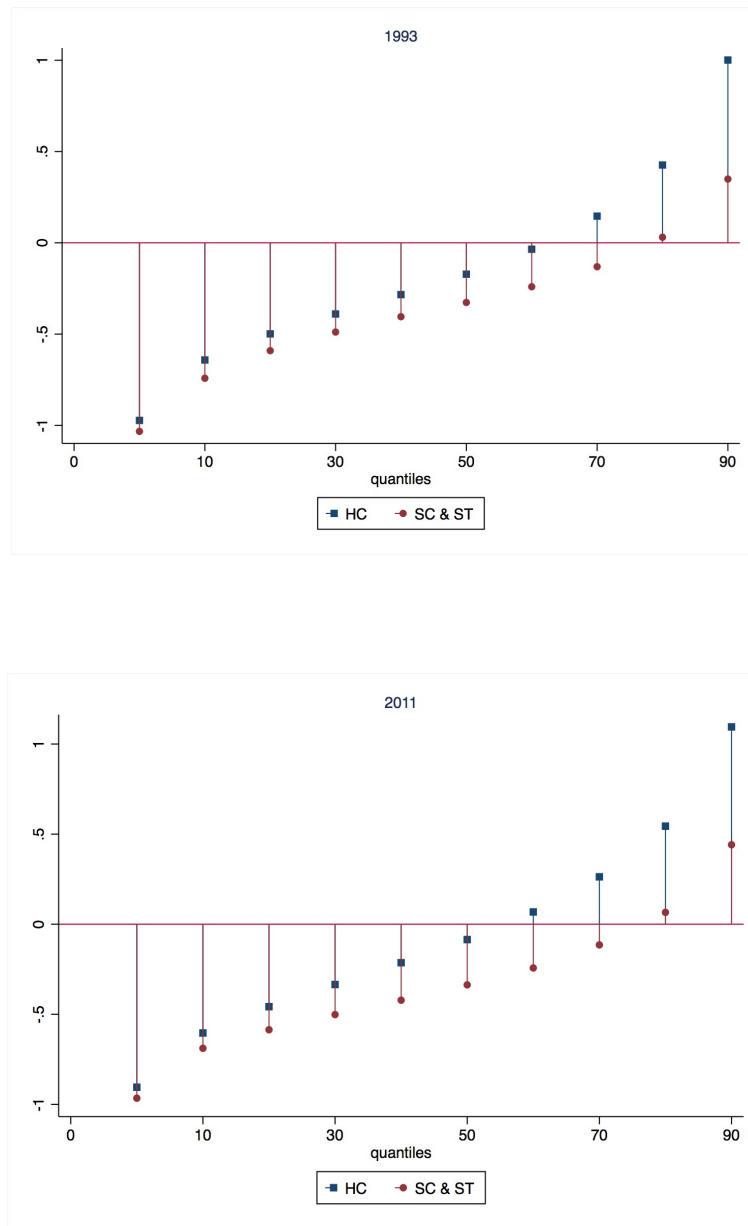
To conceptualize this idea, given the information on consumption (as a proxy for income), we measure each individual household's consumption level from a reference point, say mean of the consumption expenditure distribution of their own group, and divide by its standard deviation (see the detailed steps in Appendix A). One can change this reference point from mean to another reference point such as median or maximum/saturation level (as in Kumar et al., 2009) of consumption expenditure to calculate this distance. This, however, will not change the interpretation of the estimates. This can be represented in the following way as:

$$dis_{it}^c = \left( \frac{x_i^c - \bar{x}}{\sigma} \right) \quad (2)$$

where  $dis_{it}^c$  is distance from mean consumption level of a household's own group as reference in a given period  $t$ .  $x_i^c$  is individual consumption expenditure,  $\bar{x}$  is the mean consumption expenditure, and  $\sigma$  is standard deviation of consumption in the respective group. Measuring deviation of individual consumption in this manner has two attractive properties. First, it is invariant to changes in units. Second, one can directly observe the effect of one  $\sigma$  change in independent variable on the dependent variable. The steps involved in normalization is discussed in Appendix A.

Using equation 2, we calculate unconditional consumption expenditure distribution differences for both social identities from the mean of their own group in the beginning and the recent survey year during the reform period which are plotted in Figures 4 and 5. The interesting facts these graphs provide are the following. First, it suggests that alienation has risen at the extreme top quantiles whereas consumption expenditure differences at any quantile from the 30th to 60th quantiles and for each social identity remain the same. Second, the widened difference at 90th quantile between privileged to marginalized groups and Hindu to non-Hindus over time has given a clear indication of greater alienation at the top end of the distribution. At the bottom end i.e., at 10th quantile, and for both social identities, differences in consumption expenditure remain the same, suggesting that the poor has not remarkably experienced the same benefit from growth as individuals at the top end. Third, comparing the shift in consumption expenditure with reference to zero line, our finding suggests that the privileged (HC) group has benefited more compared to marginalized groups (SC and SC) especially after 60th and above quantiles. A similar pattern is also observed between Hindus and non-Hindus, suggesting that

Figure 4: Normalized consumption expenditure distribution of privileged and marginalized groups in 1993 and 2011.

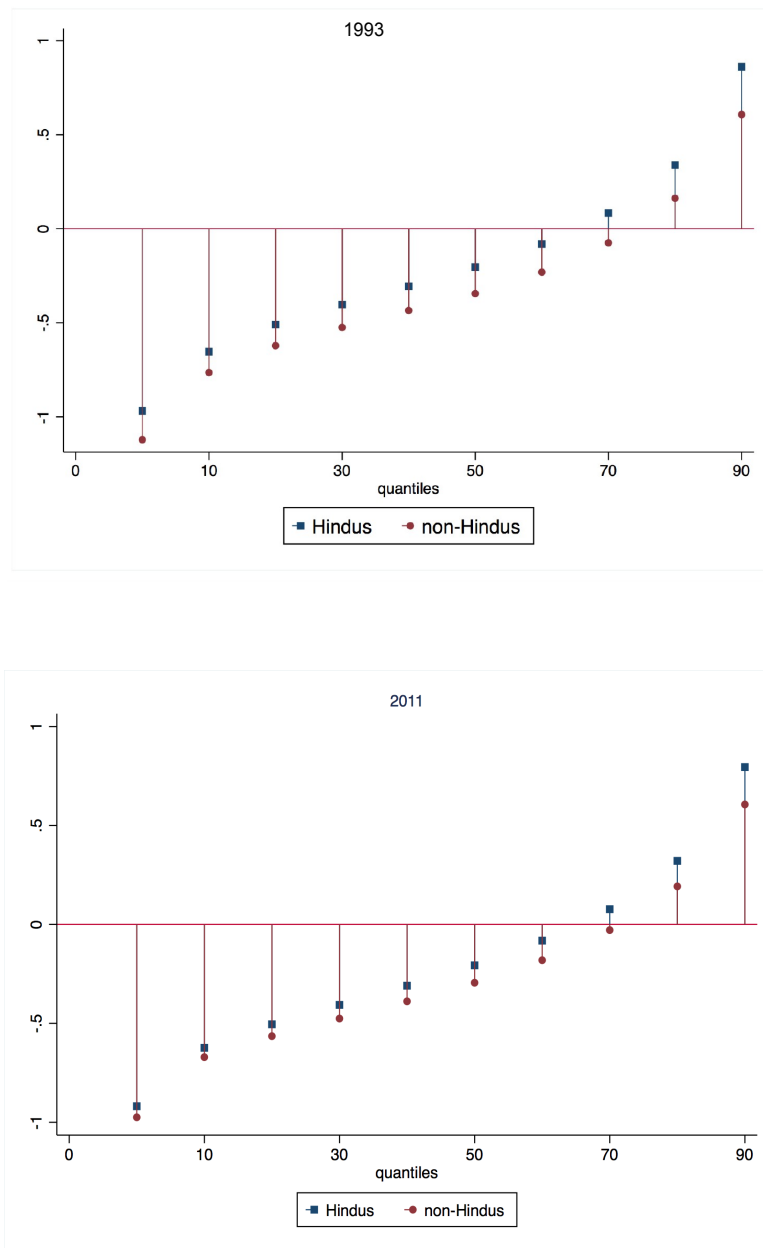


Source: Authors' calculation

Hindus have benefited more from growth compared to non-Hindus.

In our empirical results, we use this normalized consumption for distance and looked at the impact of social identities on consumption distance. For better intuition of our results, two important points are worth noting. First, as we move from bottom to top end of the distribution, one should expect the same sign of a particular covariate. Second, the coefficient difference from median to lower quantiles and from median to higher quantiles, in absolute terms, should decline over time with both the level of consumption and the normalized consumption expenditure. In

Figure 5: Normalized consumption expenditure distribution of Hindus and non-Hindus in 1993 and 2011.



Source: Authors' calculation

Table 2, we present the expected sign for our covariates.

## 4 Empirical findings and analysis

India's growth dynamics in the post-economic reform period since the early 1990s present an interesting contrast in terms of real GDP per capita, education, and consumption expenditure, as

Table 2: Expected sign of coefficients across the quantiles

<i>Variables</i>	<i>Expected sign</i>
Marginalized groups	-
Non-hindus	-
Education	+
Age	+
Regular earnings	+
casual labour	-
Non-agriculture	-
Others type work	+

Source: Authors' predictions

there has been an uneven rise in consumption distance across social groups.<sup>10</sup> At the aggregate level, we show that the rising consumption distance appears to coexist with higher level of per capita income at region level within India. Using state-level per capita income data for those 3 survey years, we show that although there is a positive correlation between alienation and development, there has been a downward slope change over time (see Figure 8). Besides, it is likely that education might have helped improving consumption pattern among different social groups and thereby lowering alienation.

We begin our analysis by presenting the expected relationships between consumption level and social identities along with the household level control variables, showing the advantage of QR approach over a mean-based OLS regression. Consumption distance constructed using the approach in the previous sub-section could reflect how social identities are important in determining consumption behaviour across distribution and over time. In Table 6, we present the result using consumption expenditure across the distribution whereas in Table 3, we use normalized value of consumption expenditure, as a dependent variable, and present the estimates using quantile regression; however for the discussion purpose and to save space we only present three quantiles - 10th (lower), 50th (median) and 90th (higher). The results for other quantiles, using the level of consumption expenditure as a dependent variable, are graphically presented in Appendix A (see Figures 5-8). The solid red line corresponds to the OLS coefficients and red dashed lines are confidence bands. The black dashed-dot lines are quantile coefficients and the dark shadow area reflects confidence intervals.

Our aim is to show the effect of social identities on normalized consumption expenditure i.e., on consumption distance. We, therefore, using Table 3, as a baseline result, discuss the rest of the results. In our baseline results, we assume that education is exogenous and thus does not directly govern individual consumption expenditure. Under this assumption, both OLS and QR estimation will result in unbiased and efficient estimates.

<sup>10</sup>Since the inception of economic reform in 1991, the growth of real GDP per capita has shown an unprecedented rise averaging about 5.6% per annum, with literacy improvement among young and household consumption expenditure with manifold growth over the years.

## 4.1 Baseline results

As mentioned before, the mean-based OLS approach only examines the effect at mean, and thus ignores the entire distribution. Therefore the variability in consumption distance at various points of the consumption distance distribution for different social groups can be better revealed through the QR based conditional distribution approach.

- *Impact of caste on consumption distance*

Examining the effect of caste, as a social identity, on consumption distance, we find some important and interesting facts.

First, looking at the median quantile coefficients, we do not observe large variation over time in marginalized groups. The negative coefficients for marginalized group across the distribution and over time, however, suggest that the marginalized groups are not better off than their counterpart privileged group even after 25 years of reform. In fact, our finding suggests that there is no confounded effect of redistributive policies<sup>11</sup>, which have been in place since the 1950s in order to improve their social status. The possible reason for such wide difference in their consumption distance could be linked to existing historical hierarchies between marginalized and privileged groups which lead them to accept low paying jobs even if marginalized groups are endowed with better human capital. As a result of this, they experience persistent gap in consumption expenditure relative to their counterpart privileged group.

Second, comparing the coefficient difference from the median quantile to the bottom quantile, and the median quantile to the top quantile, and in each period, we observe that those individuals who consume above the mean consumption are more integrated but they are still alienated compared to their counterpart at the same quantile (Table 3). For instance, comparing the coefficient value (-0.056) at the 10th quantile to the coefficient value (-0.079) of median quantile, and from median quantile to the coefficient value (-0.167) at the 90th quantile in 1993, showing the presence of strong relative inequality at the top end of the distribution.

- *Impact of religion on consumption distance*

Looking at the impact of religious affiliation as a social identity on consumption distance, our finding shows that non-hindus are alienated relative to hindus at any quantile, although there is an improvement at the median quantile in each survey round. Our finding strongly suggests that non-Hindus even after thirty years of reform are still placed in a low socio-economic position. These findings thus lead to two important implications. First, it links to political incentivization issues of economic resources which are normally used during electoral campaigns and frequently debated in the literature on the relative bias of allocation of resources between Hindus and non-Hindus. Second, our results link to conflict theory on social identity and provide reasons for social riots between two groups. For example, Varshney (2001) observed that difference in economic status between Hindus and Muslims in Ahmedabad, the Indian city of Gujarat, led to the much-condemned conflict which killed people from both religions. Another implication of

---

<sup>11</sup>In India, affirmative policies are well known in the form of reservation policy or quota system for scheduled castes and tribes, where a certain percentage of jobs or school vacancies must be reserved for members of these groups.



these findings also links to the recent argument of aspirations failure (Ray, 2003; and Dalton et al., 2014), which suggests that if differences in consumption level persist for a long time, then Hindus may dominate non-Hindus in the long run.

- *Impact of education on consumption distance*

It is theoretically argued in the literature that the rise in human capital formation in any society can offset various external negative consequences of both economic and social development. In this regard, we have also examined whether the rise in human capital has helped in reducing social inequality or not. There is no doubt that since the 1990s, i.e., following economic reforms, India has experienced unprecedented gains in educational attainment both at aggregate level and across the states. Indeed, upward mobility in consumption expenditure across all educational categories is observed (see Table 1). This has, then, raised two important questions in the context of this research. First, do households with higher level of education across the consumption distribution improve an individual's relative position in the social hierarchy and thus lower consumption inequality in a society? Research on this view, using education as a conspicuous good or signaling status, showed that lack of education has not only hindered the social mechanisms but has also affected the national productivity (Khamis et al., 2012). An interesting account of such practice can be found, for instance, in the work of Srinivas (1956) and Ambedkar (1944) where both individual socially-inherited position and different levels of educational attainment have been used to understand its interdependence from the growth perspective.

In light of our research questions, the estimated results in Table 6 strongly suggest that the effect of education has not only improved the level of consumption across the distribution but it has also helped lowering consumption distance across the distribution and over time (see Table 3). For instance, its effect is noticeable more at higher (90th) quantile compared to lower (10th) quantile of the consumption distance. We observe that the education difference from median to 90th quantile in 2011 has declined by one-third compared to the period 1993 for the same quantiles where it is not the same from 10th quantile to 50th quantile and over time.

Table 3: Effect of social identities on consumption distance across distribution and over time

	1993					2004					2011				
	OLS	10th	50th	90th		OLS	10th	50th	90th		OLS	10th	50th	90th	
<i>Privileged groups (excluded)</i>															
Marginalized groups	-0.110*** (0.009)	-0.056*** (0.005)	-0.079*** (0.006)	-0.167*** (0.020)	-0.145*** (0.008)	-0.050*** (0.003)	-0.106*** (0.005)	-0.232*** (0.020)	-0.116*** (0.009)	-0.051*** (0.004)	-0.090*** (0.005)	-0.151*** (0.017)			
<i>Hindus (excluded)</i>															
Non-hindus	-0.122*** (0.012)	-0.064*** (0.006)	-0.100*** (0.008)	-0.139*** (0.026)	-0.159*** (0.010)	-0.062*** (0.005)	-0.122*** (0.007)	-0.222*** (0.027)	-0.087*** (0.010)	-0.033*** (0.005)	-0.073*** (0.006)	-0.118*** (0.021)			
Education	0.082*** (0.001)	0.026*** (0.001)	0.047*** (0.001)	0.138*** (0.002)	0.047*** (0.001)	0.021*** (0.001)	0.044*** (0.001)	0.076*** (0.003)	0.073*** (0.001)	0.023*** (0.000)	0.046*** (0.001)	0.109*** (0.002)			
Age	0.003*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.004*** (0.001)	0.001*** (0.000)	-0.003** (0.000)	0.001 (0.000)	0.002*** (0.001)	0.004*** (0.000)	0.001*** (0.000)	0.003*** (0.000)	0.006*** (0.001)			
<i>Self-employed (excluded)</i>															
Regular earnings	0.146*** (0.026)	0.105*** (0.013)	0.112*** (0.016)	0.130** (0.056)	0.110*** (0.009)	0.014*** (0.004)	0.050*** (0.006)	0.231*** (0.023)	0.038*** (0.009)	0.017*** (0.004)	0.036*** (0.005)	0.054*** (0.017)			
Casual labour	-0.104*** (0.025)	-0.087*** (0.013)	-0.071*** (0.016)	-0.104* (0.054)	-0.018 (0.011)	-0.024*** (0.005)	-0.036*** (0.007)	-0.049* (0.029)	0.038*** (0.011)	-0.023*** (0.005)	0.001 (0.006)	0.037* (0.022)			
Non-agriculture	-0.225*** (0.028)	-0.146*** (0.014)	-0.171*** (0.017)	-0.245*** (0.059)	0.088*** (0.009)	0.012*** (0.004)	0.084*** (0.006)	0.180*** (0.024)	-0.088*** (0.016)	-0.029*** (0.007)	-0.044*** (0.009)	-0.141*** (0.033)			
Other types work	0.197*** (0.027)	-0.003 (0.014)	0.153*** (0.017)	0.428*** (0.059)	0.452*** (0.011)	0.069*** (0.005)	0.257*** (0.007)	1.020*** (0.028)	0.225*** (0.014)	0.041*** (0.006)	0.154*** (0.008)	0.445*** (0.028)			
Constant	-0.410*** (0.030)	-0.633*** (0.015)	-0.381*** (0.019)	-0.058 (0.063)	-0.191*** (0.019)	-0.744*** (0.008)	-0.401*** (0.013)	0.609*** (0.049)	-0.594*** (0.018)	-0.754*** (0.008)	-0.555*** (0.010)	-0.169*** (0.036)			
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Observations	69152	69152	69152	69152	109320	109320	109320	109320	89623	89623	89623	89623			
R-squared	0.122	0.101	0.832	0.1524	0.058	0.0366	0.0442	0.0614	0.0911	0.0889	0.0729	0.1155			

Note: Authors' calculations based on NSS data. Dependent variable is consumption distance calculated from the mean consumption level and bounded with the standard deviation. Heteroscedasticity corrected robust standard errors are reported in brackets. \* Significant at  $p < 0.10$ , \*\* Significant at  $p < 0.05$ , \*\*\* Significant at  $p < 0.01$ .

Our finding is consistent with the theoretical insights of Gradstein and Justman (2002), who have suggested that improving human capital should minimize social distance. This also means that households those who are endowed with the low level of human capital are less socially integrated compared to those who are endowed with the higher level of human capital. The interaction between education and social identity shows that the households with higher level of education in both groups have significantly benefited in minimising their distance as reflected in lower coefficients over time (see Table 7 of Appendix B). The interactive effect therefore reveals that education has played a significant role in reducing the consumption disparity among those who are consuming above the mean consumption level.

- *Effect of household types on consumption distance*

Analysing the relationship between different household types on consumption distance, we find mixed results for some household types. However, a strong presence of social alienation, regardless of the household types, is observed for those who are above the median consumption level. Overall, it is observed that households those who are regular earners are less alienated than those who are classified in other categories. A slope difference test is also performed to test the slope inequality from 10th to 50th quantile and from 50th to 90th quantile with the level of consumption (see Table 3).

The results in this section show the impact of socio-economic indicators on consumption distance at the mean and across distribution. In sum, our results strongly suggest that individual social identity matters to govern the consumption differences across distribution while the mitigating effect has been noticed between educational attainment level and consumption distance.

- *Robustness test - I*

As argued above, changing the individual reference point from mean to another reference point such as the median does not change the direction and magnitude of the main results already discussed. To confirm this, several other alternative reference points are also considered. The obtained results from these reference points are presented in Appendix B in Table 7.

## 4.2 Dealing with Endogeneity - IV quantile regression approach

Our baseline result indicates that education has played a significant role for both social identities in reducing their consumption distance over time. Households with higher level of education tend to be more integrated in our baseline results. However, possible endogeneity of education suggests that the education variable could be correlated with the error term, which we address by using the instrumental variable regression. To confirm this endogeneity, we run the Wu-Hausman test<sup>12</sup>, which rejects the null of no endogeneity of education (see Table 4). Chernozhukov and

---

<sup>12</sup>To show the endogeneity of education, the reduced form of the baseline model is run by taking education as a dependent variable with all the other exogenous variables, and the predicted value (say *eduhat*) is obtained. Then, the residual (*edures*=*education*-*eduhat*) is added to the original model and tested for significance in the baseline regression. As the coefficient on '*edures*' is statistically different from zero, we can conclude that education is indeed endogenous.

Hansen (2005 and 2008) instrumental variable quantile regression (IVQR<sup>13</sup>) is further used to obtain consistent and robust coefficient estimates across the distribution. The underlying idea of the IVQR approach is described in Appendix A.

As part of the instrumental variable approach, we need to identify valid instruments for education which satisfies the exogeneity assumption. In order to satisfy the exclusion restriction on our endogenous variable of interest (education), we need at least two instruments and therefore we use total number of schools and total school enrollment in each district in every state for the two survey rounds - 2004 and 2011. Our estimation is now more robust in addressing the endogeneity of education (years of schooling at household level). This data come from "http://schoolreportcards.in" starting from 2000 onwards, and we therefore test the endogeneity of education only for two survey years - 2004 and 2011. The rationale behind choosing the number of schools and school enrolment as instruments is that it is directly related to the years of schooling at household level and therefore independent from the level of individual consumption. These instruments suggest that districts with higher number of schools and higher level of enrollment tend to have households with higher level of education, but they do not govern the consumption pattern directly. Thus the total number of schools and total enrollment are the most relevant instruments to solve possible endogeneity of education. The validity test for the instruments<sup>14</sup> at mean level and the results obtained from IVQR are presented in Table 4 and across the distribution.

Overall, we find the instruments relevant which have the expected effect on the endogenous variable. The IVQR results show that the effects of social identities, education and other control variables on consumption distance are consistent with our baseline results. Compared to Table 3, the estimated coefficient is large in magnitude and statistically significant at least at 5% level for all control variables. The larger effect implies a downward bias in our baseline estimates, which according to our interpretation above, would suggest that improving human capital helps minimise alienation in the society and its larger effects are noticeable at higher end of the distribution.

The relevance and validity of the instrument used is also confirmed by the under-identification and weak-identification tests (0.05) along with the over-identification test. The estimated test results are presented in Table 4. The first-stage results demonstrate that our set of instruments is indeed capable of explaining a large fraction of the variation. The F-tests easily reject the null hypothesis of no effect from the instrument set. In addition to the F-statistics reported in the first-stage regressions, we find that both the Kleibergen & Paap (2006) LM- and F-test reject the null hypothesis of under-identification and weak identification. The Hansen over-identification J-test examines whether the instruments are orthogonal to the error in the regression, and the high p-value suggests that the instruments are valid as we do not reject the null hypothesis of no over-identification.

In general, the instrumental variable approach appears to be supportive of our baseline

---

<sup>13</sup>The IVQR approach is known to be computationally efficient for a small number of endogenous variables (which is typically the case in our investigation). It should also be noted that the IVQR captures the impact of changes in covariates upon a conditional distribution of dependent variable, in very much the same way that the QR regression measures.

<sup>14</sup>To the best of authors' knowledge, there are no sound test procedures to establish the strict identification test at each quantile. Therefore, we only satisfy the relevance and exogeneity condition at mean and compare the validity of results close to the median quantile (50th) IVQR.

findings. Improvements in educational attainments minimize consumption distance across social identities, emphasising education's welfare-promoting and status-changing roles in consumption dynamics. This is in line with the literature that an improvement in human capital can minimize the effects of social alienation (Gradstein and Justman, 2002). In other words, our results suggest that the households endowed with low level of human capital are less diversified and socially less integrated compared to those who are endowed with higher level of human capital. As educational attainment rises from primary to tertiary level, the positive effect of education on consumption distance distribution suggests that households with higher level of education tend to be more integrated than being alienated, as they are able to consume more than the group mean. In light of this evidence, it is not unusual to anticipate that with better level of educational attainment, the consumption level gap for caste (privileged-marginalised) and religion (Hindus-Muslims) based groups should be lower over time.

Table 4: Instrumental variable results at mean and across the distribution

	2004				2011			
	OLS	10	50	90	OLS	10	50	90
<i>Second-stage</i>								
Marginalized groups	-0.283 (0.025)	-0.179 (0.027)	-0.189 (0.031)	-0.371 (0.043)	-0.011 (0.002)	-0.164 (0.052)	-0.194 (0.045)	-0.235 (0.063)
Non-Hindus	-0.232 (0.023)	-0.183 (0.035)	-0.143 (0.021)	-0.298 (0.033)	-0.007 (0.002)	-0.149 (0.041)	-0.162 (0.044)	-0.474 (0.023)
Education	0.051 (0.015)	0.024 (0.013)	0.031 (0.009)	0.044 (0.025)	0.031 (0.001)	0.03 (0.013)	0.043 (0.018)	0.117 (0.041)
<i>First-stage</i>								
Marginalized groups	-1.664 (0.023)				-1.476 (0.034)			
Non-Hindus	-1.138 (0.031)				-1.443 (0.043)			
Number of school (log)	0.597 (0.019)				0.321 (0.032)			
School Enrollment (log)	0.280 (0.022)				0.237 (0.025)			
N	94844				66900			
Other control variables	Yes				Yes			
First stage R-squared (centred)	0.1536				0.1479			
Wu-Hausman (F -test/P-value)	40.424/0.000				36.777/0.001			
Weak-identification test (F-test)	70.98/000				61.72			
Under-identification tests:LM test (P-value)	0.002				0.001			
Hasen J Over-identification test (P-value)	0.7311				0.2136			
Second stage R-squared (centred)	0.1435				0.1363			

Note: (i) Dependent variable is consumption distance. Heteroscedasticity corrected robust standard errors are reported in brackets. (ii) Education (number of schooling years) is instrumented by number of schools and school enrollment at district level. \* Significant at  $p < 0.10$ , \*\* Significant at  $p < 0.05$ , \*\*\* Significant at  $p < 0.01$ .

To confirm this, we use the counterfactual decomposition approach as proposed by Machado and Mata<sup>15</sup> (hereafter MM, 2005) in section 4.3 where we examine systematically the consumption differential for both social identities across the consumption distribution in unmatched and matched samples.

### 4.3 Digging deeper: Counterfactual approach

In the previous sub-sections, we discussed our first research question as: *to what extent social identities determine consumption distance across the distribution at household level*. The question that needs to be explored is *what would have occurred if the observed characteristics of the privileged or majority religion group were assigned to the other group as a counterfactual for a group-level analysis*. The main idea here is to understand whether the estimated consumption gap between two groups can be lower if a counterfactual group is created by assigning the covariates of the privileged group to the marginalised group (or from Hindus to non-Hindus).

We have already looked at the unconditional pattern of consumption expenditure gap for both social identities - caste and religion, through density plots which show large consumption gap in both groups and over time, indicating an uneven society. Thus the MM counterfactual decomposition can help confirm whether the estimated gap can be lowered if the characteristics of the high consumption group are assigned to the low consumption group. In Table 5, the results obtained from the counterfactual decomposition are presented.

First, the consumption gap rises over time (see column: 1, 5, and 9 in Table 5) for both identities (religion and caste). Considering the case of religion in 1993, the estimated gap is higher across quantiles and we observed approximately two-fold higher estimated consumption gap at the 90th quantile (0.198) compared to the 10th quantile (0.095). Both unobserved (column 3, 7, and 11 in Table 5) and observed (column 4, 8, and 12 in Table 5) coefficient effects have also marked a sharp rise. We observe that the proportion of total gap in consumption is due to the unobserved inequality (column 3, 7, and 11 in Table 5) and it is higher in 2011 compared to 1993 which explains why the estimated overall gap has grown twice as much.

What is notable is that the estimates of the gap exhibit monotonic rise for the minority religious group which is consistent across bottom and top quantiles. In particular, we observe that for 1993 the estimated gap is 0.095 at 10th quantile, which has increased to 0.153 and 0.184 in 2004 and 2011 respectively. A similar trend is observed for 90th quantile, although the rate of change at this quantile is noticeably smaller than at 10th quantile. The observed overall trend comes from the unexplained part (column 3,7, and 11 in Table 5) for all quantiles. Our aim here was to show that even after exchanging the characteristics of the high consumption group to low consumption group, the big part of the estimated consumption gap (59% at the median quantile) is still being explained by the high consumption group (Hindus) consuming more than the counterfactual group (Muslims with Hindu characteristics) possibly due to their inherent social position; nevertheless the counterfactual minority group with exchanged characteristics is certainly better off relative to their own characteristics, explaining 40% of the gap at the median quantile in 2011.

When caste differences are taken as a measure of social identity, the decomposition of esti-

---

<sup>15</sup>The steps of Machado and Mata and the matching approach are discussed in detail in Appendix A

mated consumption gap depicts a similar pattern that the privileged group still consumes more than the counterfactual marginalised group, explaining around 77% of the estimated consumption gap at the median quantile in 2011-12, suggesting other inherent identity-specific social factors as possible contributors to consumption gap between groups. We find an overall rise in the gap over time at lower quantile (from 0.201 in 1993 to 0.310 in 2011), with a slight decline (0.194) in 2004. At higher quantile (90th), we do not find any marked rise in the gap, although, at median quantile, a monotonic rise is still observed over the years.

- *Robustness test - II*

As a robustness, we further investigate how sensitive the above results are to a correction of specification bias in unmatched samples. To correct this, we employ  $\tilde{N}$ opo (2008) matching approach and then re-estimate the counterfactual model for both social identities. In Table 8, we report the results of counterfactual decomposition at various quantiles after matching individual characteristics for each social identity, i.e., caste affiliation of individuals. The decomposition is undertaken between privileged and marginalized groups for all survey years and across the distribution. An important and striking result that emerges after matching is that there is observed persistence and monotonic rise in consumption gap across quantiles (for each survey year). This can be seen by comparing the 10th, 50th and 90th quantiles of consumption gap (see column: 1, 5, and 9 of Table 8). For instance, in 1993-94, the gap at 10th quantile is 0.201, which has increased to 0.245 at median quantile and observed almost twice at 90th quantile (0.389) compared to 10th quantile.

Similar trend is observed for 2004-05 and 2011-12 surveys. We also observe a persistent rise in consumption gap over time between privileged and marginalized groups during 1993-94 and 2011-12 at 10th quantile (from 0.201 to 0.262). At 90th quantile, this gap also manifested an increasing trend from 0.389 in 1993-94 to 0.444 in 2011-12. However, we observed a significant and larger increase in gap at median quantile (from 0.245 to 0.390 between these two sample years).

In sum, the trend of rising consumption gap is similar to the unmatched sample case (Table 5); however the extent of increase in consumption gap is different - for instance, in Table 5 and for privileged-marginalized case, the temporal rise in consumption gap goes from 0.251 in 1993-94 to 0.382 in 2011-12 at 50th quantile, which is slightly smaller than the one for matched sample (which goes from 0.245 to 0.390 in 1993-94 and 2011-12 respectively). We also notice that at the lower quantile, the consumption gap is larger in unmatched sample between 1993-94 and 2011-12 than in the matched sample. These inferences of our finding are graphically presented in Figure 6 and compared with Figure 7.

Our overall conclusion from the matching exercise is as follows: (i) (across quantile and over time) the overall trend in consumption gap by social identities is similar to the matched one. This establishes robustness of our results irrespective of sample design and differences in characteristics. (ii) By matching individuals' characteristics, we find slightly smaller persistence effects of consumption gap and that the results between matched and unmatched vary considerably at lower and higher quantiles. This latter result points to the possible distortionary effect that the matched samples may experience at low and high end of the distribution.

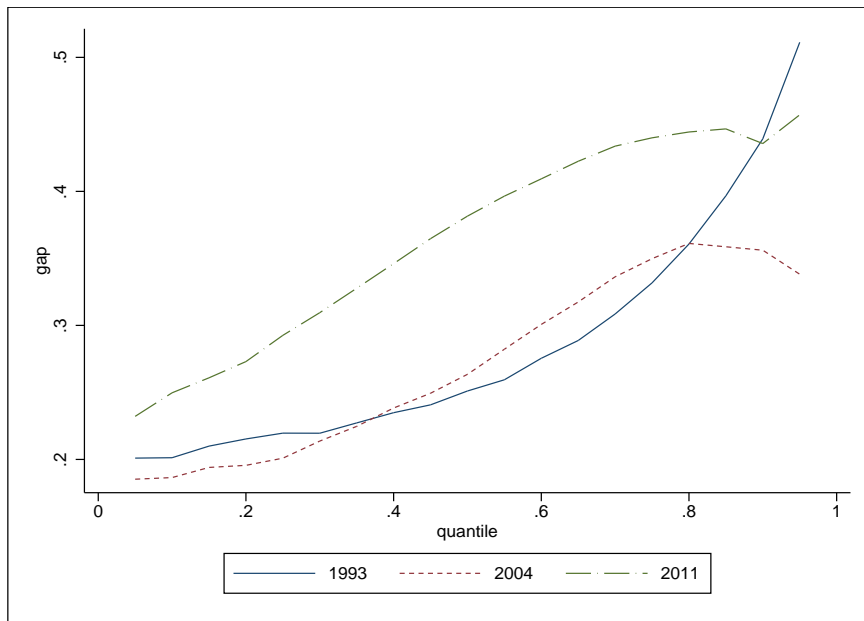


Table 5: Counterfactual distribution results between privileged and marginalized groups (Unmatched samples)

Quantile	1993				2004				2011			
	Gap (1)	Endowment (2)	Coefficients (3)	Interaction (4)	Gap (5)	Endowment (6)	Coefficients (7)	Interaction (8)	Gap (9)	Endowment (10)	Coefficients (11)	Interaction (12)
Hindu-Muslims												
10	0.095	0.044 (0.021) 46%	0.045 (0.023) 47%	0.006 6%	0.153	0.062 (0.032) 40%	0.085 (0.027) 55%	0.006 4%	0.184	0.070 (0.023) 38%	0.103 (0.024) 56%	0.011 6%
50	0.123	0.055 (0.019) 45%	0.065 (0.018) 53%	0.003 2%	0.157	0.065 (0.021) 41%	0.087 (0.020) 56%	0.005 3%	0.172	0.068 (0.022) 40%	0.102 (0.021) 59%	0.002 1%
90	0.198	0.080 (0.038) 41%	0.113 (0.036) 57%	0.005 2%	0.209	0.072 (0.039) 34%	0.127 (0.040) 61%	0.010 5%	0.212	0.060 (0.031) 28%	0.149 (0.034) 71%	0.003 1%
Privileged-Marginalized												
10	0.201	0.063 (0.020) 31%	0.132 (0.019) 65%	0.007 3%	0.194	0.086 (0.025) 44%	0.102 (0.023) 53%	0.005 3%	0.310	0.095 (0.020) 31%	0.210 (0.023) 68%	0.002 1%
50	0.251	0.109 (0.021) 43%	0.154 (0.019) 61%	-0.012 -5%	0.264	0.133 (0.021) 51%	0.144 (0.019) 54%	-0.013 -5%	0.382	0.084 (0.024) 22%	0.293 (0.020) 77%	0.002 1%
90	0.439	0.171 (0.040) 39%	0.250 (0.032) 57%	0.019 4%	0.356	0.173 (0.044) 49%	0.180 (0.040) 50%	0.003 1%	0.436	0.091 (0.033) 21%	0.356 (0.035) 82%	-0.011 -3%

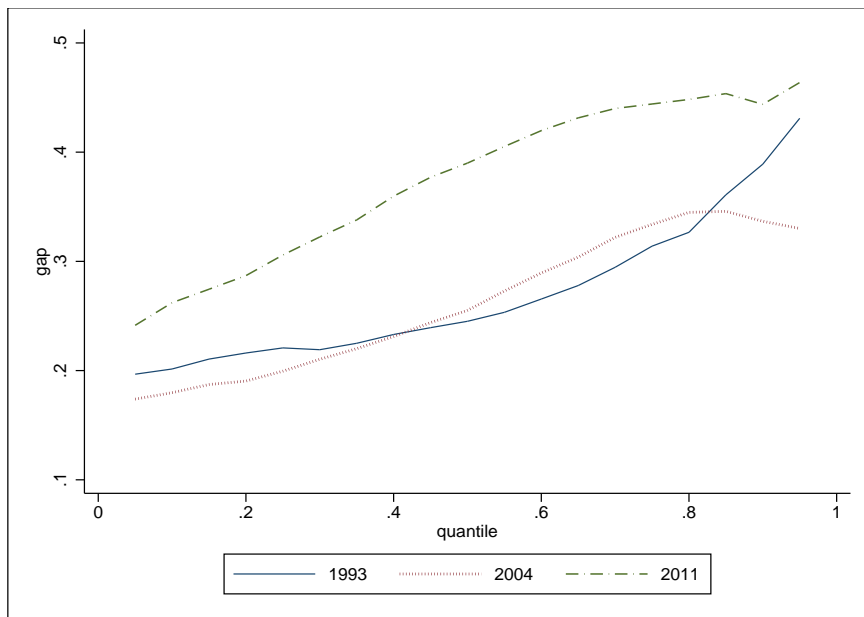
Note: (i) bootstrap standard errors are in parenthesis, and proportions to observed gap are mentioned under standard errors. (ii) The terms Endowments and Coefficient reflect observed and unobserved inequality in the table.

Figure 6: Consumption gap between Privileged and Marginalized groups - unmatched samples



Source: Authors' calculation

Figure 7: Consumption gap between Privileged and Marginalized groups - matched samples



Source: Authors' calculation

## 5 Conclusions and Policy Implications

- *Conclusion*

The main aim of this paper was to incorporate the theoretical idea of Akerlof (1997)'s social distance theory and to understand how individual affiliation to any particular social group explains their alienation. In the recent years, concerns have been raised about lack of social cohesion despite the continued high growth in many emerging market economies. Our empirical evidence on multimodal distribution of consumption expenditure lends support to the idea of multiple equilibria or multiple clusters/modes in the consumption distribution. Given little empirical evidence on this issue of alienation, in this paper using household survey data from India, we estimated the effects of inherited social identity on consumption distance as a measure of alienation. The main reason for using consumption deviation from group mean is that group members in a particular identity are endowed with similar attributes and experience similar social conditions in their respective community groups.

We calculated distance for each group from the mean consumption of their own group, which clearly implies that the groups are facing within-group inequality but we cannot ignore between-group inequality. We examined whether one group's distance (from a specific norm or threshold) is lower than the other group. Even after 25 years of economic reform, we find that the identified social groups (marginalised and minority religious groups) remain alienated in the distance distribution. Having controlled for the effects of education, age, and household type, we found that there is significant evidence of a monotonic rise in consumption distance across quantiles and over time, and the distance of two groups (marginalized groups in terms of caste, and for non-Hindus in terms of religion) relative to their counterpart gets minimised with the help of higher educational attainment. In addition, we re-estimated our base line regressions considering two exogenous instruments namely number of schools and school enrollment in order to satisfy the exclusion restriction on our endogenous variable of interest (education). Our estimation is more robust in addressing the endogeneity of education (years of schooling at household level). Both instruments suggest that districts with higher number of schools and higher level of enrollment tend to have households with higher level of education.

Esteban and Ray (1999) observed that unequal distribution of wealth - especially in developing countries - facilitates bribery and corruption of politicians and triggers intense rent-seeking. Khamis et al. (2012) related consumption inequality to the concept of status signalling and observed large and significant difference across groups with distinctive social identities. They describe that differences in the consumption pattern across groups are mostly linked to preferences and norms of society. In this context, considering consumption differences, our results remain consistent with the literature while supporting the broad theoretical conclusions of Cont and Lowe (2010) that multiple equilibria could result from conflicts between various social groups.

If social identity is the only reason that governs individual consumption pattern, then what would have happened to average consumption differences between groups if one exchanges their characteristics namely educational and demographic factors. In this regard, we used the counterfactual regression analysis and observed that even if we exchange their characteristics, the counterfactual deprived groups (marginalized groups in the case of caste, and non-Hindus

in the case of religion) are still not able to consume as much as their counterparts (although slightly higher than their own consumption), and therefore one's social identity does matter in explaining where a household is in the consumption distribution.

Widening income/consumption gap between groups in our study suggests the reasons for social conflict between groups over opportunities or availability of public goods and thus strongly links to the concept of polarization<sup>16</sup>. Besides, such inequality leads to group conflict, where poor individuals are feeling strong antagonism against the rich in the same group, providing potential reasons for social discrimination and possible conflict. Overall, the widening gaps could indirectly imply policy ineffectiveness, which requires further investigation. Since, social cohesion (alienation) is a multi-dimensional concept, it requires a direct measure or non-economic dimensions such as trust, civic cooperation and confidence in law, to understand its complete effect in a given society. We did not incorporate these non-monetary dimensions because of data limitations.

- *Policy relevance*

Our results from the policy perspective point out that the rise in consumption gap between groups weakens the mechanism of social cohesion, which in turn could affect the current path of economic development. If the goal of public policy is to make economic growth inclusive, then the promotion of social cohesion through the strengthening of social networks, trust and tolerance, and motivating people to show solidarity to obey social rules and to engage them with different channels of economic development, are promising roads to inclusive development. We stress that our results on the effect of education in minimizing distance are only suggestive of policy actions, particularly in the areas of fiscal, employment and social policies to improve access to opportunities for marginalised sections of the society that can help reduce the growing consumption distance identified in this paper.

An implication of our results suggests that the adopted reservation policies for marginalized groups are probably effective to some extent in low consumption groups, whereas in privileged groups, some other factors are contributing towards the higher persistence of consumption gap. The lower gap between privileged and marginalized groups at the bottom-end of the distribution, thus, gives a clear indication of the success of reservation policies in protecting marginalized groups from discriminatory practices in various public sector organisations. Besides, as shown in this paper, education expansion for the marginalised group should be promoted by the government which can be alienation-reducing.

## References

- [1] Akerlof, G.A. (1997), Social Distance and Social Decisions, *Econometrica*, 65(5), 1005-27.
- [2] Ambedkar, B. R. (1944), *Annihilation of Caste with a Reply to Mahatma Gandhi*, (3rd edition).

---

<sup>16</sup>Polarization here only refers to the alienation that individuals and groups feel from one another as opposed to what is described in Duclos et al. (2004).

- [3] Andrews, D. and Leigh, A. (2009), More inequality, less social mobility, *Applied Economics Letters*, 16 (15), pp. 1489-92.
- [4] Bandyopadhyay, S. (2012), Convergence Clubs in Incomes across Indian States: Is There Evidence of a Neighbours' Effect? *Economics Letters*, 116, 565 -70.
- [5] Banerjee, A. and Duflo, E. (2007) The economic lives of the poor. *The Journal of Economic Perspectives*, 21, no. 1, 141- 67.
- [6] Basu, S. and Mallick, S. (2008) When does growth trickle down to the poor? The Indian case. *Cambridge Journal of Economics*, 32 no. 3, 461 - 77.
- [7] Bogardus, Emory S. (1933), A Social Distance Scale, *Sociology and Social Research*, 17:26571.
- [8] Chernozhukov, V., and Hansen, C. (2005), An IV model of quantile treatment effects, *Econometrica* (73), 245-61.
- [9] Chernozhukov, V., and Hansen, C. (2008), Instrumental variable quantile regression: A robust inference approach, *Journal of Econometrics* (142), 379-98.
- [10] Cont, R. and M. Lowe (2010), Social distance, heterogeneity and social interactions, *Journal of Mathematical Economics*, 46: 572-90.
- [11] Dalton, P. S., Ghosal, S. and Mani, A. (2016), Poverty and Aspirations Failure. *The Economic Journal*, 126: 165-88.
- [12] Corak, M. (2013), Income inequality, equality of opportunity, and intergenerational mobility, *The Journal of Economic Perspectives*, pp. 79 -102
- [13] Deaton, A. (1997), Poverty Among Children and the Elderly in Developing Countries, Research Program in Development Studies (Unpublished manuscript).
- [14] Deaton, A. and J. Muellbauer (1980), An Almost Ideal Demand System, *American Economic Review* 70(3), 312-26.
- [15] Deaton, A. and J. Muellbauer (1986), On Measuring Child Costs: With Applications to Poor Countries, *Journal of Political Economy* 94: 720- 44.
- [16] Dev, S.M. (2013), India development report, Oxford University Press, New Delhi (2013)
- [17] Duclos, J-Y, J. Esteban, and D. Ray (2004), Polarization: Concepts, Measurement, Estimation, *Econometrica* 72, 1737-72.
- [18] Dutta, I., Madden, P. and Mishra, A. (2014), Group inequality and conflict, *The Manchester School*, 82 (3), 257-83.
- [19] Esteban, J. and D. Ray (1999), Conflict and Distribution, *Journal of Economic Theory* 87, 379 - 415.

- [20] Esteban, J. and D. Ray (2011), A Model of Ethnic Conflict, *Journal of the European Economic Association* 9, 496 - 521.
- [21] Esteban, J., Mayoral, L., Ray, D. (2012), Ethnicity and Conflict: An Empirical Study, *American Economic Review*, 102(4): 1310 - 42.
- [22] Genicot, G and Ray, D. (2006), Contracts and externalities: How things fall apart, *Journal of Economic Theory*, Elsevier, vol. 131(1), pages 71-100
- [23] Gradstein, M. and M. Justman (2002), Education, Social Cohesion, and Economic Growth, *American Economic Review*, 92(4): 1192 -1204.
- [24] Jayadev, A., and S.G. Reddy (2011) Inequalities between Groups: Theory and Empirics, *World Development*, 39 (2): 159-173.
- [25] Khamis, M., N. Prakash, and Z. Siddique (2012), Consumption and social identity: Evidence from India, *Journal of Economic Behavior & Organization*, 83 (3): 353 - 371.
- [26] Kijima, Y. (2006), Why did wage inequality increase? Evidence from urban India 1983-99, *Journal of Development Economics*, 81(1): 97 - 117.
- [27] Koenker, R. and Bassett, G. (1978), Regression Quantiles, *Econometrica*, 46 (1): 33 - 50.
- [28] Kumar, T.K., Mallick, S.K. and Holla, J. (2009) Estimating consumption deprivation in India using survey data: a state-level ruralurban analysis before and during the reform period. *Journal of Development Studies*, 45 (4): 441 - 470.
- [29] Lahiri, A., V. Hnatkovska, and S.B. Paul (2012), Castes and Labor Mobility, *American Economic Journal: Applied Economics* 4(2).
- [30] Machado, J. and J. Mata (2005), Counterfactual decomposition of changes in wage distributions using quantile regression, *Journal of Applied Econometrics*, 20:4, pp. 445 - 465.
- [31] Mallick, S. K. (2014), Disentangling the Poverty Effects of Sectoral Output, Prices, and Policies in India. *Review of Income and Wealth*, 60: 773 - 801.
- [32] Mitra, A. and D. Ray (2014), Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India, *Journal of Political Economy* 122, 719 -765.
- [33] Munshi, K. and M. Rosenzweig (2006), Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy, *American Economic Review*, 96(4), 1225 -1252.
- [34] Motiram, S., and Sarma, N. (2014), Polarization, inequality, and growth: The Indian experience. *Oxford Development Studies*, 42(3), 297 - 318.
- [35] Ñopo (2008), Matching as a Tool to Decompose Wage Gaps, *The Review of Economics and Statistics*, MIT Press, 90(2), 290 -299.

- [36] Rosenzweig, M. R. and Zhang, J. (2006), *Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's 'One Child' Policy*, IZA Discussion Papers 2082, Institute for the Study of Labor (IZA).
- [37] Srinivas, M. N. (1956), A note on sanskritization and westernization, *The Far Eastern Quarterly*, 15 (4), pp. 481- 496
- [38] Varshney A. (2001), Ethinc Conflict and Civil Society: India and Beyond, *World Politics*, 53 (3): 362 - 98.

## Appendix A

- Steps involved in constructing consumption distance

A detailed description of the consumption data is provided below along with the data normalization strategies in operationalising the idea of social distance in terms of consumption distance. We have taken individual consumption deviation from a reference point such as the mean of the consumption expenditure ( $\bar{C}$ ) from their own group, and then divided by their group standard deviation ( $\sigma$ ). For example, Hindus individual consumption distance is calculated in the following way:

*Step 1:* First consider State =  $i$  ( $i=1, 2, \dots, n$ ) in the period  $t = 1993$  ( $t=1993, 2004, \text{ and } 2011$ ).

*Step 2:* Identify Hindu individuals in State 1.

*Step 3:* Calculate individual deviation from the mean and divide by their standard deviation.

*Step 4:* Since both rural & urban have different consumption levels, the deviation for all urban Hindus in State 1 is calculated first and then deviation for rural Hindus is calculated next and then across two different social groups based on caste. A similar exercise is performed for non-Hindus in the same state.

*Step 5:* Repeat Steps 1 to step 4 for all other states and for the given time periods.

In addition to this and for robustness purpose, we have tested our baseline results (see Table 7) for the following model specifications:

1. Using all the above 5 steps, individual deviations are calculated from *mean* of the consumption level after considering the equal adult-level division of consumption in the same household (i.e., equivalence method). To check this assumption, we have taken the square root scale, as an equivalence scale, and then the regression is performed across the distribution (see Table 7).
2. Using all the above 5 steps, we also evaluate what happens if we keep median as a reference point instead of *mean* but keep the spread or dispersion bound same. In other words, what happens if one fixes the ( $\sigma$ ) as a bound? (see Table 7)
3. Using all the above 5 steps, we further evaluate individual consumption deviation being calculated from the median and then consider what happens if one changes the dispersion bound from ( $\sigma$ ) to *max-min*? (see Table 7)



- Instrumental variable quantile regression approach

Let us assume that we are interested in the scalar potential outcome  $Y_D$  (in our case, variability in consumption distance) conditional on a vector of variables,  $X = x$ . Also assume that we have a vector of binary indicators of treatment status, denoted as  $D$ . Also describe by  $q(\cdot)$  the conditional  $\tau$ -quantile function. Then conditional on  $X = x$ , the outcome  $Y_D$  can be presented by the following quantile function:

$$Y_{(D,x)} = q(D, x, U_D) \text{ where } U|x, z \sim u(0, 1) \quad (3)$$

In the above,  $U$  is a non-separable error and  $z$  is a vector of excluded instruments correlated with endogenous variable,  $D$ , but not correlated with the outcome variable ( $Y$ ). We can describe  $D$  as:

$$D = \delta(X, Z, V)$$

where  $z$  is a vector of excluded instruments and correlated with endogenous variable ( $d$ ) but not correlated with the outcome variable ( $Y$ ). Also note that  $V$  is a vector unobservable characteristics and can depend on  $u$ . A further note  $D = \delta(X, Z, V)$  is in order. Chernozhukov and Hansen (2005) notes that conditioning on the instrument as in  $D(\cdot)$  may appear to be a natural strategy for estimating quantile treatment effect. In general, variation of treatment effects across households creates many problems for 'conditioning on  $Z$ ' approaches. This problem is resolved in Chernozhukov and Hansen (2005) by imposing rank similarity conditions and quantile transforms. For the sake of convenience, we assume a linear model for the IVQR estimator  $q(D, x, U_D)$ , where we describe:

$$q(d, x, \theta) = d'\alpha(\theta) + x'\beta(\theta) \quad (4)$$

with  $q(\cdot)$  being assumed to be a strictly increasing in  $\theta$ . Our interest is in obtaining the treatment effects defined by  $q(d, x, \theta) - q(d^0, x, \theta)$  holding unobserved heterogeneity  $U_D$  fixed at  $\theta$ . The presence of instrumental variables in this scenario leads to a set of moment conditions (identification equation) described as:

$$P[Y \leq q(d, x, \theta)|z, x] = \theta \quad (5)$$

which allows us to estimate  $\alpha(\theta)$  and  $\beta(\theta)$ . Under the assumption of ranking invariance and linearity, Chernozhukov and Hansen (2005, 2008) derive the objective function as:

$$\operatorname{argmin} E(\rho_\theta[y - d'\alpha(\theta) - x'\beta(\theta) - z'\gamma(\theta)]) \quad (6)$$

- Counterfactual decomposition approach

To understand the concept of the counterfactual approach, consider two groups,  $A$  (Privileged/Hindu) and  $B$  (Marginalized/non-Hindus), those who are inherited with different types of characteristics and thus consume differently. Their consumption (outcome) variable  $C$  is linearly related to the covariates,  $X$ , and the error term  $\epsilon$  is conditionally independent of  $X$  and can be written in the following way:

$$C_A = X_A\beta_A + \epsilon_A \quad (7)$$

$$C_B = X_B\beta_B + \epsilon_B \quad (8)$$

where  $X$  is a set of standard control variables such as age, household size, education level, household type and state. The consumption gap between two groups can be calculated as

$$C_A - C_B = X_A\beta_A - X_B\beta_B \quad (9)$$

Since both groups are endowed with certain distinct characteristics and thus one can ask what would *group B* consume if they were endowed with the same characteristics like their counterpart *group A*. This can be calculated as

$$C^* = X_A\beta_B \quad (10)$$

In terms of counterfactual equation, the equation 9 thus can be written as

$$\begin{aligned} C_A - C_B &= X_A\beta_A - C^* + C^* - X_B\beta_B \\ &= X_A\beta_A - X_A\beta_B + X_A\beta_B - X_B\beta_B \\ &= X_A(\beta_A - \beta_B) + \beta_B(X_A - X_B) \end{aligned}$$

The quatile counterfactual decomposition assumption is that the regression quantiles are  $\beta^A(u)$  for group  $A$  and  $\beta^B(u)$  for group  $B$  for each  $u \in [0, 1]$ ; i.e.,  $Q_u(C_A|X_A = x_A) = x_A\beta^A(u)$   $u \in [0, 1]$  and similarly for group  $B$ . The distribution of  $C_A$  conditional on  $X_A = x_A$  is completely characterized by the collection of regression quantiles  $[\beta^A(u); u \in [0, 1]]$ , and likewise for the distribution of  $C_B$  conditional on  $X_B = x_B$ . The counterfactual method generates a sample from the unconditional distribution of  $C_A B$  as follows:

1. Sample  $u$  from a standard distribution .
2. Compute this  $\hat{\beta}_B(U)$ , i.e., estimate the  $u^{th}$  regression quantile of distribution  $C_B$  on  $x_B$ .
3. Sample  $x_A$  from the empirical distribution  $\hat{G}_{X_A}$ .
4. Compute  $\hat{C}_{AB} = x_A\hat{\beta}^B(u)$ . Repeat steps 1 to 3  $M$  times.

The underlying methodological approach is as follows:

$$\underbrace{Q_\tau(\hat{C}_A|\hat{\beta}_A - \hat{C}_B|\hat{\beta}_B)}_{gap} = \underbrace{[Q_\tau(\hat{C}_A|\hat{\beta}_A) - Q_\tau(\hat{C}_A|\hat{\beta}_B)]}_{coefficient} + \underbrace{[Q_\tau(\hat{C}_A|\hat{\beta}_B) - Q_\tau(\hat{C}_B|\hat{\beta}_B)]}_{characteristics} + \epsilon \quad (11)$$

- Matching approach  $\tilde{N}$ opo (2008)

The potential problem in any decomposition approach is misspecification due to differences in support of the empirical distribution of individual characteristics between two groups (in our case this is between two social identities: i.e., between privileged and marginalized groups/Hindus and non-Hindus). To overcome this problem, we use  $\tilde{N}$ opo (2008) matching algorithm to solve the problem of specification bias. The step-by-step procedure of this matching algorithm<sup>17</sup> is explained next:

- Step 1: Select one marginalized individual from the sample (without replacement).
- Step 2: Select all the privileged individuals that have the same characteristics as the marginalized individual previously selected.
- Step 3: With all the individuals selected in step 2, construct a synthetic individual whose consumption is the average of all of them and match the individual to the original marginalized group.
- Step 4: Put the observations of both individuals (the synthetic privileged and marginalized groups) in their respective new samples of matched individuals.
- Repeat steps 1 through 4 until we exhaust the original marginalized individuals sample.

The new data set contains the observations of matched privileged and marginalized individuals based on the same empirical distribution of probabilities for characteristics. The advantage of such matching approach is the modelling assumption that the same individuals with the same observable characteristics should consume the same, regardless of their social status.

---

<sup>17</sup>For details on econometric specification, see  $\tilde{N}$ opo (2008)

## Appendix B

Figure 8: Relationship between Caste (A & B) and Religion (C& D) with level of development

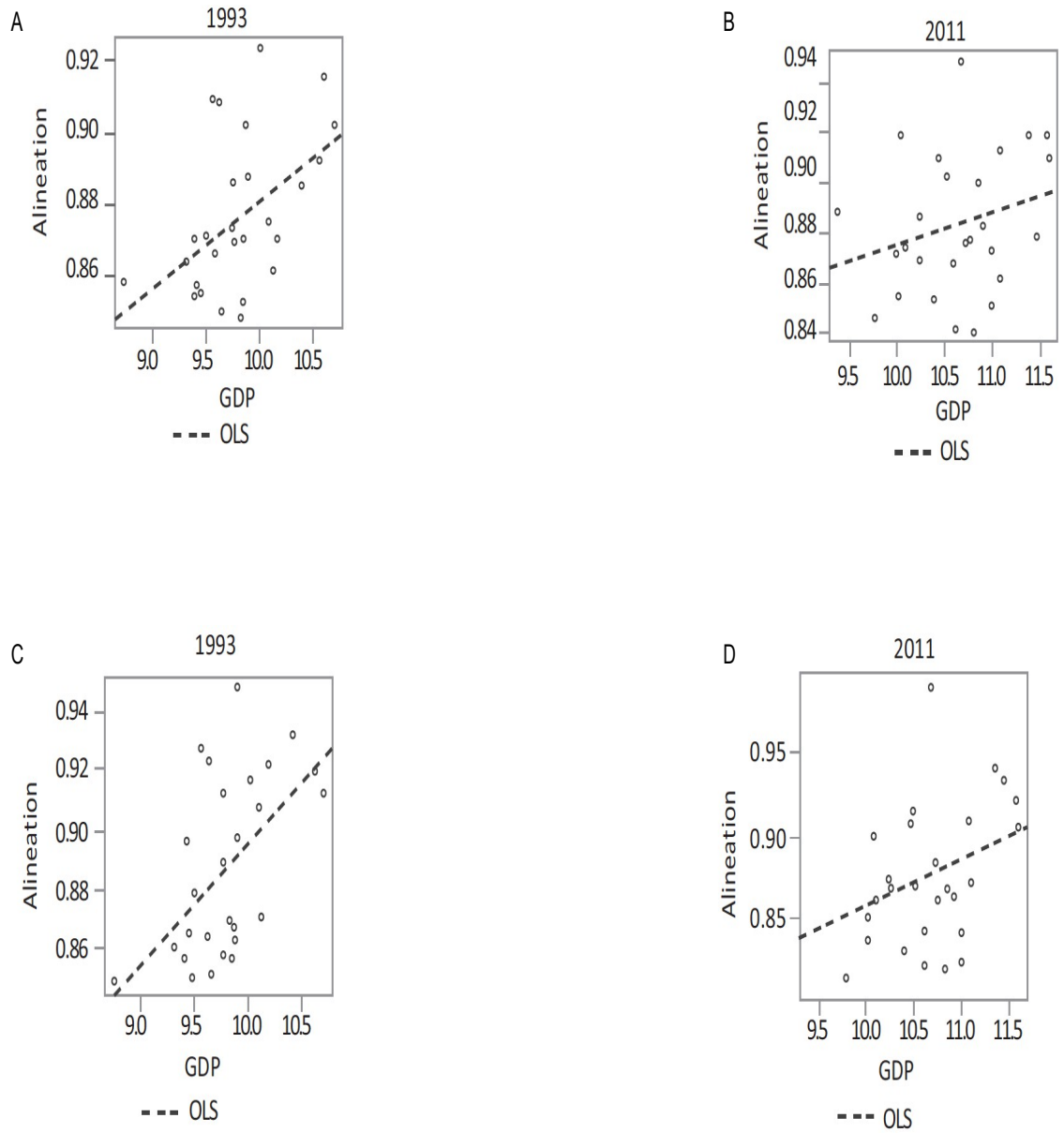


Figure 9: The effect of marginalized group relative to privileged group on consumption distance across the distribution

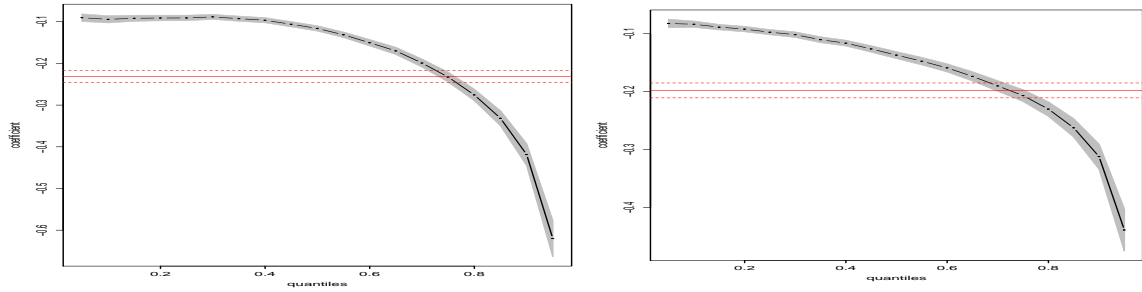


Figure 10: The effect of non-Hindus relative to Hindus on consumption distance across the distribution

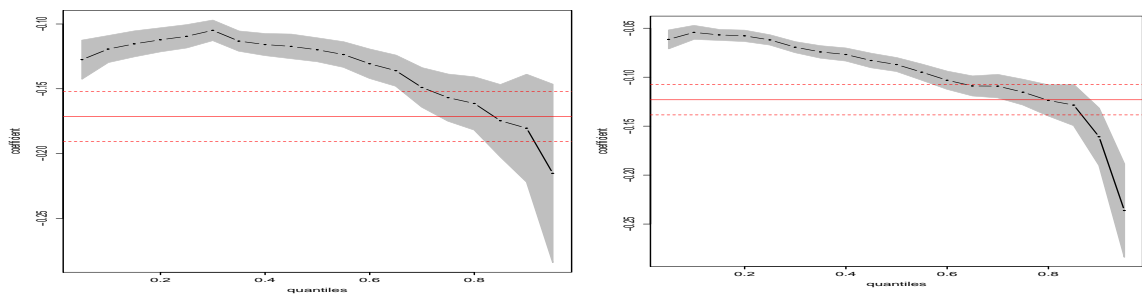
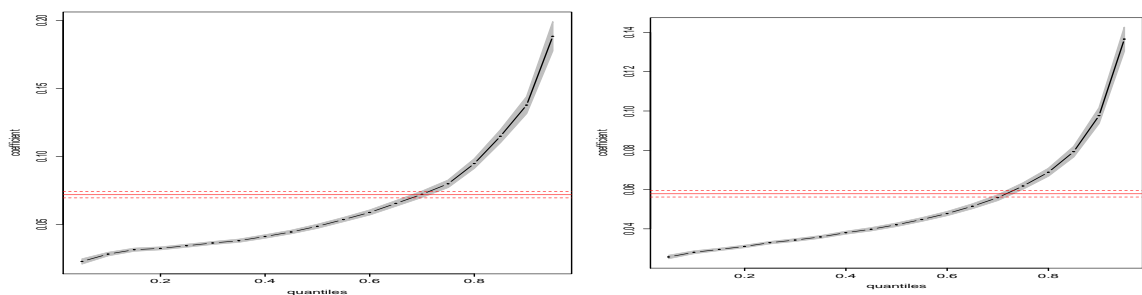


Figure 11: Effect of education on consumption level across the distribution in 1993



(i) The black line is estimated coefficient, whereas the shaded gray lines are confidence interval at 95% level and at different quantiles. (ii) Thick dashed black line is OLS and both dashed lines are confidence interval at the mean in each figure.

Table 6: Effect of social identities on the level of consumption across distribution and over time

	1993			2004			2011					
	10	50	90	10	50	90	10	50	90			
<i>Privileged groups (excluded)</i>												
Marginalized groups	-72.194*** (4.195)	-22.817*** (1.275)	-46.051*** (2.177)	-104.916*** (8.594)	-105.203*** (3.339)	-34.739*** (1.381)	-66.827*** (1.831)	-169.394*** (7.959)	-74.340*** (5.317)	-28.453*** (1.578)	-51.668*** (2.763)	-105.191*** (10.344)
<i>Hindus (excluded)</i>												
Non-hindus	-43.565*** (5.464)	-10.571*** (1.660)	-26.158*** (2.836)	-47.342*** (11.194)	-60.059*** (4.544)	-12.199*** (1.879)	-32.051*** (2.491)	-96.108*** (10.831)	-38.665*** (6.259)	-4.218** (1.857)	-16.033*** (3.252)	-50.914*** (12.177)
Education	47.285*** (0.513)	10.878*** (0.156)	26.304*** (0.266)	81.919*** (1.050)	29.771*** (0.580)	9.876*** (0.240)	23.024*** (0.318)	49.029*** (1.382)	53.055*** (0.630)	12.406*** (0.187)	29.420*** (0.327)	78.501*** (1.226)
Age	1.417*** (0.129)	0.548*** (0.039)	0.881*** (0.067)	1.177*** (0.264)	0.492*** (0.115)	0.015 (0.047)	0.532*** (0.063)	1.112*** (0.274)	1.899*** (0.168)	0.518*** (0.050)	1.174*** (0.087)	1.742*** (0.326)
<i>Self-employed (excluded)</i>												
Regular earnings	29.023** (11.894)	30.221*** (3.615)	30.347*** (6.175)	55.848** (24.369)	108.844*** (4.116)	16.350*** (1.702)	49.156*** (2.257)	217.585*** (9.812)	65.417*** (5.467)	11.830*** (1.622)	33.181*** (2.841)	111.548*** (10.635)
Casual labour	70.154*** (11.671)	13.378*** (3.547)	23.196*** (6.058)	70.858*** (23.910)	-5.546** (2.186)	-13.147*** (2.144)	-10.995*** (2.843)	10.462 (12.361)	-54.857*** (6.740)	4.940** (2.000)	-12.418*** (3.502)	-76.825*** (13.111)
Non-agriculture	5.554 (6.689)	9.755** (3.856)	5.357 (6.587)	10.043 (25.997)	-60.396*** (4.268)	-35.156*** (1.765)	-35.381*** (2.340)	-84.886*** (10.174)	-96.564*** (10.435)	-31.592*** (3.096)	-68.277*** (5.422)	-157.383*** (20.300)
Other types work	74.694*** (12.451)	35.227*** (3.784)	61.693*** (6.464)	141.848*** (25.509)	135.130*** (4.983)	3.522* (2.061)	39.930*** (2.732)	411.133*** (11.878)	211.298*** (8.811)	35.675*** (2.614)	138.905*** (4.578)	414.181*** (17.141)
Constant	155.228*** (13.139)	106.484*** (3.993)	181.775*** (6.820)	333.032*** (26.918)	395.874*** (7.517)	174.141*** (3.109)	296.713*** (4.121)	676.780*** (17.918)	207.824*** (10.731)	163.906*** (3.184)	249.507*** (5.576)	494.188*** (20.876)
N	69152	69152	69152	69152	109320	109320	109320	109320	89623	89623	89623	89623
R <sup>2</sup>	0.1542	0.0584	0.1049	0.2088	0.0729	0.025	0.0514	0.087	0.1172	0.0506	0.0959	0.1567
Slop difference test		492.85***				343.63***					774.72 ***	
10th-50th			374.45***				261.55***					514.2 ***
50th-90th												

Note: Authors' calculations based on NSS data. Dependent variable is level of consumption. Heteroscedasticity corrected robust standard errors are reported in brackets. \* Significant at  $p < 0.10$ , \*\* Significant at  $p < 0.05$ , \*\*\* Significant at  $p < 0.01$ .

Table 7: Effect of social identities on consumption distance across distribution with alternative reference points

	1993				2004				2011			
	OLS	10th	50th	90th	OLS	10th	50th	90th	OLS	10th	50th	90th
<i>Model 1</i>												
Marginalized groups	-0.071** (0.014)	-0.027*** (0.007)	-0.021* (0.009)	-0.042** (0.029)	-0.024** (0.014)	-0.014** (0.006)	-0.027* (0.009)	-0.069** (0.035)	-0.033** (0.015)	-0.022*** (0.007)	-0.033*** (0.009)	-0.048** (0.0291)
Non-hindus	-0.114*** (0.012)	-0.063*** (0.006)	-0.097*** (0.008)	-0.131*** (0.026)	-0.143*** (0.010)	-0.058*** (0.005)	-0.110*** (0.007)	-0.205*** (0.027)	-0.083*** (0.010)	-0.031*** (0.005)	-0.069*** (0.006)	-0.110*** (0.022)
Education	0.088*** (0.001)	0.028*** (0.001)	0.052*** (0.001)	0.148*** (0.003)	0.065*** (0.002)	0.026*** (0.001)	0.060*** (0.001)	0.099*** (0.005)	0.076*** (0.001)	0.024*** (0.000)	0.048*** (0.001)	0.115*** (0.002)
Marginalized x Edu	-0.033*** (0.003)	-0.009*** (0.001)	-0.026*** (0.002)	-0.068*** (0.006)	-0.027*** (0.003)	-0.009*** (0.001)	-0.027*** (0.002)	-0.037*** (0.006)	-0.016*** (0.002)	-0.006*** (0.001)	-0.012*** (0.001)	-0.029*** (0.005)
Non-hindus x Edu	-0.029*** (0.004)	-0.008*** (0.002)	-0.020*** (0.002)	-0.041*** (0.008)	-0.015*** (0.003)	-0.004*** (0.001)	-0.010*** (0.002)	-0.030*** (0.010)	-0.012*** (0.002)	-0.005*** (0.001)	-0.004*** (0.001)	-0.016*** (0.002)
<i>Model 2</i>												
Marginalized groups	-0.096*** (0.010)	-0.037*** (0.003)	-0.058*** (0.005)	-0.174*** (0.025)	-0.118*** (0.008)	-0.030*** (0.002)	-0.066*** (0.003)	-0.210*** (0.021)	-0.097*** (0.008)	-0.043*** (0.003)	-0.065*** (0.004)	-0.158*** (0.018)
Non-hindus	-0.138*** (0.012)	-0.055*** (0.004)	-0.093*** (0.007)	-0.201*** (0.033)	-0.141*** (0.010)	-0.042*** (0.003)	-0.089*** (0.005)	-0.224*** (0.027)	-0.112*** (0.009)	-0.023*** (0.003)	-0.066*** (0.005)	-0.211*** (0.022)
<i>Model 3</i>												
Marginalized groups	-	-	-	-	-	-	-	-	-0.110*** (0.009)	-0.053*** (0.004)	-0.089*** (0.005)	-0.152*** (0.017)
Non-hindus	-	-	-	-	-	-	-	-	-0.081*** (0.010)	-0.022*** (0.004)	-0.064*** (0.006)	-0.112*** (0.021)
<i>Model 4</i>												
Marginalized groups	-	-	-	-	-	-	-	-	-0.006*** (0.001)	-0.003*** (0.004)	-0.004*** (0.000)	-0.006*** (0.001)
Non-hindus	-	-	-	-	-	-	-	-	-0.004*** (0.001)	-0.002*** (0.000)	-0.003*** (0.001)	-0.006*** (0.001)

Model 1: Consumption distance is calculated from the mean and interacted with the education and identity. Model 2: Consumption expenditure is expressed on the basis of Square root scale and deviation is taken from the mean consumption. Model 3: Consumption distance is calculated from the median but bounded with the standard deviation. Model 4: Consumption distance is calculated from the median but bounded with the max-min Heteroscedasticity corrected robust standard errors are reported in brackets. \* Significant at  $p < 0.10$ , \*\* Significant at  $p < 0.05$ , \*\*\* Significant at  $p < 0.01$ .

Table 8: Counterfactual distribution results between privileged and marginalized groups (**Matched samples**)

Quantile	1993						2004						2011		
	Gap (1)	Endowment (2)	Coefficients (3)	Interaction (4)	Gap (5)	Endowment (6)	Coefficients (7)	Interaction (8)	Gap (9)	Endowment (10)	Coefficients (11)	Interaction (12)	Coefficients (11)	Interaction (12)	
10	0.201	0.062 (0.022)	0.129 (0.021)	0.010	0.180	0.081 (0.029)	0.089 (0.028)	0.009	0.292	0.061 (0.018)	0.230 (0.015)	0.001	0.230 (0.015)	0.001	
50	0.245	0.104 (0.021)	0.160 (0.020)	-0.019	0.255	0.133 (0.025)	0.145 (0.024)	-0.022	0.390	0.061 (0.022)	0.321 (0.021)	0.008	0.321 (0.021)	0.008	
90	0.389	0.155 (0.036)	0.219 (0.035)	0.015	0.337	0.151 (0.033)	0.174 (0.032)	0.012	0.444	0.093 (0.035)	0.364 (0.034)	-0.013	0.364 (0.034)	-0.013	
		40	56	4		45	52	3		21	82	-3	82	-3	

Note: (i) Bootstrap standard errors are in parenthesis and proportions explaining observed gap are mentioned under standard errors. ii) The terms Endowment and Coefficient reflect observed and unobserved inequality in the table.

Table 9: Comparison of unmatched and matched sample estimates in terms of percentage changes

Quantile	GAP						Percentage change					
	Unmatched			Matched			Unmatched			Matched		
	1993	2004	2011	1993	2004	2011	2004-1993	2011-1993	2011-2004	2004-1993	2011-1993	2011-2004
10	0.201	0.180	0.262	0.201	0.186	0.250	-0.074	0.241	0.339	-0.108	0.301	0.458
25	0.221	0.199	0.306	0.220	0.201	0.293	-0.085	0.332	0.456	-0.097	0.385	0.533
50	0.245	0.255	0.390	0.251	0.264	0.382	0.050	0.520	0.448	0.041	0.591	0.528
75	0.314	0.334	0.424	0.332	0.350	0.440	0.054	0.326	0.258	0.063	0.351	0.270
90	0.389	0.337	0.444	0.429	0.356	0.436	-0.170	0.015	0.223	-0.134	0.141	0.318

Note: Authors' calculations.