REVIEW OF DEVELOPMENT ECONOMICS
Manuscript No: #2382, Acceptance Date: March 14 2004

Caste, Inequality, and Poverty in India*

Vani K. Borooah

RRH: Inequality and Poverty in India
LLH: Vani K. Borooah

Abstract

This paper analyses inequality and poverty in India within the context of caste-based discrimination. It does so by decomposing the difference between (caste) Hindu and Scheduled Caste (SC) and Scheduled Tribe (ST) households in: their average household incomes; their probabilities of being in different income percentiles; their probabilities of being at different levels of poverty into: a 'discrimination effect', which stems from the fact that a household's income level, into which its (income-generating) profile translates, depends on whether it is SC/ST; an 'attributes (or residual) effect' which stems from the fact that there are systematic differences between SC/ST and Hindu households in their (income-generating) profiles. The results, based on unit record data for 28,922 households, showed that at least one-third of the average income/probability differences between Hindu and SC/ST households was due to the 'unequal treatment' of the latter.

JEL Classification Numbers: C25 D31 I32 J71

Abbreviations: SC, ST
Number of Figures: 0 Number of Tables: 6
Date: 24 March 2004

Address of contact author: V.K. Borooah, School of Economics and Politics, University of Ulster, Newtownabbey BT37 0QB, Northern Ireland, UK. Tel: +44-28-90366934, Fax: +44-90366847, Email: VK.Borooah@ulst.ac.uk.

---

* Borooah: School of Economics and Politics, University of Ulster, Newtownabbey BT37 0QB, Northern Ireland, UK. Tel: +44-28-90366934, Fax: +44-28-90366847, E-mail: VK.Borooah@ulst.ac.uk. This study is based on unit record data from the 1993-94 Human Development Survey carried out by the National Council of Applied Economic Research (NCAER), New Delhi. I am grateful to the NCAER for providing me with the data. Thanks are also due to an Editor and two referees whose comments have greatly improved the paper. This paper was substantially written while I was a Visiting Fellow at the Institute of Social Sciences (ISS) at The Hague (The Netherlands) and I am grateful to the ISS and to the Royal Economic Society for supporting this research. However, needless to say, I alone am responsible for the results reported in the paper, for their interpretation and, indeed, for any of its deficiencies.

1. Introduction

In response to the burden of social stigma and economic backwardness borne by persons belonging to India's "untouchable castes" (that is, those with whom physical contact, most usually taken to be the acceptance of food or water, is ritually polluting or unclean) the Constitution of India allows for special provisions for members of these castes[1]. Articles 341 and 342 include a list of castes and tribes entitled to such benefits and all those groups included in this list – and subsequent modifications to this list – are referred to as, respectively, "Scheduled Castes" and "Scheduled Tribes". For all practical purposes the terms "Scheduled Castes" (SC) and "Scheduled Tribes" (ST) are synonymous with the "untouchable" castes.

Against this background of the relative economic and social backwardness of the 16% of India's population who belong to the SC, and the 8% who belong to the ST, Deshpande (2001) has observed: "poverty and inequality has been extensively researched but the focus of economic research has been on identifying and defining the contours of poverty and redistribution policies that target the poor as a group….this has excluded the study of other ingredients of stratification, most notably caste, and precludes any inferences about inter-group disparity based on caste" (p. 130). Such studies of caste and economic inequality that exist have been few and far between and Desphande (2000b) provides a review of such studies[2].

Most economic studies of caste in India focus – for reasons cited above – on the SC/ST versus non-SC/ST distinction. Given this common cleavage, broadly speaking, they fall into two camps: studies of caste-based discrimination; and studies of economic disparities between castes. The first group typically estimates earnings functions for SC and non-SC workers with a view to seeing how much of the earnings difference between the two groups can be explained by differences in worker qualities (for example, Banerjee and Knight, 1985). The second group is concerned with measuring the degree of economic disparity between, on the one hand, SC and ST persons/households and, on the other hand, non-SC/ST persons/households (Gang *et. al.* 2002; Deshpande, 2000a; Sagar and Pan 1994).

The novelty of this paper is that it attempts to combine the two strands of research by analysing inequality and poverty in India within the context of caste-based discrimination. It does so by decomposing the difference between (caste) Hindu and SC/ST households in: (i) their average household incomes; (ii) their probabilities of being in different income percentiles; (iii) their probabilities of being at different levels of poverty into: (a) a 'discrimination effect', which stems from the fact that a household's income level, into which its (income-generating) profile translates, depends on whether it is SC/ST; (b) an 'attributes (or residual) effect' which stems from the fact that there are systematic differences between SC/ST and Hindu households in their (income-generating) profiles[3].

This is accomplished by, first, estimating an income-generating function and using the Blinder-Oaxaca decomposition methodology (Blinder, 1973; Oaxaca 1973; Oaxaca and Ransom, 1994) to estimate the size of the 'discrimination effect' in determining average inter-group income differences Then, a multinomial model for the likelihood of being in different income quintiles (Diamond *et. al.*, 1990) is estimated and the Blinder-Oaxaca decomposition methodology is extended to multinomial probability models to estimate the size of the 'discrimination effect' in determining average inter-group differences in the probability of being in different income quintiles. Thirdly, following a suggestion by Ravallion (1996), a multinomial model for the likelihood of being at different levels of poverty is estimated and the decomposition methodology is used to estimate the size of the 'discrimination effect' in determining average inter-group differences in the probability of being at different poverty levels.

The results reported in this paper are based on unit record data for 28,922 households – which were spread over 1,759 villages, in 195 districts, in 16 states of India. The Survey, from which these data were obtained, was commissioned by the Indian Planning Commission, funded by a consortium of United Nation agencies; it was carried out by the National Council of Applied Economic Research (NCAER) over January-June 1994. Details of this Survey, hereafter referred to as the NCAER Survey, are contained in Shariff (1999).

2. The Decomposition of Household Income: A Log-Linear Model

This section "models" income inequality in the context of estimating the relative strengths of the different factors affecting the level of household income in India and examining how these vary across SC, ST and Hindu. It was hypothesised that a household's income would *inter alia* depend upon the following factors:

(i) Whether the household owned land.

(ii) The amount of land (if any) owned by the household, in acres.

(iii) The number of adult workers in the household.

(iv) The amount of (non-land) productive assets owned by the household: this amount was represented by a productive assets index[4].

(v) The educational level of the household head: 'low', if the head was illiterate; 'moderate', if the head had received schooling of primary level or below; 'high', if the head had received above primary level schooling.

(vi) The social group to which the household belonged. The study was confined to those households who were: SC; ST; or 'caste' Hindus (that is, Hindus who were neither SC nor ST)[5]

(vii) The region in which the household lived: central; south; west; east; north[6].

The structure of the econometric equation, with the natural logarithm of household income as the dependent variable, is set out below for household *i:*

$$\begin{aligned}\log(hinc_i) &= \alpha_0 + \alpha_1 \times central_i + \alpha_2 \times south_i + \alpha_3 \times west_i + \alpha_4 \times east_i \\ &+ \alpha_5 \times st_i + \alpha_6 \times sc_i + \alpha_7 \times X_i \\ &+ \beta_1 \times (st_i \times central_i) + \beta_2 \times (st_i \times south_i) + \beta_3 \times (st_i \times west_i) + \beta_4 \times (st_i \times east_i) \\ &+ \beta_5 \times (sc_i \times central_i) + \beta_6 \times (sc_i \times south_i) + \beta_7 \times (sc_i \times west_i) + \beta_8 \times (sc_i \times east_i) \\ &+ \gamma_1 \times (central_i \times X_i) + \gamma_2 \times (south_i \times X_i) + \gamma_3 \times (west_i \times X_i) + \gamma_4 \times (east_i \times X_i) \\ &+ \delta_1 \times (st_i \times X_i) + \delta_2 \times (sc_i \times X_i) + \varepsilon_i\end{aligned} \quad (1)$$

The variable X_i in equation (1) is a "representative covariate" from those defined in (i)-(vii) above. The α coefficients in equation (1) measure the separate effects of the regional variables ($\alpha_1 - \alpha_4$), the social group variables (α_5, α_6) and X_i (α_7) on household income[7]; the β coefficients allow the effect of a household belonging to a particular social group to vary by the region in which it lived; the γ coefficients allow the effect of X_i, to vary by the region in which the household lived; and the δ coefficients allow the effect of the value of X_i on the income of household *i* to vary by the household's social group; ε_i is an error term.

In general, the $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4$, α_7 and the $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ coefficients of equation (1) may be thought of as the "Hindu" coefficients; $\alpha_5, \beta_1, \beta_2, \beta_3, \beta_4$, and δ_1 may be thought of as the *additional* effects from being a ST household and $\alpha_6, \beta_5, \beta_6, \beta_7, \beta_8$, and δ_2 may be thought of as the *additional* effects from being a SC household. So, for example, the mean (log) income of a Hindu household living in the North, whose head was illiterate, which did not own any land or any (other) productive asset[8] was predicted, from equation (1), to be: α_0. Call this the "baseline income”; the *ceteris paribus* clauses in the subsequent paragraph refer to this “baseline scenario”[9].

The study reported in this paper - with certain differences in econometric methodology - is similar to van de Walle and Gunewardena's (2001) study of income inequality in Vietnam, with reference to its 'majority' (*Kinh*) and minority (non-*Kinh*) ethnic populations. Given the small number of locations, and the large number of households in our study, compared to that of van de Walle and Gunewardena (2001), the effects of region and ethnicity were accommodated, as described above, by employing interaction terms. This allowed not only separate intercept terms for each region and each ethnicity, but also separate slope coefficients. By contrast, estimating the income generating equation separately for the different ethnic groups, as did van de Walle and Gunewardena (2001), meant that they were confined to testing, via a Chow test, for the inter-ethnic equality of the entire coefficient vector without being able to focus on the subset of the covariates whose coefficients were significantly different between the groups.

The distinction between Hindu, SC and ST coefficients, contained in equation (1), may be made explicit by setting out the income determining equation (1) separately for the three social groups as:

$$\log(hinc_i^k) = \boldsymbol{\theta}^{\mathbf{k}\prime}\mathbf{X}_{\mathbf{i}}^{\mathbf{k}} \tag{2}$$

where, in equation (2), $\boldsymbol{\theta}^{\mathbf{k}}$ is the coefficient vector for social group k and $\mathbf{X}_{\mathbf{i}}^{\mathbf{k}}$ is the vector of values of the determining variables, pertaining to household *i* in social group *k*: *k = Hindu, SC, ST*. Following the methodology of Blinder (1973) and Oaxaca (1973), the difference in mean household income between any two social groups (say, Hindu and SC) may be decomposed as:

$$\overline{\log(hinc^{H})} - \overline{\log(hinc^{SC})} = \boldsymbol{\theta}^{\mathbf{H}\prime}(\overline{\mathbf{X}}^{\mathbf{H}} - \overline{\mathbf{X}}^{\mathbf{SC}}) + (\boldsymbol{\theta}^{\mathbf{H}} - \boldsymbol{\theta}^{\mathbf{SC}})'\overline{\mathbf{X}}^{\mathbf{SC}} \quad (3)$$

or as:

$$\overline{\log(hinc^{H})} - \overline{\log(hinc^{SC})} = \boldsymbol{\theta}^{\mathbf{SC}\prime}(\overline{\mathbf{X}}^{\mathbf{H}} - \overline{\mathbf{X}}^{\mathbf{SC}}) + (\boldsymbol{\theta}^{\mathbf{H}} - \boldsymbol{\theta}^{\mathbf{SC}})'\overline{\mathbf{X}}^{\mathbf{H}} \quad (4)$$

In equation (3), the difference in mean incomes between Hindu and SC households is decomposed by asking what the mean income of SC households would have been had they "been treated" as Hindu households; in equation (4), this difference is decomposed by asking what the mean income of Hindu households would have been had they "been treated" as SC households.

2.1 Estimation Results from the Log-linear Model

The mean income of Hindu households, at Rs. 32,016 was 59% higher than that of ST households and 68% higher than that of SC households. Compared to SC and ST households, a greater proportion of Hindu households owned land, and the mean size of land owned was considerably greater for land-owning Hindu households than for land-owning SC or ST households. The value of the productive assets index was also considerably greater for Hindu households and the level of education of the household head was higher than that of the heads of SC and ST households[10].

Table 1 shows the results from estimating equation (1), above, with zero restrictions imposed on all those coefficients in equation (1) whose associated t-values were less than unity[11]. The mean baseline household income (defined above) was estimated as Rupees (Rs.) 12,972 per year, which is the exponential of the intercept estimate. *Ceteris paribus* being a SC or a ST household (as opposed to being Hindu) also reduced mean household income by Rs. 2,531 for SC households and by Rs. 2,074 for ST households. With the other components of the baseline scenario unchanged, the reduction in SC and ST income, from the baseline level, depended on the region in which the household was located: for SC households, living in the West reduced their incomes by Rs. 938, and living in the East raised their incomes by Rs. 1,179, compared to the other regions; for ST households, living in the West reduced their incomes by Rs. 777, and living in the East raised their incomes by Rs. 1,393, compared to the other regions.

The additional income from being a land owner was lower for SC and ST households than for Hindu households. For SC households in the Central region, for example, the additional income from being a land owner was Rs. 2,490 compared to Rs. 3,408 from Hindus. Another difference between SC and Hindu households was that the marginal contribution to mean household income of an additional worker was lower in SC households than in Hindu households: within the context of the baseline scenario, an additional worker would contribute Rs. 1,277 to a SC household but Rs. 1,586 to a Hindu household.

Table 2 shows that the (log) difference between the mean incomes of Hindu and SC households was 0.411. When SC households were treated as Hindus, 36% of this difference (0.150 out of 0.411) could be attributed to 'unequal treatment' (that is, the income-determining coefficients were different for Hindu and SC households so that the values of the income-generating attributes of SC households yielded a higher mean income when they were evaluated using Hindu, as opposed to SC, coefficients). When Hindu households were treated as SC households, 32% of this difference (0.133 out of 0.411) could be attributed to 'unequal treatment', as defined above.

In terms of the difference in mean incomes between Hindu and ST households, 46% and 39% of the difference could be attributed to “unequal treatment” when, respectively, ST households were treated as Hindus and Hindu households were treated as ST - with the remainder being due to “attribute differences” between Hindu and ST households.

3. A Multinomial Probability Model of Income Distribution

Diamond *et. al.* (1990) proposed a multinomial probability model of income distribution which predicted the conditional (upon characteristics) probability of a household/individual being in a particular income percentile. They argued that this method provided a better fit to the overall distribution than the traditional log-linear method. This study estimated such a model using the household income data from the NCAER Survey.

Suppose that household incomes have been stratified into percentiles. The basic question that a multinomial probability model of income distribution seeks to answer is: what is the probability that a household, with a particular set of characteristics, will be found in a specific income percentile? While there are many distributions from which these probabilities may be derived, perhaps the only tractable one is that of multinomial logit (Diamond *et. al.*, 1990).

The Oaxaca (1973) and Blinder (1973) method of decomposing group differences in means into an “explained” and a “residual” component has been extended to explaining group differences in probabilities, derived from models of discrete choice with *binary* outcomes, by *inter alia* Nielsen (1999). This section extends this methodology to models of discrete choice with *multiple* outcomes.

There are N households (indexed, $i=1...N$) which can be placed in G mutually exclusive and collectively exhaustive groups $g=1..G,$ each group containing N_g households. Then, under a multinomial logit model, the likelihood of a household, from community g, being in income quintile j is:

$$\Pr(Y_i = j) = F(\mathbf{X_i^g}\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g}) \qquad (5)$$

where: $\mathbf{X_i^g} = \left\{X_{ik}^g, k = 1...K\right\}$ represents the vector of observations, for household i of group g, on K variables which determine the likelihood it being in a particular quintile, and $\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g} = \left\{\beta_{jk}^g, k = 1...k\right\}$ is the associated vector of coefficient estimates for that group and for that quintile outcome.

The average probability of a household from group g being in quintile j is:

$$\overline{P}_j^g = P(\mathbf{X_i^g}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g}) = N_k^{-1}\sum_{i=1}^{N_k} F(\mathbf{X_i^g}\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g}) \qquad (6)$$

Define the "reference" coefficient vector, $\hat{\boldsymbol{\beta}}_\mathbf{j}^* = \sum_{g=1}^{G} \alpha^g \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g}$ for weights: $0 \le \alpha^g \le 1,\ \sum_{g=1}^{G} \alpha^g = 1$ and define the 'discrimination effect' for group g in quintile j as:

$$D_j^g = N_g^{-1}[\sum_{i=1}^{N_g} F(\mathbf{X_i^g}\hat{\boldsymbol{\beta}}_\mathbf{j}^*) - N_g^{-1}\sum_{i=1}^{N_k} F(\mathbf{X_i^g}\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g})] = P(\mathbf{X_i^g}, \hat{\boldsymbol{\beta}}_\mathbf{j}^*) - \overline{P}(\mathbf{X_i^g}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g}) \qquad (7)$$

The 'discrimination effect' then expresses the change in the average probability of households from community g being in quintile j when the attributes of the households in the community (the $\mathbf{X_i^g}$) are evaluated using the 'reference coefficients', $\hat{\boldsymbol{\beta}}_\mathbf{j}^*$ rather than the community-specific coefficients, $\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{g}$. For some groups, this change will be positive; for others, it will be negative.

Now for any two communities, say Hindu ($g=H$) and Scheduled Caste ($g=SC$):

$$D_j^{SC} - D_j^H = [P^*(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^*) - P^M(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC})] - [P^*(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^*) - \overline{P}^H(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H})]$$

which implies:

$$\bar{P}_j^H - \bar{P}_j^{SC} = P(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H}) - P(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC}) = [D_j^{SC} - D_j^H] + [P(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^*) - \bar{P}(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^*)] \qquad (8)$$

where the term on the left hand side of equation (8) is the difference in mean probabilities of Hindu and SC households being in quintile j.

If Hindus and SC have the same estimated coefficient vector, so that $\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H} = \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC} = \hat{\boldsymbol{\beta}}_\mathbf{j}$, then:

$$\bar{P}_j^H - \bar{P}_j^{SC} = P(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}) - P(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}) \qquad (9)$$

and the difference between Hindus and SC households in their observed proportions in income quintile j would be entirely due to differences in attributes.

However, suppose that $\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H} \neq \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC}$ and that $\alpha_j^H = 1$, so that the 'reference' coefficient vector is $\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H}$. Then, by equation (8), $D_j^H = 0$ so that:

$$\bar{P}_j^H - \bar{P}_j^{SC} = [P(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H}) - P(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC})] + [P(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H}) - P(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H})] \qquad (10)$$

On the other hand, if $\alpha_j^{SC} = 1$ (so that the 'reference' coefficient vector is $\hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC}$) then $D_j^{SC} = 0$ and equation (8) becomes:

$$\bar{P}_j^H - \bar{P}_j^{SC} = [P(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{H}) - P(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC})] + [P(\mathbf{X_i^H}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC}) - P(\mathbf{X_i^{SC}}, \hat{\boldsymbol{\beta}}_\mathbf{j}^\mathbf{SC})] \qquad (11)$$

In equation (10), the difference in sample means is decomposed by asking what the average probability of SC households being in quintile j would have been, *had they been treated as Hindu households*; in equation (11), it is decomposed by asking what the average probability of Hindu households being in quintile j would have been, *had they been treated as SC households*.

3.1 Empirical Results From the Multinomial Logit Model

The multinomial logit model was estimated over 28,922 households (distinguished between Hindu, SC and ST households) such that, for each household, the dependent variable of the model assumed exactly one of the values: 1,2,3,4 and 5, depending upon the quintile in which the household, on the basis of its income, was placed. The specification of the equation was as shown in equation (1), that is, with *all* the interaction effects included[12].

The multinomial logit estimates are not shown, both for reasons of economy of space and because the coefficients are difficult to interpret in terms of the underlying probabilities of being in the different quintiles[13]. Instead, using the estimates as in equations (10) and (11) above, Tables 3 and 4 decompose, respectively, the difference in average probabilities, between Hindu and SC, and between Hindu and ST, households of being in the top and bottom quintiles into a 'discrimination effect' and an 'attributes effect'.

Table 3 shows that 26.3% of Hindu households, but only 10.7% of SC households and 12.3% of ST households had incomes in the top quintile; conversely, 26.6% of SC households and 26.4% of ST households, but only 15.7% of Hindu households, had incomes in the bottom quintile. If SC and ST households were treated as Hindus (in the sense that their income-generating attributes were evaluated at Hindu coefficients) their presence in the top quintile would rise to, respectively, 15.5% and 17.2% and their presence in the bottom quintile would fall to, respectively, 22.7% and 20.7%.

Thus, of the total difference between Hindu and SC households, and between Hindu and ST households, in their probabilities of being in the *top* income quintile, 29% (for SC households) and 35% (for ST households) could be ascribed to a ‘discrimination factor’ working against SC and ST households. Conversely, in Table 4, of the total difference between Hindu and SC households, and between Hindu and ST households, in their probabilities of being in the *bottom* income quintile, 36% (for SC households) and 53% (for ST households) could be ascribed to a ‘discrimination factor’ working against SC and ST households.

If the results are viewed from the perspective of Hindu households being treated as SC households, then the presence of Hindu households in the top quintile fell from 26.3% to 20.6% and rose from 15.7% to 18.1% in the lowest quintile; when they were treated as ST households, the presence of Hindu households in the top quintile fell from 26.3% to 22.2% and rose from 15.7% to 20.8% in the lowest quintile. Consequently, from this perspective, 37% and 29%, respectively, of the total difference between Hindu and SC households, and between Hindu and ST households, in their probabilities of being in the *top* income quintile could be ascribed to a ‘discrimination factor’ working against SC and ST households; of the total difference between Hindu and SC households, and between Hindu and ST households, in their probabilities of being in the *bottom* income quintile, 22% (for SC households) and 47% (for ST households) could be ascribed to a ‘discrimination factor’ working against SC and ST households.

3.2 Multinomial Probability Model of Poverty

Ravallion (1996), reviewing current practice in ‘poverty modelling’, was critical of the method of defining a *binary* variable (poor, non-poor), on the basis of a poverty line, and then using logit or probit methods to ‘explain’ the probability of being poor in terms of observed household characteristics. He argued that unlike the usual binary response model, in which the latent variable was not observed, data on household incomes clearly exist: so, estimating a binary (poor/non-poor) model in the presence of income data for individual households was, in effect, to ‘throw away’ data by pretending they did not exist. Instead, Ravallion’s (1996) proposal was to follow the methodology of Diamond, *et. al.* (1990) by estimating a *multinomial* model of poverty based on *different* poverty lines. This suggestion is developed in this subsection.

Let $z_1 > z_2 ... > z_J$ represent J ‘poverty lines’, defined in terms of household income, such that household i is ‘poor’ at ‘level j’ if: $z_{j-1} < hinc_i \leq z_j$. For example, if *J=3*, there are three levels of poverty, the level of poverty being indicated by the value assumed by a variable Y_i : a household is ‘not poor’ if $hinc_i > z_1$: $Y_i = 0$; a household is ‘mildly poor’ if $z_2 < hinc_i \leq z_1$: $Y_i = 1$; a household is ‘moderately poor’ if $z_3 < hinc_i \leq z_2$: $Y_i = 2$; a household is ‘very poor’ if $hinc_i \leq z_3$: $Y_i = 3$

Under a multinomial logit formulation - in which the log-odds ratio of being poor at level *j* ($Y_i = j,\ j = 1,2,3$), relative to being non-poor ($Y_i = 0$), can be written as a linear function of $\mathbf{X_i} = \{X_{ik},\ k = 1...K\}$, the vector of values, for the household, of K ‘poverty conditioning’ variables - the difference between Hindu and SC households (and between Hindu and ST households) in their, respective, average probabilities of being poor, at different levels of poverty, can be decomposed, using the methodology set out earlier, as the sum of two parts:

(i) a part which stems from the fact that coefficient differences between Hindu and SC (or ST) households mean that a given household profile translates into different probabilities of being poor for households from the two communities

(ii) a part which arses from the fact that the profiles of Hindu households differ from those of SC (or ST) households

3.3 Empirical Results From the Multinomial Poverty Model

The poverty lines, which were defined in terms of the median income (μ) of the 28,922 households studied, were: $z_1 = 0.75\mu$; $z_2 = 0.5\mu$; $z_3 = 0.25\mu$. Accordingly, a household was: 'not poor', if its income exceeded three-fourths median income ($Y_i = 0$); 'mildly poor', if its income lay between three-fourths and half of median income ($Y_i = 1$); 'moderately poor', if its income lay between half and one-fourth of median income ($Y_i = 2$); 'very poor', it its income was equal to, or less than, one-fourth of median income ($Y_i = 3$).

Table 5 shows that, on the basis of μ=Rs.17,202, nearly three-fourths of Hindu households, but just over half of SC and ST households were, on the above definition, 'not poor'; less than 15% of Hindu households, but over 20% of SC and ST households, were 'mildly poor'; one in ten Hindu households, but nearly one in five SC and ST households, was 'moderately poor'; lastly, 4% of Hindu households, but 6% of SC and ST households, were 'very poor'. In summary, the incidence of poverty, at every level of poverty, was greater for SC and ST households than for Hindu households.

The multinomial estimates, for reasons set out earlier, are not shown. However, one may use the estimated equations to test the stability of the coefficients across the different levels of poverty - thus relaxing the first-order dominance assumption implicit in attaching a single parameter to the "poverty-conditioning" variables – by testing the null hypotheses: $\boldsymbol{\beta}_\mathbf{r} = \boldsymbol{\beta}_\mathbf{s}$, $r, s = 0,1,2,3$ $r \neq s$. Then, as Ravallion (1996) suggests, "one may want to specify a set of regression functions, the parameters of which vary according to the segment of the distribution one is considering" (p. 1335). The estimated equations did not accept the above null hypotheses for all *r,s* thus justifying the analysis of poverty at different levels of poverty rather than on the basis of a simple poor/non-poor distinction.

One can, as described earlier, decompose the difference in average probabilities, between Hindu and SC and ST households, of being at different poverty levels, into a 'discrimination effect' and an 'attributes effect' . These results showed that if SC and ST households had been treated as Hindus (in the sense that their household profiles were evaluated at Hindu coefficients): the proportion of 'non-poor' SC and ST households would rise to, respectively, 61% and 64%; the proportion of 'mildly poor' SC and ST households would fall to, respectively, 18% and 17%; the proportion of 'moderately poor' SC and ST households would fall to, respectively, 16% and 15% ; and the proportion of 'very poor' SC and ST households would fall to, respectively, 6% and 5%[14].

Table 6 collates the results to show the strength of the 'discrimination factor', in terms of the probabilities of being at different poverty levels, operating against SC and ST households. This discrimination factor (expressed in percentage terms) is defined as the proportion of the total difference, between Hindu and SC/ST households, in their average probabilities of being at a poverty level, which can be explained by coefficient differences between the communities. Thus the first line of Table 6 suggests that, of the total difference between Hindu and SC households, and between Hindu and ST households, in their average probabilities of being 'non-poor', 39% (for SC households) and 58% (ST households) could be ascribed to a 'discrimination factor', when Hindu households were treated as SC/ST households. When, however, the profiles of SC/ST households were evaluated using Hindu coefficients, the corresponding figures were 27% (SC) and 46% (ST).

Table 6 indicates that the strength of the discrimination factor, in shaping differences between Hindu and SC, and Hindu and ST, households in their average probabilities of being poor/non-poor, was considerably stronger for ST than for SC households. Moreover, for both SC and ST households, the strength of the discrimination factor was higher when SC/ST households were 'treated as' Hindus than when Hindu households were treated as SC or ST households.

4. The Decomposition of Income Inequality and Poverty

The previous sections attempted to 'explain' observed household income in terms of household characteristics. This section takes the observed household incomes and attempts to measure: (i) the contribution of the different social groups to the inequality in the distribution of these incomes; (ii) the differential risk of the different groups of being poor.

Suppose that the sample of $N=28,922$ households is divided into K mutually exclusive and collectively exhaustive groups. Then overall income inequality may be decomposed as the sum of 'within-group' and 'between-group' income inequality using Theil's (1967) Mean Logarithmic Deviation (MLD) index[15].

Similarly, suppose that, given a poverty line, there are M poor households in the total of N households and that, in group k, M_k , of the N_k , households in the group, are poor ($k=1...K$). Then, using the poverty index due to Foster *et. al.* (1984) the value of the overall poverty index can be expressed as the weighted average of the subgroup values.

The poverty risk of group k, ρ_k , is the ratio of a group's contribution to poverty to its contribution: $\rho_k > 1$ ($\rho_k < 1$) means that it contributes more (less) to poverty than its population share warrants. If the norm for poverty risk is taken to be unity, then, say, $\rho_k = 1.3$ means that the poverty risk for members of group k is 30%

above the norm; similarly, $\rho_k = 0.82$ means that the poverty risk for members of group *k* is 18% below the norm.

4.1 Empirical Results

When the 28,922 households were subdivided according to whether they were Hindu, SC or ST, 7% of all-India inequality could be explained by inequality between the three groups; in terms of the regions, this percentage contribution varied from a high of 8% in the Central region, the North and the West to a low of 1% in the East. When the 28,922 households were subdivided by social group *and* by whether or not they owned land (a total of six subgroups), 15% of all-India inequality could be explained by inequality between the three groups; in terms of the regions, this percentage contribution varied from a high of 20% in the South and the West to a low of 3% in the East[16]. To put the results in perspective, Cowell and Jenkins (1995) found, for income inequality in the USA, that four factors relating to the household head – age, sex, race and employment status – collectively explained less than one-fourth of overall US inequality.

There was, on an all-India basis, greater inequality among Hindus than within the ST or the SC groups: the values of MLD index were 0.35, 0.31 and 0.42 for, respectively, ST, SC and Hindu households. These results echo, on an all-India basis, Deshpande's (2000a) conclusions for Kerala : "the existence of an elite group, or upper class, is much more pronounced in the 'Others' category than it is in the SC or ST category" (p. 325).

Saggar and Pan (1994) found, for the states of Bihar and West Bengal, that inequality was greatest among, respectively, ST and SC households. Our results, based on household income data and on a different data set confirm their finding: for Bihar, on our calculations, the values of the MLD index were 0.34 (ST), 0.24 (SC) and 0.29 (Hindu) while, for West Bengal, they were 0.26 (ST), 0.27 (SC) and 0.28 (Hindu).

Using half of the all-India median income as the poverty line, the Head Count Ratio showed that, on an all-India basis, one-fourth of SC and ST, but only 15% of Hindu, households had incomes below the poverty line[17]. The poorest regions were the East and the Central regions with, respectively, 27% and 20% of households who were poor; the least poor regions were the North and the West with, respectively, 11% and 17% of households who were poor.

SC and ST households had a poverty risk that was, on an all-India basis, 30% above the norm: their proportionate contribution to poverty in India was nearly one-third greater than their proportionate presence in the sample of 28,922 households; conversely, Hindu households, on an all-India basis, had a poverty risk that was 23% below the norm: their proportionate contribution to poverty in India was nearly one-fourth smaller than their proportionate presence in the sample of 28,922 households.

Sagar and Pan (1994) calculated poverty rates, using mean per-capita household expenditure, for SC, ST and nonSC/ST households for the four states: Assam, Bihar, Orissa and West Bengal. Our results support their finding that, of these states, Assam had the lowest incidence of poverty and that Orissa and Bihar had the highest incidence of poverty.

5. Conclusions

This paper showed that at least one-third of the average income/probability differences between Hindu and SC/ST households was due to the 'unequal treatment' of SC/ST attributes ('discrimination'). In consequence, the economic position of SC and ST, relative to Hindu, households is a major source of concern. Of course, a part of the reason for the economic backwardness of SC and ST households is that India is a poor country and, naturally, SC and ST households share in its poverty. But, another reason for this backwardness is that in India, inequality and poverty are not doled out fairly. It is an attempt to quantify the degree of this unfairness that provides the *raison d'etre* of this paper.

Although the paper has made frequent use of the term 'discrimination', the interpretation of such a term is open to question. On the one hand, it could mean that "equal persons are treated unequally": two persons with identical endowments are, by virtue of belonging to different communities, rewarded differently. On the other hand, it could mean that the quality of endowment inputs supplied to the different groups is different and the data fails to capture these qualitative differences. For example, a Hindu and a SC/ST household may cultivate the same amount of land with the Hindu household receiving a higher return because its land is more fertile.

It should be emphasised *that at its core*, the treatment of SC/ST persons by Hindus is predicated on the belief that SC/ST persons are inferior to 'caste' Hindus. Not to put too a fine a point on it, such treatment is predicated on a devaluation of their worth as human beings. This, then, has considerable resource implications for SC/ST families: in villages, it is not uncommon for them to be denied access to Hindu wells; in village schools, their children are often made to sit away from caste Hindu children and are routinely referred to as *bhangis*[18]; their women are frequently humiliated and violated; and their houses are located in the low-lying (and, therefore, most liable to flooding) parts of villages[19].

Consequently, SC/ST persons, compared to Hindus, are more likely to be ill, less likely to be adequately educated, more likely to cultivate marginal land and more likely to live in a climate of fear and oppression. In short, the quality and quantity of their economic, educational, and psychosocial endowments are likely to be considerably inferior to that of Hindus. One solution to this is the passive one of making caste-based discrimination illegal. But a more effective solution is to empower SC/ST persons so that they may force a redistribution of resources in their favour. The success of the non-Brahmin movement in southern India meant that caste inequality was addressed there by positive discrimination in favour of non-Brahmins, in education and in job. As recent elections in northern India have shown, there too the ballot box is becoming a powerful instrument for the empowerment of India's 'untouchables'.

References

Ai, Chunrong. and Edward Norton, *Interaction Terms in Nonlinear Models*, Working Paper No. 2, Triangle Health Economics Working Paper Series: www.unc.edu/the, 2001.

Banerjee, Biswajit and J.B. Knight, "Caste Discrimination in the Indian Urban Labour Market", *Journal of Development Economics* 17 (1985): 277-307.

Blinder, Alan S., "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources* 8 (1973): 436-455.

Cowell, Frank and Stephen Jenkins, "How Much Inequality Can We Explain? A Methodology and an Application to the United States", *The Economic Journal* 105 (1995): 421-30.

Deshpande, Ashwini, "Does Caste Still Define Disparity? A Look at Inequality in Kerala, India", *The American Economic Review* (Papers and Proceedings) 20 (2000a): 322-23.

Deshpande, Ashwini, "Recasting Economic Inequality", *Review of Social Economy* 58 (2000b): 381-99.

Deshpande, Ashwini, "Caste at Birth? Redefining Disparity in India", *Review of Development Economics* 5 (2001): 130-44.

Diamond, Charles, Curtis. J. Simon and John T. Warner, "A Multinomial Probability Model of Size Income Distribution", *Journal of Econometrics* 42 (1990): 43-61.

Foster, James, Joel Greer and Erik Thorbecke, "A Class of Decomposable Poverty Measures", *Econometrica* 52 (1984): 571-6.

Gang, Ira, Kunal Sen and Myeong-Su Yun, *Caste, Ethnicity and Poverty in Rural India*, Discussion Paper No. 629, Bonn: Institute for the Study of Labour (IZA), 2002.

Nielsen, Helena S., "Discrimination and Detailed Decomposition in a Logit Model", *Economics Letters* 61 (1998): 115-20.

Oaxaca, Ronald L., "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review* 14 (1973): 693-709.

Oaxaca, Ronald L. and Mark R. Ransom, "On Discrimination and the Decomposition of Wage Differentials", *Journal of Econometrics* 61 (1994): 5-21.

Ravallion, Martin, "Issues in Measuring and Modelling Poverty", *Economic Journal* 106 (1996): 1328-43.

Saggar, Mridul. and Indranil Pan, "SCs and STs in Eastern India: Inequality and Poverty Estimates", *Economic and Political Weekly* 29 (1994): 567-74.

Sainath, P., *Everybody Loves a Good Drought: Stories from India's Poorest Districts*, New Delhi: Penguin, 1996.

Shariff, Abusaleh., *India Human Development Report*, New Delhi: Oxford University Press, 1999.

Srininvas, M.N., "An Obituary On Caste As A System", *Economic and Political Weekl* 38 (2003): 455-460.

Sundaram, K. and Suresh D. Tendulkar, "Poverty in India in the 1990s: An Analysis of Changes in 15 Major States", *Economic and Political Weekly* 38 (2003): 1385-1394.

Theil, Henri., *Economics and Information Theory*, Amsterdam: North-Holland, 1967.

van de Walle, Dominique and Dileni Gunewardena, "Sources of Ethnic Inequality in Vietnam", *Journal of Development Economics* 65 (2001): 177-207.

Table 1
Regression Estimates of the Household Income Equation: 28,922 Households

Hindu Coefficients		***Additional effects from Being Scheduled Caste***	
Determining Variables	*Coefficient Estimate (t value)*	*Determining Variables*	*Coefficient Estimate (t value)*
Central	-0.464 (15.7) [15.2]	Scheduled Caste (SC)	-0.217 (9.7) [9.3]
South	-0.469 (14.4) [13.3]	SC×West	-0.094 (2.7) [2.7]
West	-0.342 (9.7) [9.5]	SC×East	0.107 (3.6) [3.5]
East	-0.498 (16.2) [16.3]	SC×No. Adult Workers	0.032 (4.1) [4.0]
Head educated: > Primary***	0.287 (10.2) [10.4]	SC×Owns Land	-0.028 (1.2) [0.7]
Head educated: ≤ Primary	0.068 (2.2) [2.1]	SC×Acres	0.002 (7.4) [1.9]
>Primary×Central	0.072 (2.2) [2.2]	SC×Productive Assets	-0.012 (4.5) [3.5]
≤ Primary×Central	0.206 (5.8) [5.7]	***Additional effects from Being Scheduled Tribe***	
>Primary×South	0.140 (3.9) [3.7]	Scheduled Tribe (ST)	-0.174 (5.1) [4.9]
≤ Primary×South	0.235 (6.3) [6.2]	ST×West	-0.077 (2.3) [2.3]
>Primary×West	0.117 (2.8) [2.7]	ST×East	0.125 (3.7) [3.5]
≤ Primary×West	0.166 (4.1) [4.1]	ST×>Primary	0.152 (4.5) [4.1]
>Primary×East	0.358 (8.8) [8.7]	ST×≤ Primary	0.05 (1.6) [1.7]
≤ Primary×East	0.224 (5.4) [5.3]	ST×No. adult Workers	-0.0279 (2.9) [2.6]
No. Adult Workers	0.115 (19.7) [18.4]	ST×Owns Land	-0.126 (4.3) [3.3]
No. Adult Workers×Central	0.016 (2.1) [2.1]	ST×Acres	0.003 (1.8) [4.5]
No. Adult Workers×South	0.026 (2.8) [2.4]	ST×Productive Assets	-0.008 (1.8) [1.6]
Owns Land	0.046 (2.2) [2.0]		
Owns Land×Central	0.303 (12.6) [7.7]		
Owns Land×South	0.380 (14.8) [14.6]		
Owns Land×West	0.381 (12.1) [8.3]		
Acres	0.005 (32.2) [12.3]		
Acres×Central	-0.003 (16.9) [3.2]		
Acres×West	-0.002 (12.7) [3.1]		
Productive Assets	0.035 (16.4) [15.7]		
Productive Assets×Central	0.007 (3.0) [2.0]		
Productive Assets×South	-0.016 (4.7) [4.5]		
Productive Assets×West	0.023 (5.5) [3.6]		
Productive Assets×East	0.010 (2.4) [2.3]		
Intercept	9.471 (363.6)		

Notes to Table 2:
Dependent variable is the logarithm of household income
Figures in ( ) are t-values based on 'conventional' standard errors
Figures in [ ] are t-values based on White-corrected standard errors in the presence of heteroscedasticity
Adjusted R^2=0.364; Root Mean Square Error=0.705

Table 2
The Decomposition of Inter-Group Differences
In Mean Household Income

	Sample Average	*Community s treated as community r*		*Community r treated as community s*	
	$\overline{\log(hinc^r)} - \overline{\log(hinc^s)}$	$(\boldsymbol{\theta}^r - \boldsymbol{\theta}^s)'\bar{\mathbf{X}}^s$	$\boldsymbol{\theta}^{r\prime}(\bar{\mathbf{X}}^r - \bar{\mathbf{X}}^s)$	$(\boldsymbol{\theta}^r - \boldsymbol{\theta}^s)'\bar{\mathbf{X}}^r$	$\boldsymbol{\theta}^{s\prime}(\bar{\mathbf{X}}^r - \bar{\mathbf{X}}^s)$
***r*=Hindu**	**9.953-9.542=**	**9.692-9.542 =**	**9.953-9.692 =**	**9.953-9.820 =**	**9.827-9.542 =**
***s*=SC**	**0.411**	**0.150**	**0.261**	**0.133**	**0.278**
***r*=Hindu**	**9.953-9.556=**	**9.737-9.556 =**	**9.953-9.737 =**	**9.953-9.799 =**	**9.799-9.556 =**
***s*=ST**	**0.397**	**0.181**	**0.216**	**0.154**	**0.243**

Table 3
The Decomposition of Inter-Group Differences
in the Proportion of Households in the Top Income Quintile:
Multinomial Probabilities Decomposition

	Sample Average	*Community s treated as community r*		*Community r treated as community s*	
	$\bar{P}_5^r - \bar{P}_5^s$	$P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_5^r) - P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_5^s)$	$P(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_5^r) - P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_5^r)$	$\bar{P}(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_5^r) - \bar{P}(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_5^s)$	$P(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_5^s) - P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_5^s)$
***r*=Hindu**	**0.263-0.107=**	**0.153-0.107 =**	**0.263-0.153 =**	**0.263-0.206 =**	**0.206-0.107 =**
***s*=SC**	**0.156**	**0.046**	**0.110**	**0.057**	**0.099**
***r*=Hindu**	**0.263-0.123=**	**0.172-0.123 =**	**0.263-0.172 =**	**0.263-0.222 =**	**0.222-0.123 =**
***s*=ST**	**0.140**	**0.049**	**0.091**	**0.041**	**0.099**

Table 4
The Decomposition of Inter-Group Differences
in the Proportion of Households in the Bottom Income Quintile:
Multinomial Probabilities Decomposition

	Sample Average	*Community s treated as community r*		*Community r treated as community s*	
	$\bar{P}_1^r - \bar{P}_1^s$	$P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_1^r) - P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_1^s)$	$P(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_1^r) - P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_1^r)$	$\bar{P}(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_1^r) - \bar{P}(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_1^s)$	$P(\mathbf{X}_i^r, \hat{\boldsymbol{\beta}}_1^s) - P(\mathbf{X}_i^s, \hat{\boldsymbol{\beta}}_1^s)$
***r*=Hindu**	**0.157-0.266=**	**0.227-0.266 =**	**0.157-0.227 =**	**0.157-0.181 =**	**0.181-0.266 =**
***s*=SC**	**-0.109**	**-0.039**	**-0.070**	**-0.024**	**-0.085**
***r*=Hindu**	**0.157-0.264=**	**0.207-0.264 =**	**0.157-0.207 =**	**0.157-0.208 =**	**0.208-0.264 =**
***s*=ST**	**-0.107**	**-0.057**	**-0.050**	**-0.051**	**-0.056**

Table 5
Proportion of Poor Households

	All Households (%)	*Hindu Households (%)*	*SC Households (%)*	*ST Households (%)*
Not Poor	64	71	54	53
Mildly Poor	17	14	21	22
Moderately Poor	14	11	19	19
Very Poor	5	4	6	6

'Not Poor': households with income above 75% of median income (of all 28,922 households)
'Mildly Poor': households with income between 75% and 50% of median income
'Moderately Poor': households with income between 50% and 25% of median income
'Very Poor': households with income below 25% of median income

Table 6
Estimates of Discrimination Against SC and ST Households in Their Probabilities of Being Poor (%)

Poverty Level↓	*SC treated as Hindus*	*Hindus Treated as SC*	*ST Treated as Hindus*	*Hindus Treated as ST*
Non-Poor	39	27	58	46
Mildly Poor	40	29	63	41
Moderately Poor	41	30	57	53
Very Poor	35	8	44	22

The numbers in Table 11 represent the proportion of the total difference, between Hindu and SC/ST households, in their average probabilities of being at a poverty level, which can be explained by coefficient differences between the communities.

Notes

[1] Mainly in the form of reserved seats in the national parliament, state legislatures, municipality boards and village councils (*panchayats*); job reservations in the public sector; and reserved places in public higher educational institutions

[2] See Sundaram and Tendulkar (2003) as an example of non-caste studies of deprivation in India. The study of caste and inequality is important because, as Srinivas (2003) has pointed out, economic relations in rural India are embedded in social relations and a defining feature of such social relations is hierarchy as expressed through the caste system. Landowners were at the top of the economic ladder while the landless occupied the lowest rungs. In the ownership of land, "a pan-Indian phenomenon was the large overlap between landlessness and the 'untouchable' castes, a fact which enhanced their poverty, misery and exploitability" (p.455).

[3] Note that differences between groups in income-generating attributes can reflect historical discrimination: for example, a group is currently disadvantaged educationally because, in the past, it was denied access to education.

[4] The value of the productive assets index for a household was computed as the weighted sum of its productive assets. These assets were (with weights in parentheses): sewing Machine (2); tubewell (10); generator (5); thresher (3); winnower (3); bullock cart (4); cycle rickshaw (3); tractor (10).

[5] Over 90% of the combined SC/ST respondents gave their religion as 'Hindu'.

[6] The Central region comprising Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh; the South comprising Andhra Pradesh, Karnataka, Kerala and Tamil Nadu; the West comprising Maharashtra and Gujarat; the East comprising Assam, Bengal and Orissa; and the North comprising Haryana, Himachal Pradesh and Punjab.

[7] Note that the default household outcomes are: lives in the North; is Hindu; head is illiterate.

[8] $st_i = sc_i = central_i = south_i = west_i = awk_i = lnd_i = acr_i = pai_i = east_i = edm_i = edh_i = 0$.

[9] For example, the joint effect of being a SC household, with three adult workers and with a head educated to above primary level, living in the Central region and owning ten acres of land *ceteris paribus* was to add to/subtract from baseline income the amount:

$$\alpha_1 + \alpha_6 + \beta_5 + (\gamma_1^{awk} + \delta_2^{awk}) \times 3 + (\gamma_1^{lnd} + \delta_2^{lnd}) + (\gamma_1^{acr} + \delta_2^{acr}) \times 10 + (\gamma_1^{edh} + \delta_2^{edh}).$$

By contrast, the corresponding addition/subtraction for a Hindu household in similar circumstances would be: $\alpha_1 + (\gamma_1^{awk} \times 3) + \gamma_1^{lnd} + (\gamma_1^{acr} \times 10) + \gamma_1^{edh}$.

[10] Details available on request.

[11] On the basis of a likelihood ratio test, the joint validity of these zero restriction was not rejected with $\chi^2(11) = 11.8$.

[12] The reason for including all the interaction terms (and not excluding the 'insignificant' ones, as in Table 2) is that Ai and Norton (2001) have drawn attention to the dangers of eliminating non-significant 'interaction' terms in *non-linear* models like that of the multinomial logit.

[13] For example, a positive coefficient attached to a variable for quintile *j* means that that the odds-ratio of being in that quintile rises as the value of the variable increases; however, it does not necessarily imply that the probability of being in that quintile goes up.

[14] Details available on request.

[15] If, indeed, inequality can be 'additively decomposed' then, as Cowell and Jenkins (1995) have shown, the proportionate contribution of the between-group component (**B**) to overall inequality is the income inequality literature's analogue of the R^2 statistic used in regression analysis: the size of this contribution is a measure of the amount of inequality that can be 'explained' by the factor (or factors) used to subdivide the sample.

[16] Details available on request.

[17] Details available on request.

[18] Persons who clean toilets.

[19] See Sainath (1996) for a discussion of the lives of India's untouchable castes.