Discussion Papers in Economics

Caste Dominance and Economic Performance in Rural India

Vegard Iversen

Adriaan Kalwij

Arjan Verschoor

Amaresh Dubey

February 2010

Discussion Paper 10-01

Indian Statistical Institute, Delhi
Planning Unit
7, S. J. S. Sansanwal Marg, New Delhi 110016, India

Caste dominance and economic performance in rural India[♦]

Vegard Iversen[£], Adriaan Kalwij[$], Arjan Verschoor* and Amaresh Dubey**

£IDPM, University of Manchester; Indian Statistical Institute, Planning Unit, $ Utrecht University, *University of East Anglia, * *Jawaharlal Nehru University

ABSTRACT

Using a unique household panel data set for rural India covering the years 1993/1994 and 2004/2005 we test a key theoretical assertion of caste and its effects, namely that marginalised social groups fare worse in terms of income levels when resident in villages dominated by upper castes. We also test whether marginalised groups perform better or worse in villages where their own group is dominant. We proceed to explore the implications for income growth and for poverty incidence and persistence. After controlling for potential locational confounds, upper caste dominance confers a positive externality on other social groups. This externality is discounted by group specific 'oppression' effects which range in size from zero to 16 percent of mean income and peak for Scheduled Caste (SC) and Other Backward Classes (OBC) households. Further, we identify positive and large own dominance (village 'enclave') effects that account for as much as a quarter of mean income for SC households in the post reform years. These results are robust to how dominance is measured. We also identify pathways through which identity-based welfare disparities may be reduced; while such disparities are widening, their causes show signs of both persistence and change. Whereas education matters, land redistribution provides the key to eliminating the adverse effects of upper caste dominance. Even after factor endowment and other controls have been added, and with the notable exception of those in SC dominated villages, SCs not only perform worse than other groups but have fallen further behind during the post reform years.

[♦] For very helpful comments and suggestions, we are indebted to Farzana Afridi, Siwan Anderson, Stefan Jonsson, Richard Palmer-Jones, Rinku Murgai and Kunal Sen. A special thanks to Richard Palmer-Jones for sharing the classification of agro-ecological zones used in this paper.

'March 1949: A group of Scheduled Caste members from villages around Delhi had been thrown out of their homes by Jat landowners angered that these previously bonded servants had the cheek to take part in local elections and graze their cattle on the village commons.

June 1951: A village in Himachal Pradesh. A conference of Scheduled Castes is attacked by Rajput landlords. The SCs are beaten up with sticks, their leaders tied up with ropes and confined to a cattle pound.

June 1952: A village in the Madurai district of Madras State. A SC youth asks for tea in a glass at a local shop. Tradition entitles him only to a disposable coconut shell. When he persists, he is kicked and hit on the head by caste Hindus.

June 1957: A village in the Parbani district of Madhya Bharat. Newly converted Buddhists [previously "untouchable" Hindus] refuse to flay carcasses of dead cattle. They are boycotted by the Hindu landlords, denied other work and threatened with physical reprisals.' (Guha 2007; 380-81)

INTRODUCTION

More than 50 years on and in spite of such early resistance, the aggressive use of affirmative action and radical legislative interventions, identity-based inequality and poverty have remained stark features of post-independence India (Deshpande 2001; Kijima 2006; Gang et al 2008a). Notwithstanding that the intensity of caste[1] barriers may have softened with the passage of time,[2] extensive reservations in public sector jobs, higher education institutions and political assemblies have not sufficed to prevent Scheduled Caste and Scheduled Tribe households from being overrepresented among the country's poor, illiterate and in the former case, also the landless.[3]

[1] Caste may refer to *jati* (sub-caste) or the more general *varna*, the latter comprising four broad occupational groups with Brahmins at the top and untouchables as a separate category. Each *varna* contains innumerable *jatis* who with few exceptions practice intra-marriage (endogamy).

[2] Examples from the recent past include caste demarcators in how people dressed and spoke and what they were allowed to do. In 19th Century Kerala, "*when a Namboodiri Brahmin approached, a Paraiya labourer had to cry out in advance, lest the sight of him pollute his superior*" (Guha 2007; 287). Also in Kerala and when talking to a person of higher caste, members of lowly ranked castes were expected to use debasing words to describe themselves (Menon 1994;19). Nambissan (1996) presents historical illustrations of how Scheduled Caste children, while permitted to attend school, could be denied entry to the classroom.

[3] Scheduled Castes, comprising 16.2 % of the country's population in 2001, are former 'untouchables', while Scheduled Tribes, accounting for 8.2 % of the population in the same year, are tribes perceived as historically disadvantaged. Article 15 of India's Constitution emphasises the 'disabling' effect of low social status and article 46 underscores the collective responsibility for promoting the educational and economic interests of individuals of Scheduled Caste and Scheduled Tribe backgrounds (Bayly 1999; 68-69). The Constitution also refers to an additional category of disadvantaged citizens, an issue addressed firstly by the Other Backward Class Commission appointed by Prime Minister Nehru, and later and more decisively by the Mandal Commission (1978-80). The

While the endurance of practices expressive of a strict social hierarchy[4] might be expected to cement disparities and accentuate ill-being, key questions about the mechanisms and pathways through which caste, tribe and religious[5] identities translate into present day disadvantage have yet to be satisfactorily answered. In addition, little remains known about whether, and if so precisely in what directions, patterns, magnitudes and causes of identity-based disadvantage may be transmuting in the post reform era.[6]

In this paper we use a household panel covering the years 1993/1994 and 2004/2005 to examine two possible explanations for identity-based disadvantage in rural India. The first, the oppression hypothesis, originates in M. N. Srinivas's theory of caste dominance which portrays a caste that in addition to strong numerical presence is also economically powerful (Srinivas 1955). This oppression hypothesis is discrimination oriented, epitomises the introductory quotes and advances the view that historically marginalised groups fare worse when resident in villages dominated by upper castes. We test this hypothesis for income levels, and explore the implications for income growth and for poverty incidence and persistence.

The second, the village enclave explanation, is theoretically ambiguous and captures a situation where a historically marginalised group is dominant at the village level. Upwards mobility may then be inhibited by factors that either wholly or in part are internal to the group in question. Munshi and Rosenzweig (2006) illustrate one such mechanism albeit in the urban context of Dadar, Mumbai, where the density of *jati* based labour market networks, via their effects on educational choice, is held responsible for slow upwards mobility across generations of low caste, young men. [7] In a rural setting, a strong preference for traditional occupations or the onset of social inertia (e.g. Peyton Young 2001) could give rise to similar, 'interlocking' effects.

Empirical studies of education and labour market outcomes in (mainly immigrant) enclaves infuse more optimism about enclave potential (e.g. Edin et al 2003; Cortes

---

Mandal Commission's recommendations, which extended reservation benefits to OBCs, were declared constitutionally legitimate by the Supreme Court in 1992 after agitations and intense controversy (Parikh 1998).

[4] Such strictness resonates with Louis Dumont's (1970) portrayal of purity and pollution as immutable principles of Hindu society (Bayly 1999; 15). M. N. Srinivas (1966) gives numerous examples of more fluid inter-caste relations, arguing that analysis of social relations through the lens of *varna 'successfully obscured the dynamic features of caste during the traditional or pre-British period* (ibid. 1966;2).'

[5] Muslims, the largest religious minority, accounted for 13.4 % of India's population in 2001.

[6] Our reference here is to India's landmark economic liberalisation programme, initiated in 1991, but with its main effects kicking in only after the first panel round (See footnote 14).

[7] In contrast and using NSS data, Das (2007) finds evidence of successful self-employment enclaves among educated Muslim men in India.

2006). A less hostile village environment could prevent the psychological internalisation of self-fulfilling and negative self beliefs that might otherwise lock individuals of marginalised backgrounds into low level equilibrium traps (Akerlof and Kranton 2000; Hoff and Pandey 2006). By reducing the social distance between parties to rural transactions, own enclaves could also, as Anderson's (2005) evidence demonstrates, improve the operation of vital rural markets.

While the oppression and the *negative* enclave explanation may inhibit income growth and contribute to poverty persistence, there is a marked contrast in how policies could address and alleviate identity-based disparities. While oppression would require society-wide attitudinal transformations, the *negative* enclave explanation would call for reforms targeting the marginalised community itself. The *positive* enclave explanation adds an intriguing policy dilemma: marginalised groups performing better in their own village enclaves could weaken the case for social integration.

BACKGROUND AND CONTRIBUTION TO THE LITERATURE

In India, empirical research on caste has focused on discrimination, mainly within the labour market (e.g. Banerjee and Knight 1985; Kingdon 1998 and 2002; Dutta 2006; Iversen and Raghavendra 2006; Thorat and Attewell 2007). The evidence shows that individuals of SC and ST background are indeed at a disadvantage[8] – through lower wages, a higher propensity of being stuck in dead end jobs (e.g. Banerjee and Knight 1985) or inferior employment terms, such as casual employment (Dutta 2006; Das and Dutta 2007). Recent research has also documented how discrimination extends to upper end urban occupations and jobs (Madeshwaran and Attewell 2007; Newman and Deshpande 2007).

While important, the labour market is only one market where individuals from marginalised social groups may experience differential treatment. To date, however, little systematic knowledge exists about discrimination in credit, insurance or other key markets or particular to rural areas, markets for agricultural inputs and outputs.[9] There is also limited evidence on whether a person's caste, religious or tribal identity

[8] Gang et al (2008b) confirm that OBCs also have lower living standards relative to the mainstream population.
[9] Hatlebakk (2009) presents evidence on the relationship between caste and credit transaction terms in rural Nepal.

circumscribes the access to poverty-oriented public policy programmes or public services in general.[10]

Even so, Shah et al's (2006) recent study of untouchability in rural India found that in 45-50 percent of the villages surveyed, Dalits were prevented from selling milk to village dairy cooperatives.[11] Such 'bans' could be rooted in purity and pollution ideals and the intimate links between a person's identity and the preparation and handling of food and water; the same study found that in 30-40 percent of the same villages, Dalits were prevented from full participation in local markets and often from entering village shops.[12] Further and wellknown, SC hamlets tend to be separate from the main village and often have their own drinking water source.

In rural Karnataka, children from orthodox Brahmin households may be forced to take a bath before entering their house after interacting with peers from 'inferior' castes while in school. A rich gamut of mechanisms for differential treatment thus persists and may affect everyday social interactions, limit the willingness to transact, the terms of such transactions as well as the access to productive resources, public policy programmes and services. Put differently, oppression could, in principle, circumscribe the access to all markets (land, labour, credit, insurance, output, and input markets), affect production costs and limit investment opportunities and returns.

Our data do not permit a precise identification of discrimination within a particular market or in the access to a specific public service, but facilitate instead identification of upper caste oppression and of negative or positive village enclave effects on household economic welfare, measured by income. Further, by introducing control variables gradually in our econometric analysis, we are able to explore the pathways through which such oppression and enclave effects operate and whether the latter are transmuting in the post reform era.

Our study advances the literature as follows. Aggregating across markets, as we do, a small number of studies test for identity-based disadvantage, but do not test the effects of village level upper caste or own group dominance on the economic

[10] Exceptions include Dreze and Kingdon's (2001) study of school participation in rural India, suggesting that Scheduled Caste children have an 'intrinsic disadvantage' and a lower chance of attending school even after household wealth, parental education and motivation and school quality are controlled for. See also Thorat and Lee's (2006) study of discrimination in food related government programmes. Banerjee and Somanathan (2007) explore, among other issues, the effects of community heterogeneity on public services (goods) availability in India.

[11] The study covered 550 villages in 11 main states.

[12] As noted by Madsen (1991), Parry (1999), Iversen and Raghavendra (2006) among others, purity and pollution ideals make the connection between caste identity and the handling, preparation and serving of food acutely sensitive.

performance of different social groups. Typically, existing studies use regression analysis of nationally representative cross-sectional data and Blinder Oaxaca or alternative decomposition techniques to quantify the disadvantage associated with Scheduled Caste, Scheduled Tribe or religious identity (e.g. Kijima 2006; Gang et al 2008a). Using All-India data, Kijima's (2006) results suggest that such contrasting returns may account for up to 50 percent of the difference in mean consumption expenditure between SC/ST households and others. While Dercon and Krishnan (2007) use the ICRISAT household panel to evaluate the relative economic performance of SC/STs, their panel is limited to 204 households from six villages and two states. The lower educational attainment of SC/STs is found to fully account for their slower standard of living improvements. Dercon and Krishnan are unable to extensively test for the impact of village social structure and land ownership patterns on the relative welfare of marginalised social groups.

In an innovative study particularly close to ours, Anderson (2005) presents the first attempt to make use of sociological and anthropological notions of caste dominance in an econometric analysis. For a data-set covering 120 villages in Uttar Pradesh and Bihar, she is able to shed light on the mechanisms through which caste based disparities emerge and may be sustained. Specifically, she finds that Yadav households in villages where Yadavs are the dominant land owners have higher incomes than Yadav households in villages where the dominant land owners belong to a local upper caste. She attributes this result to the market for irrigation water's failure to operate in villages with upper caste land dominance and concludes that social distance may prevent the efficient operation of crucial rural markets.

Anderson's study uses cross-sectional data for north and central Bihar and south and south-eastern Uttar Pradesh, which are part of India's "poverty belt" and more than elsewhere in the country riddled by inter-caste tensions and conflict (Bayly 1999;345). We use a panel data set that in its base year 1993/1994 is representative for rural areas of most of India's major states (see section II). In contrast to Anderson, we estimate the impacts of upper caste and own group dominance on household income and explore the implications for growth and poverty for social groups where such effects may be expected to be most pronounced.[13] We also distinguish

[13] The Yadavs in Anderson's study are OBCs (Other Backward Classes) while the Scheduled Castes in our panel data set include numerous former untouchable, low status *jatis* such as Chamars or Satnamis (leatherworkers), Balmikis (sweepers) and Pasis in the North and Pallars and Paraiyars in the South.

externalities associated with residence in upper caste dominated villages from social group specific oppression effects within the same village regimes. Finally, our panel households are followed from 1993/1994 up to 2004/2005 which enables us to investigate whether India's economic liberalisation programme, initiated in 1991 but with its main effects being felt only after the first panel round,[14] has mitigated, augmented or otherwise changed the effects of social distance, mediated through a traditional social hierarchy, on the functioning of markets and in turn household incomes.

The paper is laid out as follows. Section II describes the data set, elaborates on the theoretical background and presents the empirical model for testing our hypotheses. Section III presents descriptive statistics on income and poverty levels and change and on education and land endowments by social group and village regime. Section IV presents the main empirical results, followed by robustness tests, and a computation of counterfactual income, growth and poverty to illustrate the order of magnitude of the oppression and enclave effects we are able to identify. Section V concludes.

II. DATA, THEORETICAL BACKGROUND AND EMPIRCAL FRAMEWORK

A. The data set

The data reported on here are derived from two large-scale household surveys that cover most of the territory of India, known as the Human Development Profile of India (HDPI) surveys. The first round, HDPI-I, took place in 1993/94 and was carried out by India's National Council of Applied Economic Research (NCAER) on behalf of UNDP. The second round, HDPI-II, took place in 2004/05 and was carried out by NCAER on behalf of the University of Maryland. The primary purpose of the surveys was to collect detailed information on a large range of human development indicators, including income, which is the variable reported on here. These surveys are the first major ones for India to measure household income in a comprehensive and refined way, using more than fifty separate components. A full description of the variables,

[14] Neither GDP growth, growth in the services sector nor private sector investment had picked up by the time the first panel round (1993/94) had been completed. For supportive evidence as well as fuller and more complex accounts of India's growth turnaround, see Sen (2007) and Panagariya (2008).

summary statistics including comparison with other major India surveys[15], and an exposition of the sampling methodology can be found in Desai et al. (2009).

A unique feature of this data is that a village questionnaire was administered in the second round in 2004/2005 and enables the construction of village social composition and land ownership distribution by jati (sub-caste). Further and similarly, the sub-division of social groups in the household questionnaires allows us, in addition to the official categories of Scheduled Castes, Scheduled Tribes, Muslims, OBCs, [16] other Hindus, and Others, to precisely identify the jati of individual households and thus to make comparisons of the economic performance of other social groups with that of upper-caste households, who mostly are Hindus.[17] These features depart notably from official data sets with collection of information on jati terminated after the 1931 Census.

The first round of the survey uses a random sample of 33,230 households located in and representative of each of the rural areas in all (then sixteen) India's major states. Initially, the aim was to re-interview 13,593 randomly selected rural households in the second round. Recontact details were, for various reasons, not always available and in the end 10,451 households in fourteen (plus three new) states participated in both rounds.[18] A residence-based sampling rule was adopted involving only households who had stayed in the village; migration and natural demise are reasons for attrition. After removing about 20 villages with missing social composition and land ownership information, we are left with a panel comprising 9,111 households spread over 679 villages.

The findings reported here are strictly speaking valid only for households who choose not to migrate (cf. Rosenzweig 2003, Baulch and Hoddinott 2000, Dercon and Krishnan 2007). However, the comparison of living standards and changes therein

[15] There is a close correspondence between the HDPI and other major surveys on mean values of all key variables; see Desai et al. (2009), Table 2.

[16] OBC lists, which include Muslims, are state-specific, regularly updated and rapidly expanding; entries often reflect political muscle rather than past discrimination. As Appendix 1 explains, manifestly political inclusions on the official list are reclassified and omitted from the OBC category used in our analysis.

[17] See Singh's (1984) account of caste among non-Hindus and Jodhka's (2004) in-depth discussion of Sikhism and caste. Punjab and to a lesser extent Haryana have many Hindu and Sikh households who all belong to the Jat community as well as Sikh SC households. Among Muslims, Fuller (1996) and other contributors to the same volume contend that while caste-like arrangements are common, few willingly admit to their existence. Jeffrey et al (2007; 43) note how '*during the pre-colonial era there were marked divisions between a very small, upper caste Muslim elite and other Muslims castes, such as weavers, carpenters and barbers*'.

[18] They are Andhra Pradesh, Bihar (+ Jharkand in round 2), Gujarat, Haryana, Himachal Pradesh, Kerala, Madhya Pradesh (+ Chhattisgarh in round 2), Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh (+ Uttaranchal in round 2) and West Bengal. Recontact details were largely lost in Assam due to a flood and in Karnataka because of human error.

across social groups – the focus of this paper – should not be much affected by this limitation: the variables caste, religion, education and income are not substantially different in the panel from those in a randomly selected rural refresher sample drawn to check the round 2 representativeness of the panel household sample (see Table 1 in the Appendix in Desai et al. 2009).[19] Furthermore, we performed a statistical test on whether or not inclusion in the panel of all households who participated in the first round is associated with our independent variable in the analysis, namely household income. Such a test provides evidence on whether or not the panel households are a selective group of households with respect to income. The test result reveals that after controlling for households' demographic composition and educational attainment, household income is not associated with selection into the panel,[20] suggesting that the panel households in our analysis, with respect to income, are a randomly selected subsample of all rural households that participated in the first round.

B. Upper caste and own dominance – theory and definitions

The caste dominance concept originates in the sociological and anthropological literature. In Srinivas's (1955; 18) own words:

'*A caste may be said to be 'dominant' when it preponderates numerically over the other castes and when it wields preponderant economic and political power. A large and powerful caste group can more easily be dominant if its position in the local hierarchy is not too low.*'[21]

The distinguishing characteristic of upper caste dominance may perhaps best be expressed as a combination of secular power and ritual status where the latter reflects the Varna hierarchical order with Brahmins topmost among four broad occupational ranks and with former untouchables (SCs) making up a separate category. The dominant social group could now be defined as the group (i) which represents a larger share of the village population than any other social group (n_d); (ii) owning more

[19] The table referred to reports the proportions of the panel household sample in round 2 and those of the refresher sample in categories of age (8 categories), gender (2), individual education (6), social group (6), place of residence (4), maximum adult education (6), and income (6). The absolute differences between the proportions of the two samples (38 comparisons in total) range from 0.04 to 5.28 percentage points, with a mean value of 1.20 and a median of 0.56 percentage points.

[20] The p-value corresponding to the null hypothesis that income is not associated with panel inclusion is equal to 0.937.

[21] Srinivas (1959) subsequently added Western education and occupations to his list of determinants of dominance.

village land than any other social group (l_d) (e.g. Dumont 1970); or (iii) both n_d and l_d (e.g. Srinivas 1955). While not exhaustive, (i)-(iii) clearly represent alternative measures of secular power.

Numerical strength could translate into village level political muscle especially after the 73rd Constitutional Amendment's elevation of the status and significance of village Panchayats.[22] However, Anderson (2005) finds no effects of population dominance on economic outcomes. As we explain below, our adopted empirical focus on land dominance partly reflects a constraint imposed by de facto village structures in rural India but also exploratory regressions supportive of Anderson's (2005) approach and Dumont's (1970) assertion that dominance is rooted in economic power captured by landownership alone.[23]

Conceptually, let the land of village j, L_j, be distributed over m groups where n_i represents the share of the village land that belongs to social group i. Hence,

$$L_j = \sum_{i=1}^{m} n_i = 1 \qquad (1)$$

Definitions: *A dominant social group has the largest share of the village land of any social group. For members of the dominant social group in village j, village j is described as own group dominated or an own enclave. If the dominant social group in village j is upper caste, village j is upper caste dominated. Upper caste dominance is an example of what we call a village regime.*

This forms the conceptual backbone for our main analysis. Two types of criticisms may be levelled against this sociologically anchored dominance measure. Firstly, it neglects fragmentation or concentration among other social groups within a village; the more diverse and fragmented the remaining social groups, the more powerful the dominant group is expected to be. In addition, our dominance measure neglects the intensity of the power the dominant group is in a position to wield over other social groups. To neutralise such reservations, we use two alternative dominance measures as robustness checks. The first is the share of village land owned by the dominant group, the second a modified Herfindahl index.

[22] Panchayats are village councils. The 73rd Amendment came into force in April 1993.
[23] These results are not presented or further discussed here, but will be made available from an author website for interested readers.

The normalised Herfindahl (H) index is a popular measure of concentration or market power within economics. Its mirror image, the fractionalization index, has been widely used in empirical studies linking community heterogeneity to conflict, collective action and the access to public goods (e.g. Olson 1971; Banerjee and Somanathan 2007; Esteban and Ray 2009). However, neither is well equipped for capturing dominance.

To see why, consider the Herfindahl index of concentration for village j which can be defined as:

$$H_j = \sum_{i=1}^{m} n_i^2 \quad \text{where} \quad H_j \, \varepsilon \, \langle 0,1] \tag{2}$$

To gauge why H_j fails, let m=2 and n_1=n_2=0.5. While the market equivalent would be a situation of considerable concentration (limited fragmentation or fractionalization), the scope for group 1 to dominate group 2 should be exactly zero. To adjust for this weakness, we introduce the following modification:

$$D_j = n_d^2 - \sum_{i \neq d} n_i^2 \tag{3}$$

where the subscript d refers to the land share owned by the dominant group. For a given share of the village land owned by the dominant group, the more fragmented is the land ownership of other groups, the higher is D_j. In the example above, its value will be exactly zero, as it should be.

To construct our village level dominance measures we combine village level information on social structure and land ownership with evidence of the locale-specific hierarchical status of precisely identified jatis. The village questionnaire administered in round 2 identifies the jati of the numerically dominant social group in each village, the percentage of village land this social group owns along with similar information for the next 4-8 most numerous social groups. Anthropological and other evidence on the status of different jatis is then used to develop a more refined upper caste definition as explained in Appendix 1. Given generally inactive rural land markets (Anderson 2005), and that land-dominant groups tend to hold a much larger share of village land than any other group,[24] we assume that the village regime is identical in rounds 1 and 2.

[24] Details are available from the authors.

C. Empirical model

Both the oppression and the enclave hypothesis refer to the extent to which the income level of households from different social groups is affected by the social identity of the dominant land owners of the village in which they reside. To statistically test these hypotheses, we model the relative differences in income by social group and village regime, controlling for location and household characteristics, as follows:

$$\begin{aligned}\ln(Y_{ht}) = {} & \alpha_{0t} + \alpha_{1t}SC_h + \alpha_{2t}ST_h + \alpha_{3t}MUS_h + \alpha_{4t}OBC_h \\ & + \beta_{1t}SC_h \times DSC_h + \beta_{2t}ST_h \times DST_h \\ & + \beta_{3t}MUS_h \times DMUS_h + \beta_{4t}OBC_h \times DOBC_h \\ & + \gamma_{0t}DUC_h + \gamma_{1t}SC_h \times DUC_h + \gamma_{2t}ST_h \times DUC_h \\ & + \gamma_{3t}MUS_h \times DUC_h + \gamma_{4t}OBC_h \times DUC_h \\ & + \pi_t X_{ht} + \theta_h + \eta_{v(h)} + \varepsilon_{ht}, \\ t=\{1993/1994, & 2004/2005\}. \end{aligned} \quad (4)$$

Subscript h denotes households, and t time. The real per capita income of a household[25] is denoted by Y and the five social groups a household can belong to are denoted by SC (Scheduled Castes), ST (Scheduled Tribes), MUS (Muslims), OBC (Other Backward Classes) and UC (Upper Caste). These are all dummy variables and take the value one if a household belongs to this group and zero otherwise. The village regime is modeled using the dummy variables DSC, DST, $DMUS$, $DOBC$ and DUC, which take the value one if this particular social group is land dominant in the village in which the household lives, and zero otherwise.

The last three terms in the right hand side of equation (4) form the error structure of the model. The first two error terms are, respectively, a random household specific effect, θ_h, that is assumed to be independently distributed across households, and a random village specific effect, $\eta_{v(h)t}$, which is assumed to be independently distributed across villages. The third error term, ε_{ht}, is an idiosyncratic error term and

[25] Throughout income is per capita per annum and in constant 1993/94 prices, converted using NSSO state-specific rural CPIs.

is assumed to be independently distributed across households, villages and time. The assumption of a random household specific effect, as opposed to a fixed effect, is required since incorporating a household specific fixed effect would make it impossible to identify oppression and enclave effects since, as noted, the village regime is constant over time and a panel household lives in the same village in both rounds.

We estimate equation (4) by Least Squares for each of the two periods separately as all parameters are allowed to vary over time. Arbitrary correlation between households within a village is accounted for when calculating the standard errors (e.g., Cameron and Trivedi 2005).

The α-parameters refer to the relative income differences between households of different social groups with *UC* as reference group. For instance, the parameter α_{1t} (x100) corresponding to the variable SC_h, is interpreted as the percentage difference in income, on average, between *SC* and *UC* households. Similarly, the enclave hypothesis refers to the β-parameters and the parameter β_{1t} (x100) corresponding to the variable SC_hxDSC_h, is interpreted as the percentage difference in income, on average, for *SC* households living in a village dominated by their own social group. The oppression hypothesis refers to the γ-parameters. The parameter γ_0 corresponding to the variable DUC_h relates to the relative income effect for all households living in a village dominated by Upper Castes (*UC*). The parameter γ_{1t} (x100) corresponding to the variable SC_hxDUC_h, say, is interpreted as the percentage difference in income, on average, between *SC* and *UC* households living in an *UC* dominant village.

Without controlling for potential locational confounds, the enclave and oppression effects could simply pick up that, for instance, upper caste dominated villages are located in areas with a favorable resource base and greater agricultural potential.[26] Another possible locational confound relates to policies and governance in the state where a village is located and a household resides. To address these important concerns, we use Palmer-Jones and Sen's (2003) agroecological zones[27] and state dummy variables as controls. Both sets of variables are included in the vector of control variables (*X*) in equation (4). In addition, X includes variables for household

[26] The relevance of locational disadvantage, which corresponds highly imperfectly with state boundaries, for poverty (and inequality) in rural India, has been extensively documented by Palmer-Jones and Sen (2003).
[27] Their map (Palmer-Jones and Sen (2003; 14-15) divides India into 19 different zones where very careful classifications of land surface capture initial conditions that indicate agricultural productivity potential.

demographic composition, education and land holdings, and for village infrastructure (the full list of variables is in Appendix 2).

All parameters of equation (4) are allowed to vary with time which makes it possible to investigate changes in oppression and enclave effects between the two rounds and, subsequently, the implications for income growth and for poverty incidence and persistence. As discussed in section B, we explore the robustness of the main results to two alternative measures of dominance and for this purpose we replace the dummy dominance variables (e.g. *DUC*) with the upper caste land share (the first alternative) or the value of the dominance adjusted Herfindahl-index (the second alternative, eq. (3)).

III DESCRIPTIVE STATISTICS

This section presents descriptive statistics on village regimes that are pertinent to the oppression and enclave hypotheses.[28] Anchored in Dumont's (1970) conception of caste dominance, as set out above, our empirical focus is on villages in which a particular social group owns the largest proportion of village land.

We first, however, present data on land and population dominance in the 679 villages in our panel.

Table I: The distribution of land and population dominance by village and social group

	Largest land-holding group in village						
	SC	ST	OBC	MUS	UC	OTH	Total
Largest population group in village:							
Scheduled Castes (SC)	**24**	2	12	1	25	4	68
Scheduled Tribes (ST)	0	**65**	3	0	2	0	70
Other Backward Classes (OBC)	1	0	**196**	3	25	10	235
Muslims (MUS)	0	0	2	**35**	4	2	43
Upper Castes (UC)	0	1	1	0	**223**	6	231
Others and none (OTH)	2	0	8	0	18	**1**	9
Total	27	68	222	39	297	26	**679**

[28] In a companion paper, we present other descriptive statistics for this panel including mean household income by state, land holdings, levels of education (of the household head), occupation and real household income per capita for different social groups and find a close correspondence between a priori expectations and summary statistics. Kerala has the highest level of household income per capita (but is represented by a small number of households in the panel, n=48), closely followed by Punjab and Haryana. Chhattisgarh, Assam and Orissa are the three bottom states. Further, marginalised social groups own less land and are less educated than others. 41% of SC households and 48 % of Muslim households have their own land; the figures for STs, OBC and UCs are 70%, 63% and 81%, respectively. One contrast between SC and Muslim households is that the latter have more irrigated land, on average. Apart from lower levels of education and consistent with Kijima (2006), marginalised communities also appear to receive lower returns on their human capital.

Source: HDPI panel, authors' calculations
Notes: Figures are number of panel villages in which the row social groups are the largest population group and the column social groups own the largest land share. The category "others" consists of villages in which either an unclassified group or no single group is land- or population-dominant.

The technical challenge posed by separate identification of land and population dominance may be gauged from the diagonal which shows that population and land dominance are strongly correlated: for each social group, if it is population dominant, in over 90 percent of cases, it is also land dominant, and vice versa. There are two exceptions to this pattern – in 44 villages SCs are population, but not land-dominant. There are also 39 villages where OBCs are population, but not land-dominant. We include separate terms to capture these two exceptions in our empirical analysis below. Table I shows that Upper caste dominance is the most common village regime, closely followed by villages dominated by OBCs. In comparison, the number of SC and Muslim dominated villages is relatively small.

Table II reports the distribution of panel households across village regimes and illustrates the extent to which panel households are clustered in 'own' dominated villages. Such clustering, which can be read off the bold diagonal, is pronounced for STs, UCs, OBCs and Muslims while the SC population is more dispersed. Relevant to the oppression hypothesis, table II also shows the presence of panel households from each social group in UC-dominated villages. 46 per cent of the panel households residing in such villages are UCs, 26.4 per cent SCs and 21.6 per cent OBCs. STs and Muslims between them account for 6 percent.

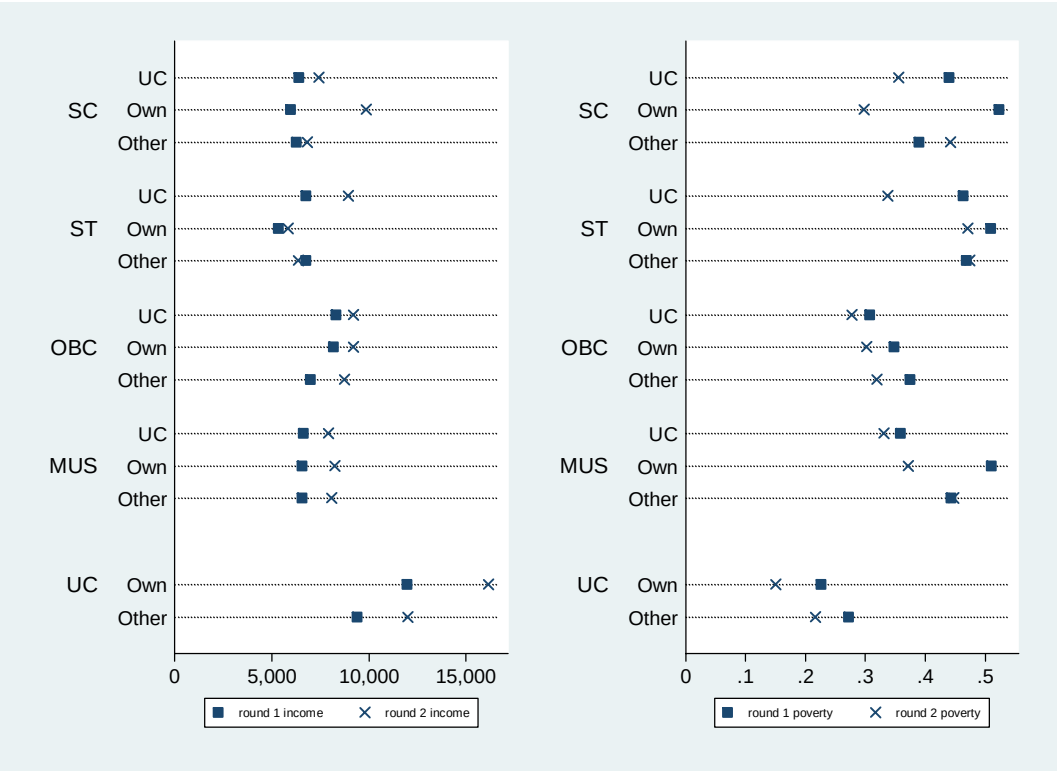
Table II: Number of panel households by social group and village regime

	Land dominant social group						
	SC	ST	OBC	MUS	UC	OTH	Total
Social group of panel households:							
Scheduled Castes (SC)	**222**	68	694	109	1,040	119	2,252
Scheduled Tribes (ST)	23	**552**	141	21	95	3	835
Other Backward Classes (OBC)	86	169	**1,608**	64	852	130	2,909
Muslims (MUS)	52	10	130	**337**	145	25	699
Upper Castes (UC)	44	61	381	29	**1,810**	91	2,416
Total	427	860	2,954	560	3,942	368	9,111

Source: HDPI panel, authors' calculations

Figure 1 reports round 1 and round 2 mean household per capita incomes and poverty headcount by social groups for villages with (i) upper caste land dominance, (ii) own group land dominance and (iii) the remaining 'other' villages.

Figure 1: Mean income and poverty headcount by social group, round and village regime



Source: HDPI panel, authors' calculations
Notes: Poverty is the share of the indicated sub-sample with income below the NSSO state-specific rural poverty lines.

Figure 1 suggests pronounced village regime effects on income levels, growth, poverty incidence and the speed of poverty reduction (persistence). In round 1, SCs and OBCs in upper caste dominated villages have marginally higher average incomes. For STs, round 1 incomes outside own enclaves were notably higher. [29] The average upper caste household was much better off in own enclaves, while Muslim incomes were roughly equivalent across village regimes. With the exception of UCs, the second round picture is strikingly different. STs appear to do much better in UC dominated villages while SCs fared much better in own enclaves. Muslims did marginally better in own enclaves. Contrasting this dynamism, SCs and OBCs in upper caste dominated villages and STs and OBCs in own enclaves made little progress.

[29] Kijima (2006;370) holds unfavourable geographic location responsible for disparities between STs and other social groups. Banerjee and Somanathan (2007) observe that districts with a high ST population have lower quality public goods and services such as roads, health and educational facilities.

In terms of average living standard improvements, enclaves appear to favour UCs and SCs very strongly and Muslims marginally; STs did remarkably well in UC dominated villages, but made little progress overall.

Were these average income changes confined to the better off within each social group or did they extend to poorer households as well? In the first round, the incidence of poverty among SCs, STs, OBCs and Muslims was lower in upper caste dominated villages than in own enclaves. Consistent with the income growth observations, the most dramatic poverty reductions appear for SCs in own enclaves and STs in upper caste dominated villages. However, in spite of modest income rises, poverty reduction among Muslims in own enclaves appears dramatic. Poorer ST households made slightly more progress than the average ST household. Consistent with the income figures, OBCs seem to have experienced limited poverty reduction between the two rounds.

The intricate ways in which social identity appears to interact with village regime, and the variations in welfare levels and changes by social group that this gives rise to, may not reflect oppression and enclave effects but instead be confounded by other factors. UC dominated villages may be clustered in areas with greater agricultural potential and SC dominated villages could be concentrated in states with more progressive policies towards Scheduled Castes or in states that experienced more (or less) income growth and poverty reduction after the 1991 reforms; we have already remarked upon the locational disadvantage of ST dominated villages.

Once potential locational confounds have been tackled, we are in a position to identify externalities associated with residence in upper caste dominated villages along with group specific oppression and enclave effects. With respect to the former, the quality of schools, health care and sanitation in UC dominated villages could all be expected to be better.[30] In addition, lower castes may emulate upper caste behaviour[31] which could strengthen educational aspirations and improve farming practices. Having rich neighbours can make it less risky to adopt high yielding seed varieties since one can learn from the good and bad experiences of wealthy early adopters (e.g. Foster and Rosenzweig 1995). Such positive externalities could exist alongside oppression effects manifested in limitations in the access to resources or

[30] Banerjee and Somanathan (2007) find that parliamentary constituencies with a concentration of Brahmins had higher level of schools and piped water in 1971.

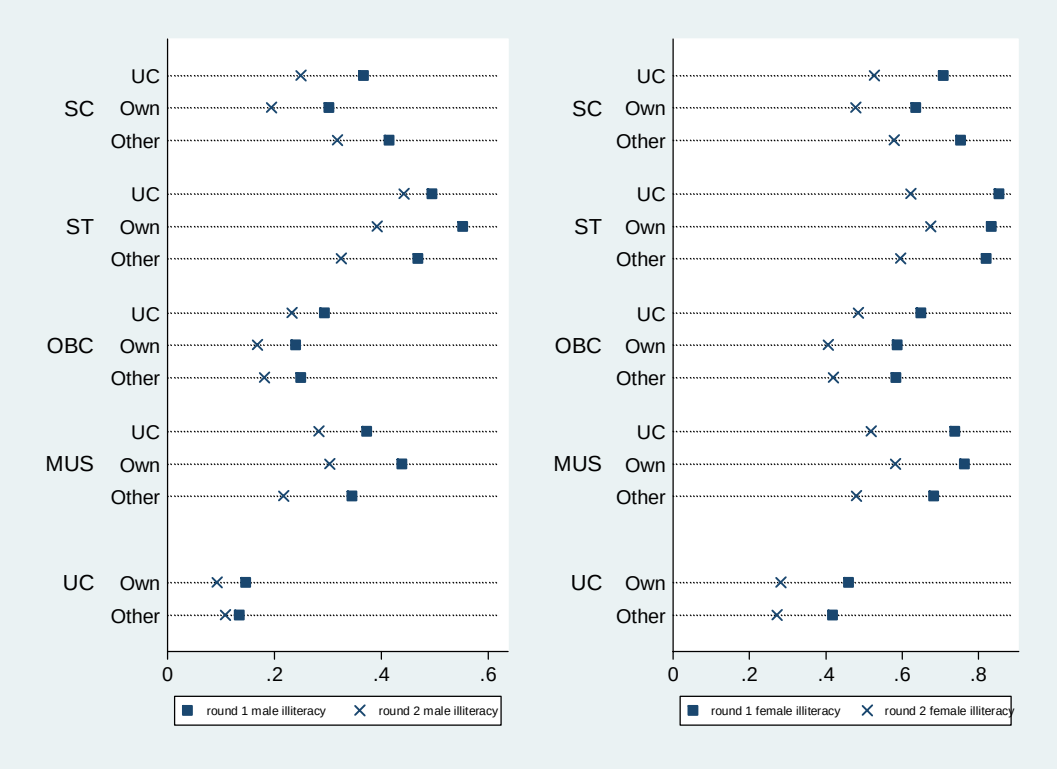
[31] One wellknown form of emulation, 'Sanskritisation', refers to a low caste changing its 'customs, ritual, ideology and way of life in the direction of a high, and frequently "twice-born' caste'. (Srinivas 1966;6)

markets, a hostile school environment, exclusion from membership in the local dairy cooperative or restrictions in the access to local credit or microcredit schemes that facilitate taking advantage of new opportunities in the post reform era.

By incrementally controlling for important factor endowments in our empirical model we obtain clues about the pathways through which each of these effects operate. A trivial possibility would be that SCs living in own enclaves do better because of larger or higher quality land endowments which directly or via more education could translate into higher incomes. If so, enclave coefficients should turn insignificant once education and household land holdings are controlled for.

Figures 2 and 3 illustrate how social identity interacted with village regime relate to two important factor endowments in rural India, namely basic education measured by male and female illiteracy and household land holdings.

Figure 2: Male and female illiteracy by social group, round and village regime



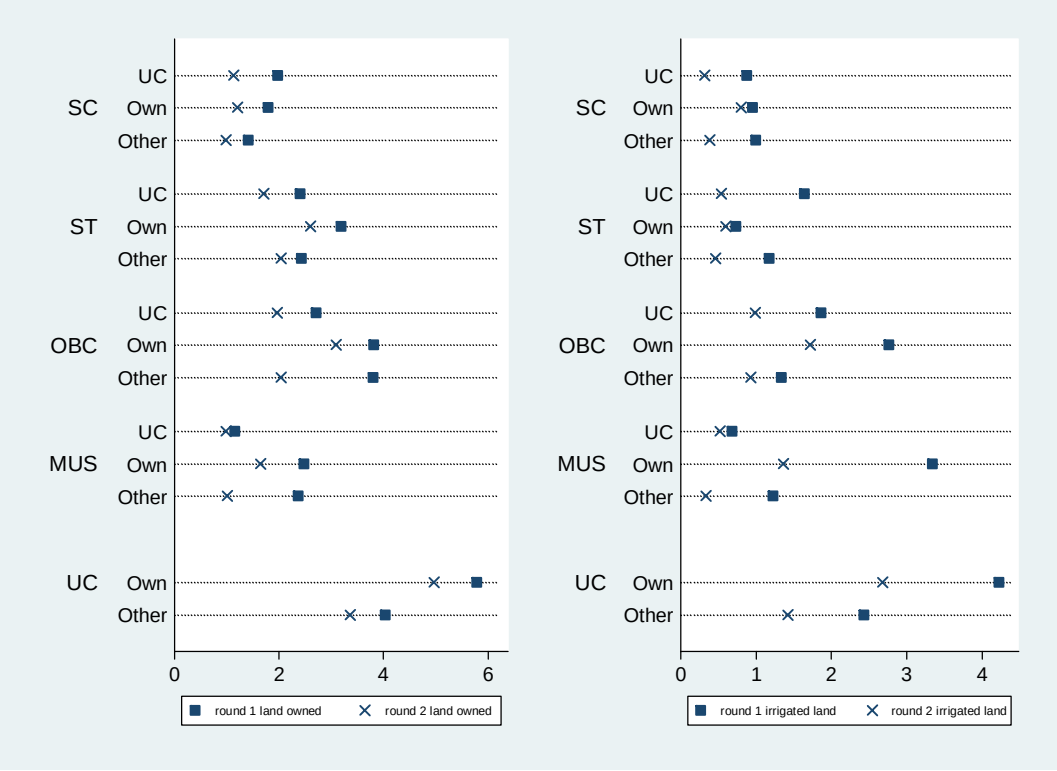
Source: HDPI panel, authors' calculations
Notes:

1. Figures are averaged across all *households* in the sub-sample indicated, and are based on the highest level of educational achievement in the household, i.e. on households of which not a single member is literate.

Among groups with low initial male literacy (SCs, STs, Muslims), there appears to have been across the board improvements with Muslims and STs in own enclaves

making more progress than those in UC dominated villages. SCs had higher and Muslims lower initial male literacy in their own enclaves. Although these observations on educational levels and progress correspond imperfectly with the income and growth patterns in figure 1, they do provide hints of positive enclave level and growth effects for SCs. While STs in UC dominated villages experienced rapid income growth, male education does not appear to be behind this spur. Female STs experienced dramatic educational progress in general, while female SCs did better and female Muslims worse in own enclaves.

Figure 3: Mean land holding (in acres) by social group, round and village regime



Source: HDPI panel, authors' calculations
Notes:

1. All figures are in acres and averaged across the entire sub-sample indicated, i.e. including those who do not own/hold any land.
2. Irrigated land includes owned and hired land (we cannot distinguish the two)

An a priori expectation of higher average round 1 land holdings in own dominated villages holds for UCs, STs and Muslims (marginally). The average SC household in UC dominated villages possessed more land than own enclave SC households only in round 1. OBCs in UC dominated villages owned less land than in own enclaves but more land than SCs in UC dominated villages. Consistent with expectations, the overall distribution of land holdings show UCs as the largest landowners followed by

OBCs, STs, Muslims (except in UC dominated villages) and SCs. Patterns are much the same, although with more pronounced differences, for irrigated land.[32]

While rich, our observations so far are inconclusive about the effects of village regimes on income, income growth and poverty incidence and persistence; our next step is to implement the empirical strategy laid out in section II.

IV EMPIRICAL RESULTS

Section A presents the estimation results, section B performs a robustness check on the main results and section C explores the implications of our results for the level of income and income growth, and for the incidence and persistence of poverty by social group and village regime.

A. Estimation results

Eq. (4) is estimated for round 1 and 2 using alternative specifications where extra variables are gradually introduced. The estimation results for these specifications are reported in Tables III and IV and in full in Appendix 2. The natural logarithm of real household income per capita is the dependent variable.

The first specification is a simple benchmark which contains 'raw' social identity dummy variables (SC, ST, MUS, OBC) with upper castes as benchmark category. We proceed by adding the enclave village regime variables capturing own group land dominance (SC x DSC, ST x DST, MUS x DMUS and OBC x DOBC), the two population (but not land) dominance dummy variables discussed in section III that we denote by SC x PSC and OBC x POBC, the dummy for demarcating villages with upper caste land dominance (DUC) and, finally, the oppression variables represented by the social group interaction terms with upper caste dominance. The latter facilitate identification of how SCs (SC x DUC), STs (ST x DUC), Muslims (MUS x DUC) and OBCs (OBC x DUC) perform within upper caste dominated villages compared to elsewhere.

Sets of related control variables are then gradually introduced: we think of agro-ecological zone indicators, state dummy variables and household composition

[32] The consistent decline in land holdings, given that we are dealing with a panel, may look puzzling. Further disaggregation confirms this trend across states, suggesting that this is a real phenomenon. The most likely explanation is that in a sufficient number of first round households to affect mean values, elderly patriarchs resided in joint households with the oldest son (and this son's wife and offspring). In the intermediate period, some of these patriarchs died – while the oldest son's household remained intact, its land holdings was split among the oldest son and his siblings.

variables[33] as 'pure' controls which are added in the specifications reported on in Table III; village infrastructure, household education and land variables we think of as possible pathways through which enclave and oppression effects operate; these are added in the specifications reported on in Table IV. The full details are available in Appendix 2.

[33] While household demography may affect well-being, demographic behaviour is unlikely to be affected by oppression and should be controlled for to identify true enclave and oppression effects.

Table III Estimation results of the effects on income of social identity, village regime and locational and demographic controls

Model:	Social identity terms (1)		Plus village regime (2)		Plus controls (3)	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
Social identity:						
HH is SC	-0.463***	-0.607***	-0.374***	-0.510***	-0.370***	-0.455***
	(-20.29)	(-24.75)	(-9.40)	(-11.91)	(-9.55)	(-11.18)
HH is ST	-0.458***	-0.601***	-0.310***	-0.462***	-0.309***	-0.374***
	(-12.80)	(-16.30)	(-4.57)	(-6.95)	(-4.92)	(-6.23)
HH is OBC	-0.267***	-0.371***	-0.291***	-0.338***	-0.216***	-0.264***
	(-11.40)	(-14.92)	(-5.87)	(-6.09)	(-4.59)	(-5.16)
HH is MUS	-0.406***	-0.530***	-0.289***	-0.446***	-0.201***	-0.324***
	(-10.81)	(-12.73)	(-4.76)	(-6.68)	(-3.35)	(-5.14)
Village regime variables:						
SC x DSC			0.032	0.268***	0.140**	0.302***
			(0.48)	(3.55)	(2.18)	(4.25)
SC x PSC			-0.031	0.062	-0.048	0.074
			(-0.59)	(1.08)	(-0.97)	(1.43)
ST x DST			-0.088	-0.056	-0.017	-0.001
			(-1.22)	(-0.78)	(-0.26)	(-0.01)
OBC x DOBC			0.170***	0.142***	0.098**	0.118***
			(3.58)	(2.73)	(2.21)	(2.56)
OBC x POBC			0.000	0.156**	-0.053	0.129**
			(0.01)	(2.30)	(-0.91)	(2.13)
MUS x DMUS			-0.023	0.160*	-0.015	0.122
			(-0.28)	(1.85)	(-0.20)	(1.58)
DUC			0.202***	0.285***	0.109***	0.107***
			(4.20)	(5.67)	(2.32)	(2.32)
SC x DUC			-0.113***	-0.149***	-0.082*	-0.135***
			(-2.30)	(-2.83)	(-1.75)	(-2.72)
ST x DUC			-0.222**	-0.030	-0.132	-0.007
			(-2.02)	(-0.27)	(-1.27)	(-0.07)
MUS x DUC			-0.187**	-0.142	-0.147*	-0.079
			(-2.11)	(-1.41)	(-1.75)	(-0.84)
OBC x DUC			-0.024	-0.127**	-0.078	-0.116*
			(-0.41)	(-1.99)	(-1.41)	(-1.95)
Controls:						
Household composition	No	No	No	No	Yes	Yes
Agro-ecological zones	No	No	No	No	Yes	Yes
State dummy variables	No	No	No	No	Yes	Yes
R squared (overall)	0.0609	0.0949	0.0663	0.1065	0.2127	0.2837
N	9111	9111	9111	9111	9111	9111

Source: HDPI-I ("round 1") and II ("round 2") surveys, panel households only; authors' calculations.
Notes: Dependent variable is the natural logarithm of annual per capita household income in constant 1993/94 prices, with round 2 figures converted using NSSO state-specific rural CPIs. Random effects, with standard errors that are robust to heteroske-dasticity and clustering within villages. ***, ** and * denote significance at 1, 5 and 10 percent respectively; robust t-statistics are in parentheses. Demographic controls are the sex of the household head, number of boys aged 0-5, girls 0-5, boys 6-14, males 15-19, females 15-19, males 20-24, females 20-24, males 25-49, females 25-49, males 50-59, females 50-59, males 60 and older, and females 60 and older. See table A2.1 in Appendix 2 for the full specification.

Table III is laid out to facilitate round 1 and round 2 comparisons. We first report broad patterns of identity-based disparities highlighting changes between round 1 and 2. We then briefly consider the enclave and oppression coefficients before and after introducing locational controls. We finally focus on the latter enclave and oppression effects and the pathways through which they operate.

Column 1 presents the relative magnitude of the raw social identity coefficients for SC, ST, MUS and OBC households, which are all significant at the 1 percent level. In both rounds, SCs and STs are, on average, the relatively most disadvantaged, having incomes compared to UCs that are about 46 percent lower in round 1 and 60 percent lower in round 2. Muslims are slightly better off with, on average, a 41 percent lower income in round 1 and a 53 percent lower income in round 2 than UCs, while OBCs, on average, are well ahead of the other three groups. On average, OBCs have a 27 percent lower income in round 1 and a 37 percent lower income in round 2 than UCs. Upper caste households are, in general, and in line with a priori expectations, much better off than everyone else. The raw coefficients also suggest that the disparity between upper castes and each of the other social groups widened between the two rounds.

In columns (2) and (3) the village regime variables interacted with households' social group are introduced, first without and then with control variables added. Prior to adding agroecological, state and household demographic controls, it would appear that residing in an upper caste dominated village not only is advantageous for upper caste households but bestows sizeable positive externalities on other social groups. The coefficient corresponding to DUC, which is statistically significant at the 1 per cent level, leaps notably in size between the two rounds. We also observe negative and statistically significant round 1 interaction terms ('oppression coefficients') for SCs, STs and Muslims. Further, the oppression effect disappears for STs and Muslims and becomes significant for OBCs in round 2.

For SCs we find a large and strongly significant positive enclave effect in round 2. The round 2 enclave effect for Muslims is weaker. In addition, significant enclave coefficients for OBCs appear in both rounds.

The responses of the 'raw' identity and village regime coefficients to the step-wise introduction of each of the three sets of 'pure' controls may be gauged in full in table A2.1 in Appendix 2. Adding Palmer-Jones and Sen's (2003) agroecological zone indicators shows that the positive externality associated with residing in an upper

caste dominated village does have a locational dimension. While remaining strongly significant, coefficient sizes are sharply reduced. Introducing state dummy variables further reduces size and eliminates the gap between round 1 and 2.

At the same time, and with one notable exception, both the size and statistical significance of the oppression effects remain intact; the coefficient for STs in UC dominated villages drops out. It is also clear that the key enclave coefficients are not locationally confounded. OBCs do better in their own enclaves in both rounds, while SCs do much better in their own enclaves in both rounds but especially in round 2. However, the weaker round 2 enclave effect for Muslims turns insignificant.

The precise implications of the positive externality and of the oppression and enclave effects for income levels, growth and poverty incidence and persistence are brought out in the computations and discussions of counterfactual income, growth and poverty in subsection C below.

These, our main results, may be qualitatively summarised as follows. UCs earn higher incomes than others in both rounds. In addition, UCs in own dominated villages perform better than other UCs. There is, moreover, a general and strong positive externality associated with residence in upper caste dominated villages. The coefficient on DUC is significant at the 1 percent level and of similar size in round 1 and 2. For the oppression coefficients, we observe that while Muslims and SCs did worse in UC dominated villages in round 1, SC disadvantage within such villages intensified while Muslims within such villages progressed in the post reform years. OBCs in UC dominated villages were also at a disadvantage in round 2, but less so than SCs. Further, OBCs and, in particular, SCs do far better in own dominated villages in both rounds, but with the enclave effect in SC-dominated villages leaping dramatically in the post reform era. Overall, our results suggest more pronounced disadvantage for SCs and OBCs in upper caste dominated villages post reform and stronger enclave effects for SCs in round 2. One possible explanation is that it is in enclaves, where discrimination by powerful groups is less likely, that marginalised groups faced fewer obstacles in the access to markets and that such access gained in importance in the post reform years.

We next focus on the pathways through which oppression and enclave effects operate and of possible change in the post reform era. We gradually control for village infrastructure, for household education and household land holdings with results reported in full in table IV.

Table IV Estimation results of the effects on income of social identity, village regime and additional controls: village infrastructure, household education and land

Controls added:	Village infrastructure		Plus education (hh)		Plus land (hh)	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
Social identity:						
HH is SC	-0.369***	-0.452***	-0.245***	-0.321***	-0.154***	-0.252***
	(-9.56)	(-11.07)	(-6.58)	(-8.23)	(-4.49)	(-6.81)
HH is ST	-0.311***	-0.374***	-0.172***	-0.234***	-0.152***	-0.203***
	(-4.95)	(-6.17)	(-2.81)	(-4.09)	(-2.72)	(-3.64)
HH is OBC	-0.21***	-0.259***	-0.141***	-0.181***	-0.129***	-0.148***
	(-4.45)	(-5.05)	(-3.08)	(-3.70)	(-3.12)	(-3.25)
HH is MUS	-0.208***	-0.324***	-0.113**	-0.217***	-0.065	-0.140***
	(-3.48)	(-5.12)	(-1.98)	(-3.59)	(1.23)	(-2.38)
Village regime:						
SC x DSC	0.131**	0.289***	0.123**	0.253***	0.06	0.205***
	(2.05)	(4.03)	(2.01)	(3.64)	(1.05)	(3.16)
SC x PSC	-0.035	0.076	-0.053	0.07	-0.062	0.059
	(-0.72)	(1.44)	(1.08)	(1.39)	(1.34)	(1.23)
ST x DST	0.008	0.019	-0.006	0.005	0.012	-0.012
	(0.12)	(0.31)	(0.10)	(0.08)	(0.22)	(0.22)
OBC x DOBC	0.095**	0.115***	0.094**	0.097**	0.067*	0.05
	(2.15)	(2.49)	(2.16)	(2.19)	(1.73)	(1.21)
OBC x POBC	-0.059	0.134**	-0.044	0.118*	-0.012	0.106**
	(-1.00)	(2.20)	(-0.77)	(1.98)	(-0.23)	(1.98)
MUS x DMUS	0.003	0.147*	0.006	0.15**	-0.048	0.087
	(0.04)	(1.88)	(0.08)	(2.01)	(-0.73)	(1.21)
DUC	0.116***	0.107**	0.124***	0.105***	0.035	0.024
	(2.48)	(2.30)	(2.76)	(2.36)	(0.89)	(0.57)
SC x DUC	-0.09*	-0.14***	-0.107***	-0.15***	0.01	-0.045
	(-1.92)	(-2.81)	(-2.39)	(-3.17)	(0.24)	(-1.01)
ST x DUC	-0.14	-0.01	-0.129	0.017	-0.032	0.112
	(-1.35)	(-0.10)	(-1.27)	(0.18)	(-0.36)	(1.26)
MUS x DUC	-0.15*	-0.094	-0.129	-0.054	0.026	0.003
	(-1.79)	(-0.99)	(-1.58)	(-0.58)	(0.34)	(0.03)
OBC x DUC	-0.091	-0.123**	-0.085	-0.113**	0.076	0.001
	(-1.64)	(-2.07)	(-1.59)	(-1.99)	(1.59)	(0.01)
Controls:						
Household composition	Yes	Yes	Yes	Yes	Yes	Yes
Agro-ecological zones	Yes	Yes	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes	Yes	Yes
Village infrastructure	Yes	Yes	Yes	Yes	Yes	Yes
Household education	No	No	Yes	Yes	Yes	Yes
Household land	No	No	No	No	Yes	Yes
R squared (overall)	0.2252	0.2877	0.2700	0.3413	0.4258	0.4181
N	9111	9111	9111	9111	9111	9111

Source and Notes: as for Table IV

Additional notes: Education variables are dummy variables used as controls for the highest level of male and female education in the household. Land refers to controls for the logarithm of owned household land measured in acres, and the logarithm of irrigated household land measured in acres. Village size is captured by village population (logarithm). The village infrastructure controls are the presence within the village of a busstop (1), or within its vicinity of a railway station (2), medical clinic (3), schools, and if so, at which level of education (4), or a market/mandi (5), as well as the type of road (footpath only, kutcha road, pucca road) that leads to the village (6). The full specification is reported in table A2.2 in Appendix 2.

Adding the village infrastructure controls detailed in the note to table IV has a close to negligible effect both on the raw identity and village regime coefficients. On the face of it and contrary to received wisdom (e.g Binstrup Andersen and Shimokawa 2006), the scope for reducing identity based disadvantage by improving village infrastructure appears more limited than expected. It is, however, quite possible that upper caste land dominance is correlated with better quality village infrastructure and that one reason for why village infrastructure variables perform so badly elsewhere reflects such a quality difference.

In line with Kijima (2006) and to shed light on the role of human capital endowment differences, we introduce dummies for the maximum female and male education within a household where the educational categories are up to primary, middle, matriculation, higher secondary and graduate plus. A hypothesis resonating with Dercon and Krishnan's (2007) findings would be that social identity disparities – by caste or religion – should be wiped out once educational attainments are controlled for. For both rounds, we note the marked reductions in the raw identity coefficients and thus in the relative disadvantage of SCs, STs, Muslims and OBCs from adding educational controls. For STs, the raw coefficient drops from -0.31 to -0.17 or by around 45 per cent. For SCs, in comparison, education nets out about 33 percent of the remaining disadvantage vis-à-vis upper caste households. Our results concur with Dercon and Krishnan (2007) in suggesting that education is crucial: at the same time, our findings suggest that education is only part of the story.

Turning to the oppression and enclave effects, controlling for household education turns the weak round 1 oppression coefficient for Muslims insignificant, while the oppression coefficient for SCs is notably reinforced. We next consider land holdings as potential oppression buffer and asset contributing to enclave advantage. In contrast to Dercon and Krishnan (2007), land appears to hold the key to eliminating oppression associated with upper caste dominance. Once household land (the natural logarithm of the acres of land owned by the household, and all the household's irrigated land) is controlled for, the positive village externality and all identity specific oppression effects are wiped out in both rounds. The interpretation is as follows: the positive externality from residence in upper caste villages essentially accrues to landholding households. Further, what one may think of as the 'traditional' and *additional* burden imposed on SC and OBC households from residing in upper caste

dominated villages could, both past and post reform, be eliminated through land redistribution. Land reform would thus wipe out the *separate* effect on income of upper caste oppression, leaving no residual effect of such oppression in any other markets or transactions. Crucially, however, this does not imply that all identity based rural disadvantage will be eliminated, since the raw coefficients, with the exception of Muslims in round 1, remain stubborn and statistically significant. In other words, even after location, demography, village infrastructure and factor endowments are carefully controlled for, the raw coefficients suggest that SCs with a similar resource base and attributes as others not only remain the worst off but have fallen further behind STs and OBCs in the post reform years. One important exception is SCs residing in own enclaves; the SC enclave coefficient remains large and strongly significant after land holdings and all other controls are added and is large enough to eliminate 80 percent of the remaining disadvantage vis-à-vis UC households. Notice that Muslims have also experienced a post-reform setback.

B. Robustness tests

As discussed in section II, we conduct two robustness tests on our main results to respond to potential reservations about the sociological foundation of our main dominance measure. The first robustness test replaces the dummy variables for upper caste and own group land dominance with the share of village land owned by the dominant group, since power may depend on the concentration of land ownership and thus how much land the dominant group owns. The second robustness test replaces the dummy variables for land dominance with fragmentation adjusted dominance measure defined by equation (3), since not only the share of land owned by the dominant group but also the extent to which the power base represented by the land ownership of other groups is fragmented could matter. Table V reports the sign and the level of significance on the estimated parameters of interest to the oppression and enclave hypothesis in the specification with 'pure' control variables only (AEZs, state dummy variables and household demographic controls).[34]

[34] We also ran regressions using alternative dominance measures for each specification reported on in Tables III and IV and Appendix 2, and obtained very similar results in terms of broad comparability with specifications using our main dominance measure to those reported here. These additional results are all available on request.

The round 1 results for these alternative specifications are in the top half of Table V and the round 2 results in the bottom half. 17 out of the 22 coefficients (11 per round) on the village regime variables when using the land dominance dummy are robust in terms of retaining sign and statistical significance (or insignificance, as the case may be) regardless of which alternative dominance measure is used.[35] It is also noteworthy that significance of coefficients is generally stronger for the more refined dominance measures, especially for those capturing oppression.

Exploiting the panel dimension of our dataset, we also investigated whether changes in coefficients on village regime variables between rounds were statistically significant, for each of the specifications reported in Tables III and IV and Appendix 2, and then each time for each of our three dominance measures, in regressions of the change in the natural logarithm of real per capita income on these variables and the appropriate set of controls. The intensifying enclave effect for SCs is statistically significant, for each dominance measure, and that for OBCs and Muslims only when we use the more refined measures. Changes in oppression coefficients are generally not significant which is consistent with our main results for STs in UC dominated villages and SCs in UC dominated villages. For the latter, non-significance in the growth regression is the equivalent of no change in the level oppression coefficient between the two rounds.

[35] The five coefficients that do not retain either (in)significance or their sign are ST x DST in round 2, OBC x DOBC in round 1, MUS x DMUS in round 2, ST x DUC in round 1 and OBC x DUC in round 1.

Table V: Qualitative summary of robustness tests

Main dominance measure:	Main dominance measure results	Land percentage of largest land holding group in village	Dominance-adjusted Herfindahl index (eq. 3)
Round 1			
SC x DSC	++	+++	++
SC x PSC	Ns	Ns	Ns
ST x DST	Ns	Ns	Ns
OBC x DOBC	++	Ns	Ns
OBC x POBC	Ns	Ns	Ns
MUS x DMUS	Ns	Ns	Ns
DUC	+++	+++	+++
SC x DUC	-	---	---
ST x DUC	Ns	-	-
MUS x DUC	-	---	--
OBC x DUC	Ns	---	---
Round 2			
SC x DSC	+++	+++	+++
SC x PSC	Ns	Ns	Ns
ST x DST	Ns	Ns	--
OBC x DOBC	+++	+++	+++
OBC x POBC	++	++	+
MUS x DMUS	Ns	+	+
DUC	+++	+++	+++
SC x DUC	---	---	---
ST x DUC	Ns	Ns	Ns
MUS x DUC	Ns	Ns	Ns
OBC x DUC	--	-	---

Notes: +++, ++, + indicates positive coefficient significant at 1, 5 and 10% respectively, ---, --, - indicates negative coefficient significant at 1, 5 and 10% respectively, Ns indicates not significant, all in the specification with social group, village regime, agro-ecological zones, state dummies and household demographic composition variables.

C. Magnitude of enclave and oppression effects

We next explore the order of magnitude of the enclave and oppression effects in terms of income, income growth, and the incidence and persistence of poverty. We do so by computing counterfactual income as if the significant coefficients on the social identity times village regime variables were equal to zero and use the coefficients from our model that includes AEZ, state dummies and household demographic controls. The enclave and oppression effects quantified here may be interpreted as aggregate effects, summed across all markets (and public services). We restrict the analysis to statistically significant enclave and oppression effects.

Table VI: Actual and counterfactual income, growth and poverty without village regime effects by social group

	Scheduled Castes		**Scheduled Tribes**		**Other Backward Classes**		**Muslims**	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
Upper-caste dominated villages								
Mean income per capita								
Actual	6,395	7,391	6,760	8,905	8,309	9,200	6,626	7,915
Counterfactual – without general village regime effect	5,735	6,641	6,062	8,002	7,451	8,266	5,942	7,112
Counterfactual – without oppression effect	6,941	8,460	6,760	8,905	8,309	10,331	7,675	7,915
Counterfactual – without general village regime and oppression effects	6,225	7,601	6,062	8,002	7,451	9,283	6,883	7,112
Growth in mean income per capita (% per year between 1994 and 2005)								
Actual	-	1.3	-	2.5	-	0.9	-	1.6
Counterfactual – without general village regime effect	-	1.3	-	2.6	-	0.9	-	1.6
Counterfactual – without oppression effect	-	1.8	-	2.5	-	2.0	-	0.3
Counterfactual – without general village regime and oppression effects	-	1.8	-	2.6	-	2.0	-	0.3
Poverty headcount (%)								
Actual	43.9	35.4	46.3	33.7	30.8	27.7	35.9	33.1
Counterfactual – without general village regime effect	49.9	42.0	51.6	41.1	35.8	32.4	44.8	40.0
Counterfactual – without oppression effect	38.7	29.9	46.3	33.7	30.8	22.9	29.0	33.1
Counterfactual – without general village regime and oppression effects	45.7	34.3	51.6	41.1	35.8	27.5	33.1	40.0
Own-group dominated villages								
Mean income per capita								
Actual	5,954	9,842	5,331	5,805	8,158	9,187	6,553	8,231
Counterfactual – without enclave effect	5,176	7,276	5,331	5,805	7,397	8,164	6,553	8,231
Growth in mean income per capita (% per year between 1994 and 2005)								
Actual	-	4.7	-	0.8	-	1.1	-	2.1
Counterfactual – without enclave effect	-	3.1	-	0.8	-	0.9	-	2.1
Poverty headcount (%)								
Actual	52.3	29.7	50.9	47.1	34.8	30.2	51.0	37.1
Counterfactual – without enclave effect	59.9	47.7	50.9	47.1	39.9	37.7	51.0	37.1

Notes: counterfactual figures are all based on counterfactual income computed for each household in villages land dominated by indicated group, using significant coefficients from the round 1 and round 2 regressions of the natural logarithm of income on village regime and social identity variables, controlling for agro-ecological zones, state dummies, and household demographic characteristics, as reported in the last column of Table 3 and in full in Appendix 2.

For round 1 and 2 income per capita and poverty, and annual income growth between the two rounds, Table VI reports, by marginalised group, actual and counterfactual

figures, separately for upper-caste dominated villages and for own-group dominated villages. For the last-mentioned villages, counterfactual figures are based on what these variables would have been without the estimated enclave effect. For the first-mentioned villages, three sets of counterfactual figures are reported. First, income, growth and poverty are computed as if there is no general village regime effect (the coefficient on DUC); next as if there is no group specific oppression effect (e.g. the coefficient on SC x DUC); and finally as if there is neither effect. So, for example, mean income per capita in round 1 for SCs living in UC-dominated villages is equal to 6,359 Rupees. Had they not benefited from the general village regime effect, it would have been 5,735 Rupees; had they not suffered from oppression, it would have been 6,941 Rupees; and if neither effect had been at work, it would have been 6,225 Rupees. The last figure is lower than their actual mean income, which shows that, in this case, the positive village regime effect is larger (in absolute terms) than the negative oppression effect.

The general village regime effect on income of marginalised groups living in UC-dominated villages is always about 10 percent, both in round 1 and in round 2: mean income would thus have been some 10 percent lower were it not for this effect. Since the effect on income is approximately the same size in both rounds, the effect on growth is negligible. The effect on the headcount percentage of poverty, on the other hand, depends on the group specific distribution of income in the vicinity of the poverty line. Muslims in round 1 benefited most and OBCs in round 2 least: poverty would have been 8.9 percentage points higher for the former and 4.7 percentage points higher for the latter, were it not for the general village regime effect.

The group specific oppression effect on income of living in UC-dominated villages, when statistically significant, tends to be larger than the general village regime effect, with one exception (SCs in round 1). Income in such villages would have been 14.4 percent higher for SCs in round 2, 12.3 percent higher for OBCs in round 2, and 15.8 percent higher for Muslims in round 1. The effect on growth is pronounced, too. SCs would have experienced 1.8 instead of 1.3 percent annual growth (22 percent over the entire period instead of 15 percent) and OBCs 2.0 instead of 0.9 percent (24 instead of 10 percent), were it not for oppression. Although the oppression effect dominates the general village regime effect for income, this is not always the case for poverty, which must be related to peculiarities of the PDF of income. It is worth noting, though, that poverty reduction would have been about 8

instead of 3 percentage points for OBCs, were it not for oppression – SCs would have experienced about the same amount of poverty reduction as they experienced actually, because the level effect in both rounds was of the same order of magnitude.

Enclave effects in the specification used are significant only for SCs and OBCs, in both rounds. For OBCs they are of the same order of magnitude (but positive) as the oppression effects remarked on above for this group. For SCs they are much larger. Income per capita would have been 13.1 percent lower in round 1, and 26.1 percent lower in round 2, annual growth 1.6 percentage points lower (20 percent less growth over the period), and poverty 7.6 and 18 percentage points lower in round 1 and round 2, were it not for the enclave effect. Poverty would have been far more persistent for SCs in own-dominated villages in the absence of this effect.

In summary, we find sizeable general village regime effects that benefit those residing in UC-dominated villages for income and poverty (but not for growth and poverty reduction), and for SCs in both rounds, for OBCs in round 2 and for Muslims in round 1 an offsetting oppression effect of the same order of magnitude, although typically larger in absolute terms than the general village regime effect in the case of income and slightly smaller in the case of poverty. Growth for SCs and OBCs is substantially negatively affected by oppression. Enclave effects are large and positive for OBCs and especially SCs in terms of income and the absence of poverty, and for SCs in terms of growth, too.

V CONCLUDING REMARKS

Using a unique panel data set for rural India covering the years 1993/94 and 2004/05, we tested the hypothesis that disadvantaged groups (Scheduled Castes, Scheduled Tribes, Muslims and OBCs) fare worse in terms of income levels when residing in villages dominated by upper castes and whether the same groups fare better or worse in own dominated villages. Our results provide strong support for the 'oppression' hypothesis and for the positive enclave hypothesis. In addition, and for all social groups, a considerable positive externality from residing in upper caste dominated villages was uncovered.

The quantitative effects on income levels, growth, poverty incidence and poverty persistence were discerned. The income levels of SCs living in upper caste dominated villages would have been 8.5 percent higher in round 1 and 14.4 percent higher in

round 2 in the absence of oppression effects, while annual income growth would have been 0.5 percentage points higher, 1.8 instead of 1.3 percent. Further, the poverty incidence would have been more than 5 percentage points lower.

The negative 'enclave' hypothesis advances the view that the slow progress of marginalised communities, rather than being attributable to external oppression, may be caused by factors internal to the community itself. In tune with Anderson's (2005) findings for Yadavs in Bihar and Uttar Pradesh, but in our case extending to marginalised groups below the pollution barrier, Scheduled Caste households in own dominated villages experienced much more rapid poverty reduction between the two rounds. In round 2, their income is 26.1 percent higher than it would have been without the enclave effect, and their incidence of poverty 18 percentage points lower, 29.7 instead of 47.7 percent.

While our enclave results contrast with Munshi and Rosenzweig's (2006) findings from Mumbai, the magnitude of the raw social identity coefficients highlight the persistence of identity based disparities in rural India; whereas upper caste 'oppression' has contributed substantially to prolong poverty and to low income among SC and OBC households there are, at the same time, significant positive externalities associated with living in upper caste dominated villages.

We also shed new light on the routes through which the welfare disparities between different social groups within and outside villages dominated by upper castes may be closed. Educational attainment, while generally important and emphasised by other studies, matters outside UC dominated villages. Overall, however, the old recipe of land redistribution holds the key to neutralising disparities attributable to upper caste dominance. This is in notable contrast to Dercon and Krishnan's (2007) finding using the ICRISAT-data set which concluded that caste based disadvantages essentially have educational origins.

Our findings go further and show that once all factor endowment differences are controlled for, the round 1 gap between SCs, STs and OBCs dramatically narrows. However, even after location, demography, village infrastructure and factor endowments are carefully controlled for, the raw coefficients suggest that SCs with a similar resource base and attributes as others not only remained the worst off but fell further behind STs and OBCs in the post reform years. The latter provides a useful empirical corrective to accounts suggestive of sustained SC progress relative to other

groups.[36] The notable exception is SCs residing in own enclaves; the SC enclave coefficient remains large and strongly significant after land holdings and all other controls are added. Notice that Muslims also experienced a post-reform setback.

REFERENCES

Akerlof, G and R. Kranton (2000): 'Economics and Identity', *Quarterly Journal of Economics,* 115 (3): 715-53.

Anderson, S. (2005): Caste as an Impediment to Trade, mimeo, Department of Economics, University of British Columbia.

Banerjee, A. and R. Somanathan (2007): 'The Political Economy of Public Goods: Some Evidence from India,' *Journal of Development Economics*, 87; 287-314.

Banerjee, Biswajit and John B. Knight (1985): 'Caste discrimination in the Indian labour market', *Journal of Development Economics*, 17, 277-307.

Baulch, B. and J. Hoddinott (2000): 'Economic Mobility and Poverty Dynamics in Developing Countries', *Journal of Development Studies* 36(6), 1-24.

Bayly, S (1999): Caste, Society and Politics in India – From the Eighteenth Century to the Modern Age, *Cambridge University Press*.

Binstrup-Andersen, P. and S. Shimokawa (2006): Rural Infrastructure and Agricultural \ Development, Paper prepared for presentation at the Annual Bank Conference on Development Economics, Tokyo, Japan, May 29-30, 2006.

Borooah, V. K. (2005): 'Caste, Inequality and Poverty in India', *Review of Development Economics*, 9(3): 399-414.

Cameron, A.C., and P.K.Trivedi (2005): Microeconometrics, Methods and Applications. New York: *Cambridge University Press*.

Chowdhry, P. (2007): Contentious Marriages, Eloping Couples: Gender, Caste and Patriarchy in Northern India, *Oxford University Press*.

Cortes, K. E. (2006): 'The effects of age of arrival and enclave schools on the academic performance of immigrant children', *Economics of Education Review*, 25: 121-32.

Das, M. B. (2007): Minority Status and Labour Market Outcomes: Testing a 'Minority Enclave' Hypothesis, Presented at the National Conference on Social Exclusion and Inclusive Policies, New Delhi, 26-27 October 2007.

Das, M. B. and P. V. Dutta (2007): Does Caste Matter for Wages in the Indian Labor Market?, mimeo, World Bank, New Delhi.

[36] Banerjee and Somanathan's (2007;308) findings on public good provision in 1991.

Dercon, S. and P. Krishnan (2007): Caste revisited: Consumption Dynamics and Exclusion in South-Indian villages 1975-2005, mimeo, Department of Economics, University of Oxford.

Desai, S., A. Dubey, B.L. Joshi, M. Sen, A. Shariff, and R. Vanneman (2009), 'India Human Development Survey: Design and Data Quality – Technical Paper No. 1,' NCAER and University of Maryland. http://ihds.umd.edu/IHDS_papers/technical%20paper%201.pdf

Deshpande, A. (2001): 'Caste at Birth? Redefining Disparity in India', *Review of Development Economics*, 5(1): 130-44.

Deshpande, A and K. Newman (2007): 'Where the Path Leads – The Role of Caste in Post-University Employment Expectations', *Economic and Political Weekly*, 13 October, 4122-4140.

Dumont, L (1970): Homo hierarchicus. The caste system and its implications, *Paladin Press*, London.

Dutta, P. V. (2006): 'Returns to Education: New Evidence for India, 1983-1999', *Education Economics,*14(4): 435-55.

Edin, P. A., P. Fredriksson and O. Aaslund (2003): 'Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment', *Quarterly Journal of Economics*, 329-357.

Epstein, T. S., A.P. Suryanarayana and T. Thimmegowda (1998): Village Voices – 40 Years of Rural Transformation in South India, *Sage Publications.*

Esteban, J. and D. Ray (2009): Linking Conflict to Inequality and Polarization, mimeo, Department of Economics, New York University.

Foster, A and M.R, Rosenzweig (1995): 'Learning by doing and learning from others: human capital and technical change in agriculture', *Journal of Political Economy*, 103: 1176-1209

Fuller, C. J. (1996): Caste Today, *Oxford University Press.*

Gang, I. N, K. Sen and M. Yun (2008a): 'Poverty in Rural India: Caste and Tribe', *Review of Income and Wealth*, 54 (1): 50-70.

Gang, I. N., K. Sen and M.Yun (2008b): Was the Mandal Commission Right? Living Standard Differences between Backward Classes and Other Social Groups in India," IZA Discussion Papers 3453, Institute for the Study of Labor (IZA).

Guha, R. (2007): India after Gandhi: The History of the World's Largest Democracy, *McMillan*.

Hatlebakk, M. (2009): 'Capacity-constrained Collusive Price Discrimination in the Informal Rural Credit Markets of Nepal,' *Review of Development Economics*, 13 (1), 70-86.

Iversen, V. and Raghavendra, P.S. (2006): 'What the signboard hides: Food, caste and employability in small, South-Indian eating places', *Contributions to Indian Sociology*, 41(3), 311-41.

Jeffrey, C. (2001): 'A Fist is Stronger Than Five Fingers: Caste and Dominance in rural North-India', *Transactions of the Institute of British Geographers,* New Series, 26 (2): 217-36.

Jeffrey, C., P. Jeffery and R. Jeffery (2007): Degrees without Freedom: Education, Masculinities and Unemployment in North-India, *Stanford University Press*.

Jodhka, S. S. (2004): 'Sikhism and the caste question: Dalits and their politics in contemporary Punjab', *Contributions to Indian Sociology*, 38 (1-2): 165-192.

Kijima, Y. (2006): 'Caste and Tribe Inequality: Evidence from India, 1983-1999', *Economic Development and Cultural Change*, 54(2): 369-404.

Kingdon, G. K. (1998): 'Does the Labour Market Explain Lower Female Schooling in India?', *Journal of Development Studies*, 35, No. 1: 39-65.

Kingdon, G. K. (2002): 'The gender gap in educational attainment in India: How much can be explained?', *Journal of Development Studies*, 39, No. 2: 25-53.

Madeshwaran, S. and P. Attewell (2007): 'Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey', *Economic and Political Weekly*, 13 October, 4146-4153.

Madsen, Stig Toft (1991): Udupi hotels: A case of ethnicity based Non-Virulent Development, Paper presented at the First Nordic Conference on South-Asian studies, 11-13 October, Nordic Association for South-Asian Studies.

Menon, D. M. (1994): Caste, nationalism and communism in South-India – Malabar 1900-1948, *Cambridge University Press*.

Mosse, D. (2003): The Rule of Water: Statecraft, Ecology and Collective Action in South-India, *Oxford University Press.*

Munshi, Kaivan and Mark R. Rosenzweig (2006): 'Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy', *American Economic Review*, 96(4): 1225-51.

Nambissan, Geetha B. (1996): Equity in Education? The Schooling of Dalit Children in India, Discussion Paper Series no. 15, Studies of Human Development in India, Project of the United Nations Human Development Programme.

Olson, M. (1971): The Logic of Collective Action – Public Goods and the Theory of Groups, *Harvard University Press*.

Palmer-Jones, R. and K. Sen (2003), 'What Has Luck Got To Do With It? A Regional Analysis of Poverty and Agricultural Growth in Rural India', *Journal of Development Studies* 40, 1-31.

Panagariya, A. (2008): India: The Emerging Giant, *Oxford University Press*.

Parikh, S. (1998): Affirmative action, Caste and Party Politics in Contemporary India, in Skrentny, J. D. (ed): Color Lines – Affirmative Action, Immigration and Civil Rights Options for America, *University of Chicago Press*.

Peyton Young, (2001): The dynamics of conformity, in S. D. Durlauf and H. Peyton Young (eds): Social dynamics, *The MIT Press*.

Rosenzweig, M. R. (2003), 'Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility,' *American Economic Review* 93(2), 112-117.

Sen, K. (2007): 'Why did the elephant start to trot? India's growth acceleration re-examined', *Economic and Political Weekly,* October 27-November 2, Vol. 43, pp. 37-49.

Shah, Ghanshyam, Harsh Mander, Sukhadeo Thorat, Satish Deshpande and Amita Baviskar (2006): Untouchability in rural India, *Sage Publications*, New Delhi.

Singh, H. (1984): Caste among non-Hindus in India, *National Publishing House*.

Singh, K. S. (1992): People of India – an Introduction, Anthropological Survey of India.

Srinivas, M. N. (1955): The Social System of a Mysore Village, in M. McKim (ed.): Village India, *University of Chicago Press*.

Srinivas, M. N. (1959): 'The Dominant Caste in Rampura', *American Anthropologist*, 1; 1-16

Srinivas, M. N. (1966): Social Change in Modern India, *Orient Longman*.

Srinivas, M. N. (1978): The Remembered Village, *Oxford University Press*.

Thorat, S. and J. Lee (2006): Dalits and the Right to Food – Discrimination and Exclusion in Food-Related Government Programmes, Working paper no. 3, Indian Institute of Dalit Studies, New Delhi.

Thorat, S. and P. Attewell (2007): 'The Legacy of Social Exclusion: A Correspondence Study of Job Discrimination in India', *Economic and Political Weekly,* 13, October.

APPENDIX I: The construction of social group variables to capture caste dominance.

The village and household questionnaires contain data on three classifications of social groups, firstly and most disaggregated by jati[37] [and name of tribe] (C1) (for Hindus, Muslims, Sikhs and STs), secondly by five broad categories (C2), namely Brahmin, OBC (Other backward classes), SC (Scheduled Caste), ST (Scheduled Tribe) and Other and finally by eight religious categories (C3), Hindu, Muslim, Christian, Sikh, Buddhist, Jain, Tribal and Other. The village questionnaire also contains information on the most (upto eight) numerous jatis, the percentage of the village population each of these jatis represent, and the percent of village land owned by each of these same jatis.

The oppression hypothesis is founded on the notion of (upper) caste dominance. If restricted to ritual rank, a simple and narrow definition would be to limit the upper caste label to *Brahmins*. Notions of upper caste advantage (and dominance) do, however, stretch beyond this top layer of the varna hierarchy.[38,39] A pragmatic alternative would be to add the "Other" category from the household questionnaire; the combination *Brahmin* (C2) plus "Other (C2)" and Hindu (C3) would then represent a broad definition of upper caste *Hindus*.

There are, however, important problems associated with the latter option; Firstly, the exclusive focus on Hindus would miss out on social groups that may be in a position to wield considerable power and influence but who belong to a different faith. To illustrate, some of the numerically important jatis in the panel transcend religious boundaries; in Punjab and as footnote 17 in the main text illustrates, there are significant numbers of Sikh and Hindu *Jat* households and Sikh and Hindu Dalit households with recent and violent inter-caste strife in Punjab involving *Jat* and Dalit Sikhs.[40] For Muslims and also noted in footnote 17, Fuller (1996) and other contributors to the same volume contend that while caste-like arrangements are common, few within the Muslim community admit to their existence. In spite of social ranks among Muslims, the much less accurate reporting of the social groups that Muslim panel households belong to, left no other option but to define Muslim households by their religion alone. A similar strategy was adopted for Scheduled Tribes. Although the tribe a household belongs is accurately reported, ethnographic evidence is not supportive of local hierarchies; STs are thus a single social category in our analysis.

Secondly, the process of "de-Sanskritisation", whereby social groups lobby to downgrade their official status in order to avail of reservation benefits has the implication that the definitions of forward castes that anthropologists and sociologists, informed by careful field observations, subscribe to, are increasingly at odds with official and survey data social group categories. The implementation of the

[37] Sub-caste.
[38] In addition, the prevalence of *Brahmin* households varies across regions.
[39] Even among *Brahmins* there are, of course, more fine-tuned internal rankings – *Gouda Saraswath* or *Konkani* Brahmins, who are fish eating residents of Karnataka's Coastal belt, have lower social status locally than the strictly vegetarian *Madhwa* or *Udupi Brahmins*.
[40] See http://hinduonnet.com/fline/fl2013/stories/20030704002703900.htm. Punjab is also the state with the highest percentage of Scheduled Castes in its population (28.9 % according to Census of India 2001). See Jodhka (2004) for more on Sikhism and caste.

Mandal Commission's (1978-80) recommendations added fresh impetus to reservations as political battleground and in the current political climate, it is not unusual to interpret the absence of 'backward' status as evidence of a social group's limited political clout. Examples of important groups who have acquired OBC status, include the 'clean-caste' *Vokkaligas*, the dominant peasant caste in Central and Southern Karnataka (e.g. Srinivas 1978; Scarlett Epstein et al. 1998), the ritually superior *Lingayats* in the same state (Bayly 1999; 294) and more recently the *Jats* in Uttar Pradesh (e. g. Jeffrey 2001) and Rajasthan; official status is therefore, in important instances, a reflection of political opportunism aimed at placating major vote banks having the unfortunate side effect of diluting official status as indicator of ritual status.[41]

Other variations in caste status are found at the lower end: *Nuniyas* and *Dhanuks*, who are OBCs in Uttar Pradesh, have Scheduled Caste status in West-Bengal. *Dhobis* (washermen), have SC status in some states but not in others. For jatis traditionally concentrated in the most degrading occupations, like leatherworkers (e.g. *Chamars*) and sweepers (e.g. *Balmikis*), SC status is less variant to state boundaries.

Further, social groups that are not OBC, SC or ST should necessarily be treated as upper castes for analytical purposes. There are intermediate social groups in many regions for whom a more fine tuned distinction is desirable. *Rods*, an important agricultural caste in Haryana, is classified as 'other' and thus forward officially and in the household questionnaire; this does not square with anthropological field observations (Prem Chowdhry, pers comm.). Further, and in tune with the Mandal commission's view and report, important agricultural castes such as the *Kurmis* of North and *Kunbis* of Central India do not enjoy the same local status as *Jats* and *Marathas*, respectively (Singh 1992; 41). For the former two, the OBC classification is therefore appropriate.

In our interpretation of upper caste which is informed by anthropological observations, it is possible for a social group to fit the upper caste definition in some states, but not in others. The adopted approach may therefore be described as adhering to ritual rank as far as the top and bottom layer is concerned, but to disconnect, whenever appropriate and for reasons already mentioned, from official categories for the more fluid middle layer. While this imposes an extra work burden, it is vital to sharply distinguish our present and small-scale endeavour from past efforts to develop extensive caste rankings for rural India. British colonial administrators have subsequently been caricatured for believing in the possibility of such a task which at the time paved the way for an obsession with caste and jati among late Victorian data collectors (Bayly 1999, chapter 3). For North-India, our classification of the most important and by far the most numerous groups (as well as households in our panel) is entirely consistent with the Mandal Commission's views: according to which the following broad groups should be treated as forward or upper castes; *Brahmins*, *Rajputs*, *Kayasthas*, *Jats*, *Marathas* and *Banias* (Singh 1992; 41).

[41] While de-Sanskritisation so far has tended to involve attempts to attain OBC-status, recent agitations by the *Gujjar*-community based on comparisons with the *Meena* community in Rajasthan aimed to downgrade their official status from OBC to ST. Similarly, in an article on UP politics, the Deccan Herald (4 March 2008), listed a number of groups whose official status were proposed 'downgraded' from OBC to SC.

Turning to the data, the easier ones to categorise are Brahmins who mostly but not always are described as *Brahmins*. Tiwari, Sharma, Tripathi, Mishra, Koshik (MP), Padhi (O), Iyer (TN) are examples of others. Categorisation of *Jats* and *Marathas* and for most of the time *Rajputs* (*Thakurs*) is straightforward, as are *Banias* (Agarwals, Guptas etc). The state-wise official lists of STs, SCs and OBCs provide a rich source of information and were extensively consulted to cross check the SC and ST classifications in the raw data.

Table A1 provides a listing of upper castes based on our definition and begins with all India upper caste jatis; these are classified as upper castes in all states. The state listing provides additional upper caste jatis, which are either sub-groups of the main jatis (*Jats* or *Rajputs*, say) or belong to a different upper caste social group (e.g *Mahajan*; *Leva Patel*). Note that the following list is based exclusively on jatis that feature in the panel data set/village level social composition data. If a state is not listed (e.g. Maharashtra), all upper caste groups in that state are included in the ALL INDIA row. Notice also that the jatis in the ALL India row are by far the most numerous in the North. A careful reader may also notice that while Andhra castes and Kerala Nayars are included in Tamil Nadu, this is not the case the other way around. This is a co-incidence – there are no upper caste households from Tamil Nadu amongst our Andhra Pradesh panel households.

	Upper castes
ALL INDIA	Brahmin, Bhumihar, Rajput (general, Thakur), Kayastha, Kshatriya, Khatri, Maratha, Jat (Sikh and Hindu), Bania (Agarwal, Gupta, Jaiswal) (plus equivalents in the South: *Vysya* in Andhra Pradesh, *Chettiar* in Tamil Nadu)
ADDITIONAL BY STATE	
Himachal Pradesh	Rajput (Suniar), Choudhary
Punjab	Rajput (Suniar), Kamboj (Sikh), Choudhary, Mahant (Sikh), Arora, Ahluwalia, Mahajan, Sood, Visnoi
Uttaranchal	Rana
Haryana	Rajput (Chauhan, Bishnoi), Jat (Jhangi), Kamboj (Sikh)
Rajasthan	Choudhary, Mahajan
Gujarat	Patel (general, Patidar, Leva, Kadava), Rajput (Jadeja [Chandravanshi], Parmar, Solanki), Darbar
Uttar Pradesh	Rajput (Chauhan, Negi [Gharwali]), Srivastava, Choudhary
West Bengal	Pokhrel, Dahal, Chettri, Mahishya, Sadgop, Roy
Orissa	Patnaik (general, Karan), Pradhan, Khandayat, Odia, Kalandi
Madhya Pradesh	Jat (Tomar), Choudhary, Maharaj
Andhra Pradesh	Reddy, Kapu [Balija, Telaga], Kamma [Naidu], Velama, Chowdary, Rajulu
Karnataka	Lingayat, Vokkaliga
Tamil Nadu	Mudaliar, Maravar, Vellalar, Nayar, Reddy, Naidu, Kamma Naidu
Kerala	Nayar (Nair)

Appendix 2

Table A2.1 Estimation results of the effects on income of social identity, village regime and demographic and locational controls

Model:	Social identity terms				Plus village regime				Plus agro-ecological zones				Plus state dummies				Plus demographic controls			
	Round 1		Round 2		Round 1		Round 2		Round 1		Round 2		Round 1		Round 2		Round 1		Round 2	
	b	T	b	t	b	t	b	t	b	t	b	t	b	t	b	t	b	t	b	t
Social identity:																				
SC	-0.463	-20.29	-0.607	-24.75	-0.374	-9.4	-0.51	-11.91	-0.381	-9.55	-0.517	-12.13	-0.384	-9.61	-0.519	-12.21	-0.37	-9.55	-0.455	-11.18
ST	-0.458	-12.8	-0.601	-16.3	-0.31	-4.57	-0.462	-6.95	-0.296	-4.44	-0.423	-6.41	-0.295	-4.42	-0.442	-6.77	-0.309	-4.92	-0.374	-6.23
OBC	-0.267	-11.4	-0.371	-14.92	-0.291	-5.87	-0.338	-6.09	-0.263	-5.38	-0.319	-5.81	-0.256	-5.23	-0.312	-5.79	-0.216	-4.59	-0.264	-5.16
MUS	-0.406	-10.81	-0.53	-12.73	-0.289	-4.76	-0.446	-6.68	-0.269	-4.4	-0.439	-6.58	-0.276	-4.49	-0.466	-7.12	-0.201	-3.35	-0.324	-5.14
Village regime:																				
SC X DSC					0.032	0.48	0.268	3.55	0.099	1.46	0.259	3.46	0.111	1.66	0.282	3.83	0.14	2.18	0.302	4.25
SC X PSC					-0.031	-0.59	0.062	1.08	-0.032	-0.6	0.059	1.06	-0.052	-0.99	0.048	0.87	-0.048	-0.97	0.074	1.43
ST X DST					-0.088	-1.22	-0.056	-0.78	-0.055	-0.79	-0.023	-0.33	-0.049	-0.69	-0.002	-0.04	-0.017	-0.26	-0.001	-0.01
OBC X DOBC					0.17	3.58	0.142	2.73	0.133	2.87	0.124	2.46	0.122	2.65	0.121	2.45	0.098	2.21	0.118	2.56
OBC X POBC					0	0.01	0.156	2.3	-0.042	-0.68	0.123	1.84	-0.058	-0.94	0.128	1.98	-0.053	-0.91	0.129	2.13
MUS X DMUS					-0.023	-0.28	0.16	1.85	-0.063	-0.79	0.084	0.98	-0.052	-0.65	0.101	1.23	-0.015	-0.2	0.122	1.58
DUC					0.202	4.2	0.285	5.67	0.128	2.63	0.176	3.47	0.116	2.39	0.112	2.3	0.109	2.32	0.107	2.32
SC X DUC					-0.113	-2.3	-0.149	-2.83	-0.108	-2.21	-0.137	-2.62	-0.103	-2.09	-0.135	-2.59	-0.082	-1.75	-0.135	-2.72
ST X DUC					-0.222	-2.02	-0.03	-0.27	-0.234	-2.2	-0.047	-0.43	-0.211	-1.97	-0.014	-0.13	-0.132	-1.27	-0.007	-0.07
MUS X DUC					-0.187	-2.11	-0.142	-1.41	-0.214	-2.43	-0.145	-1.44	-0.184	-2.08	-0.071	-0.71	-0.147	-1.75	-0.079	-0.84
OBC X DUC					-0.024	-0.41	-0.127	-1.99	-0.045	-0.78	-0.125	-1.96	-0.048	-0.84	-0.124	-1.98	-0.078	-1.41	-0.116	-1.95
Agro-ecological zones:																				
aez2									0.273	2.96	0.271	2.91	0.637	1.46	0.458	1.16	0.452	1.09	0.52	1.45
aez3									0.601	3.83	-0.217	-1.29	0.602	1.45	-0.127	-0.35	0.441	1.12	-0.173	-0.52
aez4									0.114	1.71	0.114	1.58	0.383	0.89	0.322	0.82	0.173	0.42	0.372	1.05
aez5									0.077	0.93	-0.204	-2.35	0.388	0.91	0.156	0.4	0.165	0.41	0.16	0.45
aez6									0.316	4.15	-0.05	-0.64	0.601	1.47	0.275	0.77	0.462	1.2	0.233	0.72
aez7									0.49	3.5	-0.024	-0.19	0.319	0.73	-0.208	-0.56	0.202	0.49	-0.188	-0.55
aez8									0.21	2.17	0.046	0.48	0.114	0.28	0.284	0.8	-0.071	-0.18	0.187	0.57
aez9									0.281	3.43	0.125	1.45	0.506	1.15	0.4	1	0.273	0.66	0.389	1.07

Model:	Social identity terms				Plus village regime				Plus agro-ecological zones				Plus state dummies				Plus demographic controls			
	Round 1		Round 2		Round 1		Round 2		Round 1		Round 2		Round 1		Round 2		Round 1		Round 2	
	b	T	b	t	b	t	b	t	b	t	b	t	b	t	b	t	b	t	b	t
aez10									-0.02	-0.26	-0.296	-3.67	0.21	0.5	0.291	0.77	0.017	0.04	0.278	0.81
aez11									0.034	0.43	-0.116	-1.37	0.261	0.69	0.546	1.62	0.139	0.39	0.524	1.73
aez12									-0.288	-3.68	-0.379	-4.62	-0.046	-0.13	0.193	0.64	-0.169	-0.51	0.151	0.55
aez13									-0.115	-1.21	-0.275	-2.65	0.049	0.1	0.259	0.61	-0.114	-0.26	0.272	0.71
aez14									-0.082	-1.04	0.23	2.82	0.34	0.75	0.274	0.66	0.068	0.16	0.255	0.68
aez16									-0.63	-3.37	-0.266	-1.18	-0.611	-3.35	-0.212	-1.01	-0.683	-3.76	-0.175	-0.94
aez17									-0.195	-1.31	0.365	1.94	-0.224	-0.54	0.429	1.15				
aez18									-0.247	-1.47	-0.61	-4.03	0.041	0.1	0.115	0.34	-0.015	-0.04	0.054	0.17
aez19									0.251	1.99	0.178	1.25	0.222	0.52	0.167	0.45	0.011	0.03	0.085	0.25
State dummy variables:																				
Bihar													-0.188	-0.68	-0.467	-1.79	-0.035	-0.13	-0.318	-1.29
Gujarat													-0.344	-1.55	-0.122	-0.55	-0.248	-1.2	-0.114	-0.56
Haryana													-0.088	-0.39	0.058	0.26	0.103	0.49	0.101	0.48
Himachal Pradesh													-0.494	-1.84	0.043	0.16	-0.285	-1.13	0.077	0.31
Karnataka													-0.546	-1.84	-0.27	-1.32	-0.409	-1.35	-0.111	-0.55
Kerala													0.144	0.55	0.595	2.17	0.228	0.92	0.559	2.13
Madhya Pradesh													-0.253	-1.27	-0.528	-2.66	-0.1	-0.55	-0.457	-2.49
Maharashtra													-0.321	-2.21	-0.262	-1.81	-0.26	-1.96	-0.205	-1.5
Orissa													-0.366	-2.03	-0.67	-4.49	-0.352	-2.09	-0.584	-4.2
Punjab													-0.258	-1.09	0.153	0.66	-0.102	-0.46	0.181	0.84
Rajasthan													-0.513	-2.3	-0.225	-1.02	-0.328	-1.58	-0.143	-0.7
Tamil Nadu													0.067	0.47	-0.223	-1.65	0.099	0.73	-0.16	-1.22
Uttar Pradesh													-0.39	-1.67	-0.426	-1.87	-0.206	-0.95	-0.312	-1.48
West Bengal													-0.046	-0.13	0.025	0.08	-0.122	-0.35	0.029	0.1
Uttaranchal													-0.448	-1.68	-0.214	-0.78	-0.235	-0.94	-0.114	-0.45
Chattisgarh													-0.27	-1.33	-0.627	-3.28	-0.23	-1.22	-0.564	-3.28
Jharkhand													-0.033	-0.17	-0.146	-0.84	0.042	0.22	-0.128	-0.8
Tripura																	-0.352	-0.9	0.403	1.19

Model:	Social identity terms				Plus village regime				Plus agro-ecological zones				Plus state dummies				Plus demographic controls			
	Round 1		Round 2		Round 1		Round 2		Round 1		Round 2		Round 1		Round 2		Round 1		Round 2	
	b	T	b	t	b	t	b	t	b	t	b	t	b	t	b	t	b	t	b	t
Demographic controls																				
Sex of hh head (male = 1)																	0.031	1.51	-0.012	-0.51
# males aged 0-5																	-0.146	-13.58	-0.179	-14.14
# males aged 6-14																	-0.112	-12.92	-0.142	-15.32
# males aged 15-19																	-0.005	-0.41	-0.036	-2.74
# males, aged 20-24																	0.072	4.79	0.059	3.78
# males, aged 25-49																	0.104	5.74	0.108	7.04
# males, aged 50-59																	0.178	8.03	0.135	5.92
# males, aged 60 +																	0.075	3.57	0.055	2.7
# females, aged 0-5																	-0.131	-13.21	-0.154	-11.97
# females, aged 6-14																	-0.117	-13.76	-0.134	-15.08
# females, aged 15-19																	-0.066	-4.46	-0.089	-6.89
# females, aged 20-24																	-0.016	-0.79	-0.016	-0.88
# females, aged 25-49																	0.07	3.34	0.09	4.62
# females, aged 50-59																	0	0	0.098	3.81
# females, aged 60 +																	0.014	0.59	-0.031	-1.36
# of couples in household																	-0.021	-1.24	0.05	2.92
Constant	8.958	389.79	9.192	385.09	8.83	234.67	9.008	224.26	8.787	139.06	9.092	127.22	8.82	22.53	9.043	26.69	8.946	24.04	8.944	28.69
R squared	0.061		0.095		0.066		0.107		0.115		0.161		0.137		0.200		0.213		0.284	
N	9111		9111		9111		9111		9111		9111		9111		9111		9111		9111	

Source: HDPI-I ("round 1") and II ("round 2") surveys, panel households only; authors' calculations.

Notes: Dependent variable is the natural logarithm of annual per capita household income in constant 1993/94 prices, with round 2 figures converted using NSSO state-specific rural CPIs. Random effects, with standard errors that are robust to heteroskedasticity and clustering within villages; robust t-statistics are reported.

Table A2.2 Estimation results of the effects on income of social identity, village regime and additional controls: village infrastructure, household education and land

Controls added:	Village infrastructure				Plus education				Plus land			
	Round 1		Round 2		Round 1		Round 2		Round 1		Round 2	
	B	t	B	t	b	t	b	t	b	t	b	t
Social identity:												
SC	-0.369	-9.56	-0.452	-11.07	-0.245	-6.58	-0.321	-8.23	-0.154	-4.49	-0.252	-6.81
ST	-0.311	-4.95	-0.374	-6.17	-0.172	-2.81	-0.234	-4.09	-0.152	-2.72	-0.203	-3.64
OBC	-0.21	-4.45	-0.259	-5.05	-0.141	-3.08	-0.181	-3.7	-0.129	-3.12	-0.148	-3.25
MUS	-0.208	-3.48	-0.324	-5.12	-0.113	-1.98	-0.217	-3.59	-0.065	-1.23	-0.14	-2.38
Village regime:												
SC X DSC	0.131	2.05	0.289	4.03	0.123	2.01	0.253	3.64	0.06	1.05	0.205	3.16
SC X PSC	-0.035	-0.72	0.076	1.44	-0.053	-1.08	0.07	1.39	-0.062	-1.34	0.059	1.23
ST X DST	0.008	0.12	0.019	0.31	-0.006	-0.1	0.005	0.08	0.012	0.22	-0.012	-0.22
OBC X DOBC	0.095	2.15	0.115	2.49	0.094	2.16	0.097	2.19	0.067	1.73	0.05	1.21
OBC X DOBC	-0.059	-1.0	0.134	2.2	-0.044	-0.77	0.118	1.98	-0.012	-0.23	0.106	1.98
MUS X DMUS	0.003	0.04	0.147	1.88	0.006	0.08	0.15	2.01	-0.048	-0.73	0.087	1.21
DUC	0.116	2.48	0.107	2.3	0.124	2.76	0.105	2.36	0.035	0.89	0.024	0.57
SC X DUC	-0.09	-1.92	-0.14	-2.81	-0.107	-2.39	-0.15	-3.17	0.01	0.24	-0.045	-1.01
ST X DUC	-0.14	-1.35	-0.01	-0.1	-0.129	-1.27	0.017	0.18	-0.032	-0.36	0.112	1.26
MUS X DUC	-0.15	-1.79	-0.094	-0.99	-0.129	-1.58	-0.054	-0.58	0.026	0.34	0.003	0.03
OBC X DUC	-0.091	-1.64	-0.123	-2.07	-0.085	-1.59	-0.113	-1.99	0.076	1.59	0.001	0.01
Agro-ecological zones:												
aez2	0.175	0.42	0.341	0.88	0.18	0.44	0.296	0.81	-0.026	-0.07	0.223	0.64
aez3	0.42	1.06	-0.316	-0.86	0.358	0.93	-0.33	-0.95	0.133	0.38	-0.348	-1.05
aez4	-0.096	-0.23	0.173	0.45	-0.137	-0.34	0.092	0.25	-0.195	-0.52	0.125	0.37
aez5	-0.07	-0.17	-0.009	-0.02	-0.056	-0.14	-0.066	-0.18	-0.177	-0.48	-0.058	-0.17
aez6	0.324	0.84	0.111	0.31	0.348	0.91	0.079	0.24	0.233	0.67	0.042	0.13
aez7	0.149	0.36	-0.342	-0.92	0.22	0.54	-0.294	-0.83	0.028	0.08	-0.319	-0.95
aez8	-0.046	-0.12	0.07	0.19	0.006	0.02	0.093	0.27	-0.141	-0.4	0.087	0.27
aez9	0.038	0.09	0.191	0.49	0.001	0	0.072	0.2	-0.043	-0.11	0.139	0.4

Controls added:	Village infrastructure				Plus education				Plus land			
	Round 1		Round 2		Round 1		Round 2		Round 1		Round 2	
	B	t	B	t	b	t	b	t	b	t	b	t
aez10	-0.15	-0.38	0.118	0.32	-0.211	-0.54	-0.006	-0.02	-0.265	-0.73	0.02	0.06
aez11	0.046	0.13	0.283	0.85	0.029	0.08	0.244	0.79	-0.12	-0.36	0.194	0.65
aez12	-0.125	-0.38	0	0	-0.078	-0.24	0.039	0.14	-0.184	-0.61	0.006	0.02
aez13	-0.374	-0.84	0.074	0.18	-0.303	-0.69	-0.036	-0.09	-0.159	-0.41	0.087	0.24
aez14	-0.132	-0.3	0.064	0.16	-0.287	-0.67	-0.114	-0.3	-0.178	-0.46	0.028	0.08
aez16	-0.547	-3.11	-0.129	-0.68	-0.516	-3.06	-0.146	-0.82	-0.413	-2.48	-0.102	-0.59
aez17	-0.305	-0.78	0.223	0.6	-0.268	-0.69	0.231	0.66	-0.318	-0.89	0.171	0.52
aez18	-0.016	-0.04	-0.074	-0.22	0.037	0.1	-0.031	-0.1	-0.068	-0.2	-0.005	-0.02
aez19	-0.03	-0.07	-0.049	-0.13	-0.081	-0.2	-0.116	-0.33	-0.093	-0.25	-0.049	-0.15
State dummy variables:												
Bihar	0.166	0.63	-0.253	-1.01	0.159	0.6	-0.15	-0.62	-0.15	-0.69	-0.244	-1.09
Gujarat	-0.051	-0.24	-0.068	-0.33	0.016	0.08	0.031	0.16	0.002	0.01	-0.051	-0.28
Haryana	0.307	1.41	0.136	0.64	0.395	1.81	0.234	1.14	0.224	1.15	0.184	0.98
Himachal Pradesh	-0.042	-0.17	0.136	0.55	0.116	0.46	0.259	1.08	0.011	0.05	0.225	1.01
Karnataka	-0.344	-1.15	-0.102	-0.5	-0.286	-0.93	-0.086	-0.44	-0.344	-1.28	-0.101	-0.55
Kerala	0.241	0.97	0.491	1.83	0.251	1.05	0.427	1.63	0.255	1.15	0.433	1.71
Madhya Pradesh	0.179	0.92	-0.398	-2.12	0.288	1.48	-0.255	-1.41	0.097	0.55	-0.353	-2.11
Maharashtra	-0.17	-1.23	-0.213	-1.53	-0.145	-1.05	-0.21	-1.6	-0.184	-1.49	-0.221	-1.83
Orissa	-0.302	-1.77	-0.581	-4.07	-0.275	-1.64	-0.576	-4.19	-0.271	-1.81	-0.536	-4.15
Punjab	0.076	0.34	0.222	1.02	0.174	0.78	0.312	1.48	-0.065	-0.33	0.233	1.21
Rajasthan	-0.095	-0.44	-0.08	-0.38	0.014	0.07	0.038	0.19	-0.099	-0.51	-0.041	-0.22
Tamil Nadu	0.043	0.32	-0.183	-1.35	0.014	0.1	-0.222	-1.63	0.037	0.3	-0.228	-1.78
Uttar Pradesh	0.017	0.08	-0.256	-1.18	0.115	0.51	-0.152	-0.73	-0.02	-0.1	-0.215	-1.11
West Bengal	-0.15	-0.43	-0.117	-0.37	-0.09	-0.26	-0.071	-0.24	-0.209	-0.67	-0.063	-0.22
Uttaranchal	0.012	0.05	-0.049	-0.19	0.183	0.73	0.131	0.52	-0.038	-0.17	0.018	0.08
Chattisgarh	-0.115	-0.6	-0.453	-2.5	-0.037	-0.2	-0.387	-2.25	-0.013	-0.07	-0.369	-2.29
Jharkhand	0.04	0.22	-0.079	-0.48	0.054	0.3	-0.082	-0.52	0.077	0.46	-0.002	-0.01
Tripura												

Controls added:	Village infrastructure				Plus education				Plus land			
	Round 1		Round 2		Round 1		Round 2		Round 1		Round 2	
	B	t	B	t	b	t	b	t	b	t	b	t
Demographic controls:												
Sex of hh head (male = 1)	0.031	1.51	-0.013	-0.55	0.039	1.93	-0.015	-0.66	0.039	2.16	-0.02	-0.92
# males aged 0-5	-0.146	-13.51	-0.178	-14.03	-0.139	-13.27	-0.16	-13.14	-0.149	-16.14	-0.149	-12.72
# males aged 6-14	-0.112	-12.88	-0.141	-15.2	-0.102	-12.22	-0.114	-12.73	-0.125	-16.59	-0.125	-14.64
# males aged 15-19	-0.006	-0.44	-0.035	-2.72	-0.046	-3.62	-0.073	-5.47	-0.077	-6.86	-0.094	-7.63
# males, aged 20-24	0.072	4.75	0.06	3.81	0.011	0.73	-0.015	-0.93	-0.011	-0.82	-0.032	-2.18
# males, aged 25-49	0.103	5.72	0.108	7.01	0.041	2.38	0.01	0.66	0.013	0.84	-0.01	-0.72
# males, aged 50-59	0.179	8.1	0.134	5.9	0.116	5.47	0.065	2.93	0.058	3.1	0.031	1.49
# males, aged 60 +	0.075	3.6	0.055	2.7	0.028	1.4	0.002	0.12	-0.03	-1.73	-0.038	-2.06
# females, aged 0-5	-0.131	-13.24	-0.153	-11.9	-0.124	-12.84	-0.135	-11.02	-0.132	-15.45	-0.138	-11.85
# females, aged 6-14	-0.117	-13.72	-0.135	-15.14	-0.112	-13.6	-0.117	-13.68	-0.129	-17.12	-0.131	-16.25
# females, aged 15-19	-0.067	-4.52	-0.09	-6.89	-0.131	-8.34	-0.136	-9.74	-0.135	-9.73	-0.145	-10.88
# females, aged 20-24	-0.017	-0.86	-0.018	-0.97	-0.092	-4.43	-0.103	-5.48	-0.102	-5.38	-0.112	-6.41
# females, aged 25-49	0.068	3.24	0.088	4.51	-0.005	-0.23	-0.007	-0.36	-0.036	-1.88	-0.045	-2.38
# females, aged 50-59	-0.002	-0.11	0.096	3.74	-0.055	-2.48	0.006	0.26	-0.075	-3.8	-0.021	-0.91
# females, aged 60 +	0.011	0.47	-0.034	-1.47	-0.043	-1.84	-0.104	-4.66	-0.077	-3.69	-0.131	-6.31
# of couples in household	-0.019	-1.12	0.052	3.01	0.026	1.62	0.106	6.29	0.007	0.49	0.067	4.25
Village infrastructure:												
Ln(village population)	0.017	0.88	-0.006	-0.38	0.006	0.33	-0.011	-0.78	0.011	0.64	-0.009	-0.61
School access:												
Primary	0.006	0.06	0.365	1.89	-0.017	-0.18	0.276	1.51	-0.012	-0.14	0.306	1.79
Middle	-0.194	-3.21	-0.128	-1.89	-0.184	-3.13	-0.051	-0.77	-0.126	-2.47	-0.066	-1.05
Lower secondary	0.044	0.65	-0.096	-1.57	0.058	0.86	-0.034	-0.57	0.062	1.04	-0.078	-1.41
Higher secondary	-0.012	-0.28	-0.075	-1.34	-0.006	-0.13	-0.044	-0.81	-0.023	-0.61	-0.058	-1.12
Graduate	-0.018	-0.44	-0.068	-1.27	-0.021	-0.51	-0.036	-0.69	-0.021	-0.57	-0.048	-0.97
Vocational	-0.016	-0.34	-0.028	-0.41	-0.043	-0.91	-0.013	-0.2	-0.003	-0.07	-0.019	-0.31
Medical access:												
Doctor	0	-0.2	-0.003	-1.85	0	-0.02	-0.002	-1.55	-0.001	-0.76	-0.002	-1.65

Controls added:	Village infrastructure				Plus education				Plus land			
	Round 1		Round 2		Round 1		Round 2		Round 1		Round 2	
	B	t	B	t	b	t	b	t	b	t	b	t
Clinic	-0.017	-0.49	-0.033	-1.08	-0.036	-1.09	-0.035	-1.18	-0.027	-0.91	-0.016	-0.56
Road access:												
Feeder	0.069	1.31	-0.012	-0.21	0.064	1.23	-0.018	-0.33	0.039	0.83	-0.017	-0.34
Tarmac	0.115	2.0	0.008	0.13	0.099	1.76	-0.004	-0.08	0.078	1.54	-0.008	-0.16
Bus stop	0.016	0.42	0.004	0.14	0.015	0.4	0.011	0.38	0.01	0.32	0.014	0.51
Railway station	0.1	1.96	0.105	2.09	0.063	1.29	0.068	1.43	0.068	1.54	0.057	1.28
Post office	0.11	3.04	0.029	0.95	0.106	2.99	0.015	0.52	0.109	3.4	0.017	0.63
Bank/credit market	-0.048	-1.3	-0.013	-0.39	-0.055	-1.51	-0.048	-1.52	-0.028	-0.87	-0.023	-0.75
Market/mandi	-0.066	-1.81	0.001	0.05	-0.074	-2.09	0.004	0.15	-0.062	-1.94	0.013	0.48
Max. educational achievement in the household (of those 15+):												
Up to primary					0.082	4.36	0.034	1.51	0.067	4.06	0.041	1.94
Middle					0.177	7.94	0.128	5.98	0.143	7.31	0.129	6.38
Matriculation					0.273	10.01	0.308	10.64	0.208	8.55	0.285	10.45
Higher secondary					0.342	10.07	0.363	11.71	0.28	9.05	0.31	10.54
Graduate and above					0.584	15.23	0.606	16.5	0.457	12.75	0.512	14.33
Up to primary					0.083	4.05	0.087	4.01	0.061	3.4	0.071	3.51
Middle					0.19	6.64	0.116	5.2	0.138	5.46	0.088	4.2
Matriculation					0.247	6.96	0.149	4.6	0.163	5.13	0.14	4.54
Higher secondary					0.193	3.15	0.33	8.34	0.214	3.66	0.263	7.02
Graduate and above					0.263	2.98	0.336	6.25	0.288	3.63	0.287	5.57
Household land:												
Land owned in acres									0.022	5.56	0.029	7.48
Land gross irrigated in acres									0.047	12.5	0.036	4.99
Constant	8.747	21.21	9.207	23.96	8.659	21.33	9.076	25.09	8.822	23.79	9.157	26.48
R squared	0.225		0.288		0.270		0.341		0.426		0.418	
N	9111		9111		9111		9111		9111		9111	

Source and Notes: as for Table A2.1