The Political Construction of Caste in South India

Vijayendra Rao (vrao@worldbank.org)
Development Research Group, The World Bank

And

Radu Ban (rban@worldbank.org)
London School of Economics and Development Research Group, The World Bank

August 2007

We thank seminar participants at the World Bank's research department and Karla Hoff for helpful conversations and comments. Jillian Waid and Babu Srinivas Dasari provided excellent research assistance. This paper reflects the views of the authors and should not be attributed to the World Bank, its member countries or any affiliated organization. We are indebted to the Dutch government, and the Research Support Budget of the Development Economics Vice-Presidency of the World Bank, for financial support.

Abstract

Are social institutions endogenous? Can measures of social diversity (e.g. fractionalization) be treated as exogenous variables in assessing their impact on economic and political outcomes? The caste system, which categorizes Hindus into endogamous and stratified social groups, is considered to be the organizing institution of Indian society. It is widely thought to have stayed stable for hundreds if not thousands of years -- so deeply resistant to change that it has been blamed for everything from (formerly) anemic "Hindu" rates of growth, to persistent "inequality traps." This paper uses a natural experiment -- the 1956 reorganization of Indian states along linguistic lines – to demonstrate that the number and nomenclature of castes has significantly changed in linguistically matched villages (i.e. "mistakes" in the reorganization) at the borders of these states. This shows that the caste system is not stable but a pliable institution - endogenous to political change.

Introduction:

A prolific line of research has explored the impact of ethnic and social diversity on a variety of economic and political outcomes (Alesina and La Ferrara, 2005). Economists, however, have paid far less attention to the possibility that social structures may be endogenous[1], though historians and anthropologists have made it an active area of research[2]. While economists now recognize that political institutions may be endogenous to economic change (Acemoglu, Johnson and Robinson 2005, Engerman and Sokoloff 2003), they have tended to treat social categories such as caste and race as fixed in time and exogenous. The small but growing literature on the economics of identity choice (Caselli and Coleman 2006, Bloch and Rao 2001, Akerlof and Kranton 2000), which examines how individuals chose a social identity within a given and fixed set of alternatives, has made a move in this direction. Our focus, however, is on a different question - what if the alternatives – the names and the number of choices in identity-categories - themselves were endogenous? This would have fundamental implications for the analysis and measurement of fractionalization and polarization, besides raising questions about whether such measures of diversity could be included on the right hand side of an OLS regression.

Our focus in this paper is on the Indian caste system – which has long been the archetype of a rigid and unchanging social institution that traps individuals within a hierarchical, hereditary structure which determines their economic and social status. We examine how caste can be transformed by political change. More specifically, we analyze the impact of an exogenous shift in the boundaries of states on caste structures (the number of castes, and their names) in villages affected by the change. Thus, we question the widespread assumption among social scientists that caste structures are fixed, given, and very slow to change.

This assumption – best described as a trope[3] - has long historical roots. Beginning with Alberuni, a thousand years ago, scholars have considered the caste system as the organizing institution of Indian society. A voluminous literature on the subject has evolved since then – with various

[1] A point made by Alesina and La Ferrara (2005) in their thoughtful review of the economics literature.
[2] A review of the anthropology and history literature on this would require another paper but for important work in different contexts see Sahlins (1991) for Spain, and Bayly (1999) for India.
[3] A persistent and familiar idea or theme, i.e.: an intellectual cliché.

western scholars commenting on its exotic and exploitative rules and practices from the Abbé Dubois (1806) in the eighteenth century[4], to Max Weber (1966) and James Mill (1820) in the nineteenth, and Louis Dumont (1980) in the twentieth. This "Orientalist[5]" trope of caste as a system of inherited institutional rigidity has become an integral part of the scholarly canon and is reflected in the writing of politicians, including Gandhi and Nehru, to contemporary anthropologists and economists

The description of caste in this literature – primarily drawn from ancient Sanskrit texts such as the *Manusmriti* - can be briefly summarized as follows: Hindu society has been divided for millennia into four hierarchical groups or *varnas*; led by Brahmins (the scholars and teachers), followed by Kshatriyas (rulers and warriors), Vaishyas (traders and merchants), and Shudras (artisans and laborers). A fifth group, considered Untouchable – is so low as to be outside the domain of the ritual hierarchy and relegated to occupations such as scavenging. In everyday practice castes manifest as *jatis*[6] - endogamous groups defined within regional and linguistic boundaries – which are mapped onto varna categories and thus bound within those hierarchies. This conception of caste reached its most sophisticated expression in Louis Dumont's (1980) influential description of Hindus as a species that he called *Homo Heirarchicus*[7]. To paraphrase Dumont's complex argument, the caste system is perpetuated by an ideology where upper castes justify the hierarchy because they internalize the belief that their inherited high status is inevitable as the fruit of efforts in past lives, while lower castes similarly internalize the justification of their low status. This stands in opposition to western notions of individual equality and as an alternative to the individually rational *Homo Economicus*.

The fast expanding literature on the economics of caste is not immune to the trope. To cite a few examples, Akerlof (1976) in his classic work on caste and the rat-race outlines a model of a stable and persistent caste-equilibrium where obedience to the caste-code results in the sub-optimal allocation of labor. More recently, Freitas (2006) formalizes a model of caste system that attempts to understand why "this system of social stratification" persisted for "3000 years of

[4] Whose work, apparently, was plagiarized from a manuscript by Père Couerdoux (Dirks, 2002).
[5] The argument that modern conceptions of eastern institutions are creations of, often biased, colonial western interpretations (Said, 1978).
[6] From this point on we will use the word "caste" and "jati" interchangeably.
[7] See Appadurai (1986) for an argument that connects Dumont to Orientalism, and Khare (2006) for a compilation of the important critiques of Dumont.

changing economic and social environments.[8]" Rao[9] and Walton (2004a) use the trope as an example of how poverty traps can develop if the poor internalize hierarchical norms, and Hoff and Pande (2004) use it to explain findings from a field-based experiment test within a north Indian setting[10].

Some of the literature relating the stickiness of caste structures to the process of economic development takes an even more radical turn, leading Lal (1989) to argue in his *Hindu Equilibrium* that caste has an atavistic hold on the development process which leads to rigidities that have resulted in anemic "Hindu rates of growth[11]." A similar point is made by Olson (1982) in *The Rise and Decline of Nations* who cites rigidities caused by the caste system as an important example of the institutional constraints to growth. This feeds right in to the influential views of Huntington and Harrison (2001) that *Culture Matters* because the long arm of path-dependent value-systems exert a strong hold on an individual's ability to undertake entrepreneurial and mobility-enhancing actions – thus, in effect, "blaming" a country's poor-development on its culture.

There is no disagreement, however, that caste is strongly correlated with inequality; Deshpande (2001) has shown that low caste groups face considerably more deprivation across India, and Kijama (2006) demonstrates that this disparity has persisted at least since the 1980's. Anderson (2005) finds that low-caste status results in lower incomes because of impediments to trade across groups. Banerjee and Somanathan (2006) show how caste structures affect the allocation of public goods and services in villages, with scheduled caste dominated villages converging more with upper caste dominated villages over time, than villages dominated by scheduled tribes. More recently it has been argued that in some circumstances low caste-status can result in poor

[8] To be fair, Freitas makes clear that her model is of caste in the pre-colonial environment, but as we will soon make clear it is not clear that caste was ever a stable institution. Also, this is not a judgment on the value of the paper which provides a new way of looking at persistent inequality that is an interesting extension of the Akerlof model.

[9] Yes, indeed, Rao is disagreeing with himself.

[10] Again, to be fair, Hoff and Pande are describing an important empirical phenomenon that demonstrates the pernicious effects of institutionalized inequality, and Rao and Walton make it clear that the value of Dumont to them is simply as a theoretical illustration of a larger point rather than an empirical description of the caste system.

[11] This refers to the 3.3% growth rate that stubbornly persisted in India for the first forty years of its existence before it morphed, largely in the last decade, into the Bangalorized miracle of today with 8-9% growth rates. Lal has revised some of his ideas to contend with this recent tripling of growth in the 2005 edition of this book.

performance in a manner that incorporates diminished expectations that emerge from discrimination (Rao and Walton 2004, Hoff and Pande 2004).

It has also long been recognized that there is a degree of mobility within the caste system. About the same time as Dumont, MN Srinivas (1966) argued that upwardly mobile castes, as groups, move up in the ritual hierarchy over a few generations by acquiring social and religious practices that are associated with Brahmanical castes. Srinivas called this process "Sanskritization" since Brahmins are distinguished from other caste by the extent of their access to the Sanskrit language and the sacred texts written in it. In Srinivas's view, thus, while group-based mobility exists it is slow and internally driven – sparked, perhaps, by economic or technological changes that increase the returns to the occupation that the caste traditionally performs. Therefore, while Sanskritization brings in a dynamic element into the theory of caste, akin to the literature on the economics of identity it remains consistent with the notion that the names and number of castes within a particular context are given. Economists have also recently begun to study individual mobility within the caste system. Munshi and Rosenzweig (2006) demonstrate that caste-based networks play an important role in the economic mobility of families in Mumbai and that this mobility leads to more English-based schooling for girls rather than for boys which could, in the long run, affect marriage choices and hence caste. However, they again do not argue with the contention that the basic construct of caste has stayed stable for a long time[12].

Our paper is concerned neither with inequality nor mobility within the caste system, but with its structure – the names, and the number of *jatis* into which caste society within a village is divided. It is, therefore, an analysis about the endogenous nature of the caste system as a whole – not just about particular individuals or groups within it. The paper falls squarely within a literature in history and anthropology that has raised fundamental questions about the Orientalist idea of caste. At the level of theory, Das (1981) questions whether the *Manusmriti* is a good representation of how caste is treated in Sanskrit texts, and their translation into practice. She argues that rather than a clear statement of hierarchy – textual sources that focus on caste tend to reveal a triangular formation between rulers, priests and ascetics, with a broad category of "shudras" who lie outside this triangle.

[12] In other work Munshi and Rosenzweig (2005) find that spatial and economic mobility in caste in rural India is low, and attribute this to unwillingness of households to give up access to sub-caste networks.

At an empirical level, anthropologists and historians have, over the last decade, begun a fundamental critique of the trope. Influenced by post-structuralist social theory, and building on the pioneering work of Cohn (1987), Nicholas Dirks (2002) argues in *Castes of Mind* that the modern caste system is largely a construction of British colonial rule. In tracing the history of ideas on caste, Dirks demonstrates the British colonial administrators, Risley (1912) being the most influential among them, were deeply influenced by the work of the Sanskrit scholar Max Mueller who introduced the *Manusmriti* to Western audiences and who, in turn, built on what he considered the field observations of Dubois to argue for its contemporary relevance. Over the course of the 19th century caste became a central construct by which the British categorized and controlled the numerous jatis and tribes that were becoming absorbed into their empire. Risley led the process by which this was done, largely by incorporating questions on caste into the Indian census. But the process of translating the fluid local dynamics of caste into a finite number of standardized quantitative census categories hardened the caste system and "created" a new form of caste – that was amenable to quantification, less fluid, and easier for policy makers to manage. Over time these categories were internalized and, among other things, the awareness of the large proportion of low castes in the population helped social reformers generate social movements by using the new categories to mobilize disadvantaged groups.

The historian Susan Bayly (1999), who traces the history of caste from 1500 to 2000, states that "caste is not and never has been a fixed fact on Indian life." Local caste systems were fundamentally modified every time a new ruler arrived and imposed different systems of tenure, revenue generation, and royally sanctioned rewards and punishments. Bayly says that caste and jati are "best seen as composites of ideals and practices that have been made and remade into varying codes of moral order over hundreds or even thousands of years." She focuses, in particular on four periods – the first from the 15th to the 18th century was the period of "warrior dynasts" where a strong link was forged between newly conquering rulers and Brahmins who "remade" kings into Kshatriyas and thus legitimized their rule with ritual sanction. The second period from 1700-1830 she describes as the "Brahmin-Raj" when Dumontian Brahmin-centered ethic became widespread because of the increasing domination of Brahmins in centers of political authority and trade. The third phase in the colonial phase, described above in Dirk's account, and the final phase is the post-independence phase when caste-based politics, affirmative action and the federal structures of politics have completely reshaped caste dynamics and recreated caste.

It is within these accounts of the changing nature of caste that we place our paper. It can be interpreted as an econometric extension of the Bayly-Dirks historical argument to the modern, post-independence, period. Political reorganizations result in changes in the "gaze" of government; In how governments measure, count and assign benefits – for instance via land reform or affirmative action; in how caste-based social movements are mobilized; and in other social processes that cause a change in how caste systems are structured. In this paper we will attempt to identify the causal effect of shifts in political boundaries on the caste system, and try to suggest ways in which these shifts might have occurred.

If ethnic categories can be constructed, this can have important implications for policy because it suggests that history is not destiny, as some recent work on path dependency in the development process has suggested (e.g.: Acemoglu, Johnson and Robinson 2001, Banerjee and Iyer 2005). It also raises questions about whether analyses of the effect of caste and other social and ethnic categories on welfare can treat such categories as exogenous. It posits instead that institutions that create social categories, which in turn affect fractionalization and inequality, are endogenous. If social institutions are themselves political constructions with relatively recent histories, then it should lead us to believe that fractionalization and polarization are less the result of deep-seated institutional path-dependencies than dynamic processes that both affect and are affected by economic development and political change. Instead of assuming the exogeneity of social institutions, an argument would have to be made that their pace of change is more gradual than the economic and political outcomes they are correlated with.

Our paper shares a kinship with recent work by Miguel and Posner (2006) that shows that ethnic identification in Africa is more a product of modernity than tradition and related to economic and political considerations, and Miguel's (2004) results that demonstrate how state policy can affect the ethnic relationships in a manner that influences development effectiveness. It is also related to Sahlins (1989) who conducts a comparative case study of historical shifts in social identity within linguistically matched pairs of Catalan villages along the French-Spanish border[13].

In the next section of the paper we will describe the natural experiment that underlies our identification strategy and the data that we use. Section 3 will outline the econometric methodology and the results, and Section 4 will conclude the paper.

[13] We thank Stathis Kalyvas for pointing this out.

The Natural Experiment:

The map of British India was stitched together from the remnants of the Mughal Empire. After Mughal dominance over the sub-continent disintegrated over the course of the 17th and 18th centuries, Hindu and Muslim generals, courtiers, local chieftains and other sundry dissidents started exercising dominance over territory and gradually carved out autonomous kingdoms. The British East India Company entered India in the 16th century initially for the purpose of trade. In the process of establishing trade routes and consolidating trade monopolies, they gradually began to extend control – via treaties and force – over territory. Depending on the conditions of the exchange of power and the local political situation in some places territory was directly governed by the Crown – gradually extending to large states that were known as "Presidencies." In other places, indigenous rulers were put in place, endowed with large incomes and some local autonomy, in "Princely States" that were indirectly controlled by British "Residents."

The shape of these territories closely reflected their historical antecedents. In Southern India, the state of Hyderabad was ruled by a Nizam – the first of whom was a Mughal governor who had extracted control from his erstwhile suzerains over a large portion of the empire's territory in the Deccan plateau. The state of Mysore was constructed in the early 19th century from the remnants of the kingdom of Tippu Sultan whose reign was characterized by creative and successful resistance to British rule until successive defeats in the Third (1792) and Fourth Mysore Wars (1799) which are among the most decisive battles in the history of British colonial expansion. Part of Tippu's empire was carved into Mysore state and a member of the Wodeyar family – considered the original Hindu rulers of the state – was installed on the throne. Much of the rest of South India, cobbled together by gradual expansion from the port city of Madras - was a "Presidency" under direct British rule[14] – with its capital in the port city of Madras, from which it took its name.

Indian independence in 1947 brought with it a number of social movements which promoted a unified linguistic identity. At the same time a number of leading Indian politicians and intellectuals were advocating that Indian states be reorganized along linguistic lines in the belief that they could then be more rationally governed. A commission[15] was instituted to go through the painstaking process of taking a hard look at historical logic and census data to solve the

[14] There were two other large princely states in the South – Travancore and Cochin – that fall outside the realm of our sample.

[15] The members of the commission were Justice Fazl Ali - Chairman, Dr. H. N. Kunzru, and K.M. Panikkar.

jigsaw puzzle of putting together new, linguistically unified states by merging districts that had the same majority language. The commission's report was published in 1955 and its recommendations implemented in 1956. In the South, this led to the creation of four states – Andhra Pradesh (AP) which was largely Telugu-speaking, Tamil Nadu (TN) - Tamil speaking, Karnataka (KA) – Kannada speaking, and Kerala (KE) – Malayalam speaking. AP was pieced together from Hyderabad and the Telugu speaking parts of the Madras Presidency. Karnataka was carved together by merging the erstwhile princely state of Mysore with Kannada speaking parts of Hyderabad, and the Madras and Bombay presidencies. Kerala was formed by merging the princely states of Travancore and Cochin with parts of the Madras Presidency, and the remaining Tamil speaking areas of Madras Presidency became Tamil Nadu.

The States Reorganization Commission's report (Govt. of India, 1955) details the process by which decisions were made to assign particular districts to particular states. The primary consideration was the language spoken by a majority of its residents, but this was coupled with sensitivity to fair assignments of economically valuable cities and ports, and a sense of whether the merger made historical and cultural sense. However, the fault-lines of this process are particularly apparent along the borders of the new states which were invariably multi-lingual and often with a mixed linguistic culture or identity. It is in these inevitable "mistakes" on the border of the modern South Indian states where we focus our paper.

Borders of the modern South Indian states overlaid on the old political configurations can be seen in Map 1. Along the borders there are districts that belonged to the same political entity prior to 1956, but were assigned by the Commission to different states. The villages along the modern border not only share a common geography and climate, they also share a common history – having belonged to the same political and administrative entity for over two hundred years. Thus, if we consider the arguments of Bayly (1999) and Dirks (2002), shared administrative and political histories should have caused their caste structures to be similar. In particular, till 1956, the villages had a common history of land tenure, administration, and reform dating back to, at least, the Mughal period. Since the distribution and control over land, particularly the prevalence of landlessness, are closely related to caste (Kumar 1962, Kumar 1992) the caste structures in border villages should be very similar.

The other fundamental determinant of caste structure is language. Jatis are endogamous groups – groups that are defined by closed marriage and kinship circles. Social norms dictate that grooms

and brides must belong to the same jati. Marriage and kinship circles in India, and particularly in South India, are defined within linguistic groups (Trautman, 1981), therefore caste systems are also defined within linguistic groups. Since our sample is located in the borders of linguistically defined states – there is considerable overlap between the languages spoken on either side of the border. We select blocks (sub-district level entities that are approximately equivalent to counties) on either side of the border matched by the mother-tongue of the majority of people in each block. Within these matched blocks we compare differences among villages, also matched by mother-tongue, across the border with comparable villages on the same side of the border. The specifics of the method used for matching are given in the next section, but this language-matching allows us to control for similar language and kinship structures.

One concern with this method is the possible role that cross-border migration could play in influencing the results. However, as Munshi and Rosenzweig (2005) and other scholars have noted, migration rates in rural India are rather low. Most migration is for the purpose of marriage, with women moving from their natal families to their husband's home. In the South, moreover, kinship rules require that marriages are arranged between families that are of the same jati and speak the same language (Trautman, 1981). Consequently, even if women were crossing borders to marry they would be marrying families within the same jati. This would raise the likelihood of finding similar caste structures across the border, and work towards confirming our null hypothesis. Another possible manner that migration could matter is if, in 1956, families who found themselves on the "wrong" side of the border migrated to the linguistically proximate neighboring state. For instance a Kannada speaker who was resident in Telugu–speaking AP after 1956 could have moved across the border to Karnataka. Since our design works with the best linguistic match between blocks, and between villages, it would exclude areas that had a significant proportion of such migrants.

The core idea behind the natural experiment can be understood by looking at Map 1. The Madras Presidency and Hyderabad state are the two old administrative units that are relevant for our analysis. Within these old states we pick 7 pairs of districts that were later split into different states after the reorganization. These are Bidar and Medak in Hyderabad, Dharmapuri/ Chitoor, Kasaragod/Dakshina Kanada, and Coimbatore/Pallakad in different parts of Madras Presidency. Bidar and Dakshina Kanada are now in the state of Karnataka, Medak and Chitoor are now in AP,

Dharmapuri and Coimbatore in Tamil Nadu, and Pallakad and Kasaragod in Kerala[16]. Within these districts we pick a set of blocks using the language matching strategy, and then a set of villages, randomly selected within each block, which are also matched by language – details about the sampling and matching process follow below. Thus, we are looking at a large sample of villages that share an administrative history, a shared language, and are geographically very proximate. Hence, by all logic, they should have similar caste structures. But, as we will show below – they do not.

Data and Methodology

Sampling

In order to select the blocks within these districts that were best matched on language, we compute the linguistic distance[17] for all combinations of blocks in each district pair. To choose the best matched block pairs we rank all the pairs and select the top ranked pairs – stopping when we find three (two for the Kerala - Tamil Nadu border) unique pairs for each district pair. Table 1 presents the summaries of the block matching process. Along 3 out 4 of our borders, we are able to find very well-matched blocks, indicating that the borders separated linguistically homogenous groups. The quality of the match is not as good along the Kerala – Tamil Nadu border.

The blocks are divided into several Gram Panchayats (GPs) or village government units -- each of which consists of between 1 and 6 villages depending on the state[18]. From each sampled block, in the states of AP, KA and TN, we randomly sampled 6 GPs in every block. In Kerala the

[16]We also sampled Kolar district which is the one exception to the block matching rule in our sample. Kolar was a part of erstwhile Mysore state the precursor to modern Karnataka and thus does not follow the colonial- rule matching process described above. Consequently, we do not use the data from Kolar in this analysis.

[17] The linguistic distance is the weighted sum of absolute differences in proportions of the languages spoken, as mother tongues, in the village/block. The weights are the proportion of the language spoken in both villages/blocks taken as a whole. The values for this measure range from 0 to 1, with zero being the best match possible. Algebraically, let l_{i1}, l_{i2}, be the proportion language i is spoken, as mother tongue, in village 1, and respectively 2. Let p_1 and p_2 be the population in village 1, 2. Then:

$$L(v_1, v_2) = \sum_i |l_{i1} - l_{i2}| \frac{l_{i1}p_1 + l_{i2}p_2}{p_1 + p_2}$$

.

[18] The sample was originally designed for a study of village governance in India.

population per GP is roughly double that in the other three states. For this reason in Kerala we sampled 3 GPs in every block. This procedure gave a total of 201 GPs. From these we selected a village sample. In AP, Karnataka and Tamil Nadu we sampled all villages if the GP had 3 or fewer villages[19]. We excluded all villages with less than 200 persons from our sampling frame. All hamlets with population over 200 were considered as independent villages in drawing the sample. In Kerala we directly sampled wards[20] instead of villages (as villages in Kerala tend to be very large) -- we sampled 6 wards per GP. This gave us a final village sample size of 527 villages

From every sampled block in AP, KA and TN we randomly selected 3 of our 6 sampled GPs and conducted household interviews in all sampled villages falling in these GPs. In Kerala we randomly selected 2 GPs in one block and one GP in the other block (the selection of which block to sample how many GPs from was also random), and within sampled GPs we conducted household interviews in all sampled wards. Twenty households were sampled at random from every selected village, of which four always belonged to Scheduled Caste or Tribes.

The complete sample has been used for other analyses (e.g. Besley et al. 2004), but for the purposes of this study we removed Kolar district and the blocks matched to its blocks. The reason for this elimination is that Kolar district is not matched historically to any of our districts. We further eliminated the blocks without household surveys, as the household surveys enable us to match villages by language – which we need for this analysis. Hence, the sample for this study consists of 143 villages, containing 2950 households

Data Collection

Our data on caste structures- the names and number of castes in each village in the sample – comes from a focus group discussion with 8-10 individuals. These individuals were selected to represent the main social groups in the village and for being knowledgeable about village life. The focus group discussion was led by a trained and experienced moderator who asked the group to list the names of all the castes in the village and to specify approximately how many households belonged to each caste. The method works well in Indian villages which are settlements with very long histories and low rates of migration. Consequently information of this

[19] If it had more than three villages, then we selected the village to which the president of the gram panchayat belonged and randomly selected two other villages.
[20] A sub-unit of a village.

kind is usually common knowledge[21]. Focus groups allow the possibility of poor information from any one individual to be cross-checked by the others in the group till a consensus on a response is reached. For questions on facts about village life that are common-knowledge it can therefore produce very good information.

The household survey, in addition to asking a variety of socio-economic questions, asked each household to identify its mother-tongue and all the languages spoken by the head of the household. We also asked an open-ended question about the household's caste – which allows caste identity to be identified without any prompting from the questionnaire.

Methodology

Estimating the effects of the State Border Natural Experiment:

Let adjacent states be denoted as *s* and *r*. (In our sample we have the adjacent states Andhra Pradesh-Tamil Nadu (AP-TN), AP-Karnataka (AP-KA), KA-Kerala (KA-KE) and KE-TN). Following our sampling strategy the villages in the border areas come from pairs of blocks that are similar in the frequency with which they speak languages. To obtain an even more precise language-congruence we do a further match - this time at the village level. To get the frequency of different languages in each village we use data from the household sample survey and calculate the linguistic distance at the village-level from it. We pair each village on the *s*-*t* border area with the closest village, in terms of linguistic distance, within the same state and in the bordering state. Table 2a, presents the summary of linguistic distance for our village pairs. Our village matches are particularly close along the AP – TN, AP – KA, and KA – KE border. The high degree of this linguistic proximity supports our identifying assumption that when the political border was drawn – some villages proximate to the border on both sides were exogenously separated.

Our unit of observation is a pair of villages. Let v_{is} and v_{jr} be such a pair of villages. We then estimate the following regression:

[21] As a test of this we asked the focus group to provide a number for the total population of households in the village. The 2001 Indian census also provides village population as an independent source of information. The correlation between these two numbers is 0.83 despite the fact that FGD data is at the household level and census data is at the individual level, and the possible problem that the census data uses a "revenue" village definition that is not always exactly the same as the "organic" villages we sample in our survey.

$$F(v_{is}, v_{jr}) = \alpha + \gamma_{sr} + \delta I(s \neq r) + \varepsilon_{isjr} \quad (1)$$

Where:

F(v_{is} , v_{jr}) is the function of interest. In all but one of our estimations the function is the absolute difference in a village characteristic. For example, when we are interested in the effect of the state border on the number of castes in the village, the function is the absolute difference between the numbers of castes in the two villages. When we are interested in the incidence of landlessness it is the proportion landless in each village. In one specification of equation (1), instead of the absolute difference we use the caste overlap between the two villages. The caste overlap is measured as the ratio of the population in castes that exist in both villages to the total population of the two villages. If the two villages have identical castes the overlap is equal to one, if they have no castes in common the overlap is equal to zero. An alternate measure of caste overlap is constructed by looking at the population in the 5 most numerous castes in each village and computing the overlap only among these castes.

γ_{sr} captures the ***sr*** border fixed effect

I(s≠r) is an indicator for whether the two villages are in different states.

We estimate this equation by OLS, with robust standard errors. The second village in the pair can appear in several pairs and hence we need to control for the correlation induced by this repeated appearance, we do this by clustering standard errors by the second village in the pair. In this equation our estimated parameter of interest is δ . This parameter measures the effect of the state border. For example, when we are interested in the effect of the state border on the number of castes in the village, a negative δ implies that villages on the same side of the border are more similar (in terms of number of castes) than villages on different sides of the border. Conversely, when we are interested in the caste overlap, a positive δ implies a larger caste overlap between villages on the same side of the border than between villages on different sides of the border.

Since we observe the fraction landless at village level both in the 1951 census and in our current 2002 data, we can test whether the between-within state gap changes from 1951 to 2002[22]. We accomplish this by estimating the following equation (2):

$$\left|landless_{ist} - landless_{jrt}\right| = \alpha + \gamma_{isjr} + \beta(t = 2002) + \Delta I(s \neq r) \times I(t = 2002) + \varepsilon_{isjrt} \quad (2)$$

Where:

γ_{isjr} is the village pair fixed effect.

In equation (2) we are interested in estimating Δ. This coefficient captures the extent to which the within-between state gap, in fraction landless, has changed from 1951 to 2002. As we are including village pair fixed effects in this estimation, we are controlling for any time invariant attributes of the village pairs. In addition to estimating equation (2) we also estimate the differences in landlessness in state borders separately for 1951 and 2002 using the specification in equation (1).

Estimating Correlates of Identity Choice:

In order to understand some of the processes that may drive the results we find in the natural experiment we also explore some correlates of identity choice at the individual and village level. One process that we are particularly interested in is, what we call, "caste broadening." These are processes by which narrow caste names get consolidated by individual, or by entire villages, within some broader label. For instance, individuals and villages may choose to identify a group by their language rather than their caste. Or particular castes may be identified as a larger category than "traditional" sub-caste groupings – many scheduled caste groups, for instance, prefer to be identified as "Dalit" which they consider a more empowered term rather than the traditional name by which their group was known. Upper caste Brahmin groups, similarly, may prefer to be identified as "Brahmin" rather than by the specific sub-caste of Brahmin to which

[22] The unit of observation in the 1951 census data is slightly larger than the present day village (as presumably some villages have split between 1951 and 2002). Furthermore, some present day villages could not be located in the 1951 census. To make the 1951 and 2002 data comparable we aggregated the 2002 data up to the level of 1951 villages. We also dropped the 2002 villages which could not be located in 1951.

they belong. The determinants of these broad categories are explored both at the village and the individual level.

In equation (3) we explore the correlates of village-level broad names using information from the focus group caste roster, to estimate the following equation.

$$b_{cvps} = \alpha + \gamma_p + \sigma_s + \phi K_{vps} + \beta_1 S_{cvps} + \beta_2 L_{cvps} + \beta_3 H_{cvps} + \beta_4 U_{vps} +$$
$$+ \theta_1 \sigma_s S_{cvps} + \theta_2 \sigma_s L_{cvps} + \theta_3 \sigma_s H_{cvps} + \theta_4 L_{cvps} U_{vps} + \theta_5 H_{cvps} U_{vps} + \theta_6 S_{cvps} U_{vps} + \varepsilon_{cvps}$$

(3)

Where:

b_{cvps} is an indicator for whether caste *c* in village *v*, block pair *p*, state *s* chooses a broad name,

γ_p, σ_s are, respectively, block pair, and state indicators,

K_{vps} is a matrix of village level variables[23],

S_{cvps} denotes caste category, namely whether the caste is a "Forward" caste, "Other Backward Caste" (OBC), or SC/ST,

L_{cvps} is the fraction land held in the village by caste *c*,

H_{cvps} is the fraction households in the village belonging to caste *c*,

U_{vps} measures fraction of land held by upper castes in the village – a measure of upper caste dominance.

We estimate this equation using a linear probability model. We cluster the standard errors at village level. In specification (i) we set all the θs – the interaction effects - to zero. In specifications (ii) through (vii) we estimate θ_1 through θ_6, one at a time.

We also explore the determinants of whether individual respondents choose to identify themselves by a broad caste category in response to an open-ended question using the following equation:

$$b_{ivps} = \alpha + \gamma_p + \sigma_s + \beta_1 X_{ivps} + \beta_2 V_{vps} + \beta_3 B_{vps} + \varepsilon_{ivps} \quad (4)$$

[23] Total number of households and total land area

Where:

b_{ivps} is an indicator for whether individual *i* in village *v*, block pair *p*, state *s* chooses a broad name to describe his/her caste,

γ_p, σ_s are, respectively block pair and state indicators,

X_{ivps} is a matrix of individual and household level variables[24],

V_{vps} is a matrix of village level variables[25],

B_{vps} is the fraction of households whose caste has chosen a broad name, in village *v*. Note that this fraction is computed from the PRA caste roster. Because this is a potentially endogenous variable, in the first specification we set $\beta_3 = 0$, while in the second specification we estimate β_3.

We estimate this equation using a linear probability model. We cluster the standard errors at village level.

Results:

One of the challenges in checking the validity of our natural experiment is that caste data at the village level hard to come by. In particular, it is not available prior to 1956 when the states were reorganized[26]. However, we do have data on the landlessness at the village level from the 1951 census. Several scholars, notably Kumar (1962), have demonstrated the high degree of correlation between landlessness and caste status. In particular villages with a high proportion of low castes are also likely to have a high proportion of landless families. Landlessness in 1951, therefore, provides a reasonably way of testing whether our "treatment" and "control" villages were similar in their caste structure prior to the "intervention" in 1956.

Table 4a reports results from a regression that estimates equation (1) for the difference in difference in 1951 landlessness, and compares with difference in difference for landlessness in 2002. It is clear from these regressions that landlessness in 1951 was very similar in matched

[24] Individual level variables: gender, age, education; Household level variables: landed dummy (= 1 if the household owns any land), and household size.

[25] Total number of households, total land area, fraction land held by upper castes.

[26] Detailed data on complete caste lists were last collected in the 1931 census, but they are not reported at the village level and hence cannot be used in our analysis The 1951 census collected data on scheduled castes and tribes, but this also is not reported at the village level.

villages across state boundaries in comparison with matched villages within the same state. However, as an indication that the 1956 border change did have an impact, we see in the same table that by 2002 these differences became salient – matches across the border have a greater difference than matches within the border. This is almost certainly because of differences across states in passing land reform legislation, and in the efficacy of their implementation – with Kerala being the most effective (Herring, 2007). In order to test whether the coefficients for the difference in difference in landlessness in 1951 and 2002 are significantly different from each other we report the estimates of equation (2). This estimates the triple difference in landlessness within the same state, across states and across years, in table 4b which shows that the difference in difference in 2002 is larger than the difference in 1951 at a 10% level of significance. Given, the high correlation between landlessness and caste structures these results permit the inference that caste structures in 1951 were very similar in matched villages within and across modern state boundaries because of shared legal, administrative and linguistic systems. By 2002 – the administrative structures in the reorganized states had begun to have an effect, even within linguistically similar villages across the state boundary.

We now examine the relationship between borders and caste structures and to begin with we look at the number of castes in a village. Table 3 shows that the district means of this variable do not show much variation – with exception of Dharmapuri district in Tamil Nadu with a mean of 4.23 castes per village, the other districts have averages that range from 11.56 to 14.55. Table 7 provides estimates of equation (1) for this variable, and we see that the difference in the number of castes in matched villages across state boundaries compared to matched village in the same state is 3.18, and significant at the 1 per cent level. In other words, differences across the border in the number of castes in a village are about 20 per cent greater than differences in caste numbers on the same side of the border.

We next look at the nomenclature of castes – specifically whether the overlap between caste names is greater for villages in the same state or across the state border. Table 2b reports that the caste-overlap on a state boundary ranges from an average of 19 per cent at the AP-TN border to 59 per cent at the Karnataka-Kerala border. Focusing on the five most populous castes within each village we see that that overlap percentage is 15% for the AP-TN border and 51% for the Karnataka-Kerala border. Table 5b provides estimates of equation (1) for this variable. We again see that the overlap in matched pairs of villages across the state border is about 42% less than the

overlap in the same side of the border, significant at the 1% level. Focusing on the most populous five castes in every village, the difference in the overlap is 34% also significant at the 1% level.

These are the key results of the paper. They indicate that the 1956 border change instigated a variety of processes that caused the number of castes, and their names, to change in villages that should have been well matched in their caste structures because of common histories, land distributions, languages and geography. It shows that caste structures are not set in stone but can change as a result of significant political changes. The processes described by Bayly and Dirks seem to apply in post-independence India just as they did during earlier periods of Indian history.

Does our natural experiment have an effect on measures of social diversity and inequality between castes? We begin by examining standard measures of fractionalization and polarization. Table 3 shows that fractionalization is low in one district – Dharmapuri where it is 31 per cent, and relatively high everywhere else, ranging from 61 per cent in Coimbatore to 81 per cent in Medak. Polarization shows less variation – ranging from 40 per cent in Dharmapuri to 67 per cent in Coimbatore. Table 5a reports result of estimates of equation (1) on these measures. The difference in fractionalization is about 46 per cent greater between states than within states, at the 5 per cent level of significance, but the state reorganization does not seem to have had a significant impact on polarization.

Considering the extent to which caste affiliation is expected to be correlated with inequality, we employ decomposable generalized entropy (GE) measures of inequality to measure the proportion of total land inequality that can be explained by inequality between and within castes. Given the high levels of landlessness in the villages we omit GE (α=0) which is very sensitive to values at the lower end of the distribution, and focus on GE(α =1) and GE(α =2). Any castes which report zero land are assigned a value of 0.01 acres since almost all households have legal or squatters rights over, at least, a small plot of land to build a shack. Table 3 reports that 31 per cent of overall land inequality can be explained by between caste differences according to the GE(1) measure and 24 per cent according to the GE(2) measure. Table 5a shows that the difference in the proportion of between caste inequality, using both measures, is significantly higher in villages across states than within them.

Our results have so far shown that the change in political boundaries in 1956 affected the number of castes, the names of castes, fractionalization, and the proportion of land inequality explained

by between-caste inequality. The question that remains to be answered is how did this come about? What forces did the formation of new state borders unleash that caused caste structures to change? This question is hard to answer with any certainty with our data, but we will try to shed some light on it by examining our data and drawing some insights from the literature on affirmative action, political reform, and caste movements in South Indian states. We should emphasize, however, that our efforts here should be considered more speculative than definitive.

In terms of sheer arithmetic – one process that is at work is caste "broadening." This describes the consolidation of diverse castes under a broader label because of a social or administrative process. Social movements to unify castes under a broader label have a long history both among upper castes (e.g. Conlon 1974) and lower castes (Omvedt, 1994). The process may also work via a gradual breakdown in patterns of caste endogamy with sub-castes gradually permitting marriages with other sub-castes with whom they are culturally and socially proximate (e.g. Leonard and Weller, 1980). Political calculations may also be at work with castes mobilizing themselves into broader groups in order to gain political power (Weiner 2001). Affirmative action programs – which share a deep association with caste politics – could work towards caste broadening processes in two different ways: a) by creating systems of categorization that affect nomenclature, much in the way that Dirks describes, and b) by changing the incentives faced by individual household to affiliate themselves with more identifiable caste categories in order to access benefits. Finally, land reform may have an effect by empowering lower caste groups and disempowering landed castes which, in turn, could affect kinship relationships, and social interactions across caste groups and lead to broadening processes both at the low and high ends of the caste distribution.

These processes have worked in different ways in different states. In Kerala, despite a strongly egalitarian communist movement and high levels of literacy, caste awareness remains acute in the private sphere where endogamy remains strong and has a significant effect on gender bias (Sudha et al 2005), and caste-based inequality remains salient (Deshpande, 2001). However, in Kerala's political sphere political-party affiliation matters much more than caste in determining election choices (Besley, Pande and Rao, 2006). Caste categories are not easily used in the public sphere – one could even say that it is considered politically incorrect to publicly refer to someone's caste in Kerala. Moving to the neighboring state of Karnataka, the princely state of Mysore was an early pioneer in raising caste awareness and creating both political and educational affirmative action programs for disadvantaged castes (Bhagavan 2003). Mysore was merged with the

Kannada speaking areas from Hyderabad, Bombay and Madras to form the state of Karnataka, but Mysore continues to dominate the political and cultural identity of the state. Consequently, the caste-consciousness that was achieved in Mysore can now be observed in the annexed areas of Karnataka. In particular, Dalit movements in Karnataka came early and were particularly strong (Omvedt, 1994) resulting in the wide-spread use of broadened caste names like "Adi-Karnataka," "Dalit" and "Harijan" though these terms have sometimes come to denote particular castes (Charsley, 1996).

In AP, on the other hand, despite the nascent rise of a Dalit unification movement, the divisions between the two dominant Dalit communities – Malas and Madigas – remain strong (Srinivasulu, 2002). Tamil Nadu, is a special state which was an early pioneer in the caste-broadening process because of the attention paid to caste categories in the Madras Presidency by colonial rulers and the subsequent mobilization of caste identity around these census identities during the late and early-20th century (Dirks, 2002). Issues of caste identity and affiliation remain salient issues in the state and the process of broadening has continued (Pandian, 1983).

To examine the state differences in the caste broadening process we examined the caste information reported for all the castes in the village, and all the households in the sample, and classified some of them as "broad" categories. Broadness was defined as any category that moved beyond narrow sub-caste affiliations and reflected a broader identity that was either influenced by religion (e.g. Muslim, Jain), influenced by a social movement (e.g. Adi-Karnataka, Dalit), or based on linguistic identity or occupation (e.g. Malayalee, Tenant Landholder). The choices that we made can be seen in the caste lists provided in Appendix A. Table 6 provides summary information for this variable by state and district showing that broad names are more prevalent in Karnataka and Tamil Nadu than in Kerala or AP.

Table 8 reports results from equation (3) – caste-level estimates of broad caste names. Controlling for fixed effects for each matched block-pair we see in column (1) that Karnataka is more likely to have broad caste names than any other state, Scheduled Castes and Tribes (SC/ST) and Backward Castes (BC/OBC) are less likely to have broad caste names than other castes, while poorer castes with less land are more likely have broad names, suggesting that impoverishment rather than discrimination may drive the move towards broader identity. Finally, castes that are more populous in the village are more likely to have broad names – which could be

a reflection of the process by which sub-castes consolidate under broader names and thus acquire numerical clout.

Our major interest, however, is to examine how the effects of these variables that affect caste-broadening vary across states. To examine this we interact SC/ST, BC/OBC, caste land fraction, and caste population fraction with state dummies in columns (2) through (4). Column (2) demonstrates that the narrowing process in SC/ST and Backward Castes significantly differs by state. In particular, SC/STs in Karnataka and Tamil Nadu are more likely to acquire broad names than those in Kerala – the omitted category – perhaps reflecting the fact that Dalit scheduled caste movements in Kerala achieved their impact before 1956, while they have had a more recent effect in Karnataka and Tamil Nadu (Omvedt, 1994). Backward castes in AP and Karnataka, on the other hand, are less likely to have broad caste names than those in Kerala. This maybe because affirmative action programs for backward castes in AP and Karnataka are particularly strong.

Columns (3) and (4) interact state dummies with the proportion of land owned by the caste and the caste's proportion in the village population. Both show that in AP and Tamil Nadu, the interactions result in a narrowing of caste names indicating that that the general trend towards more populous castes having broad names is tempered in these states.

A potentially important variable in the caste-broadening process is the extent to which the village is dominated by upper castes. Upper castes domination can result in a reinforcement of the status-quo by keeping feudal power structures and social norms in place, and thus making it more difficult for social movements to take hold. Column (1) shows that upper caste domination does not have an independent effect on broadening in the village data, but columns (4) and (5) show that it affects interactions with the land fraction of the caste, and with its fraction in the population. In both cases, the interaction with upper caste domination results in less broadening – suggesting that broadening processes slow down in villages which are dominated by upper castes – and are hence more feudal in character.

Having examined the correlates of the prevalence of broad caste names at the caste level, we now turn to household level data. Two processes may be at work here – individual incentives, such as status, that might affect a household's decision to choose a broad identity, and group-based choices if the choice of a broad identity is driven by building coalitions across groups via social movements and affected by administrative technologies. Table 9, column (1) provides estimates

of equation (4) without interactions. The table shows that both individual and social effects matter – the respondent's age, education and household size all raise the probability that s/he reports a broad caste identity, while the probability is reduced if the respondent belongs to a family that owns land. As in the caste-level regressions, households in Karnataka are more likely to report a broad identity. Interestingly, the strongest effect in the regression is from the caste dominance variable – the proportion of land in the village owned by upper castes. As expected, this strongly reduces the probability that the household will choose a broad caste identity.

Column (2) interacts the caste variables with state dummies to see if the scheduled and backward caste effects differ by state. As in the village-level regressions, backward caste status reduces the probability of caste-broadening in all states except Kerala. While scheduled caste/tribe status increases probability of broadening in Tamil Nadu. In column (3) we interact the respondent's land status with state dummies; we see that landed households in AP are more likely to report a broad caste name, while those in Tamil Nadu are less likely to do so. The results in Table 9 show that broadening is affected both by individual and social incentives, but that these effects tend to vary systematically by state.

Thus, our analysis of caste broadening at the village and individual level show that it could be an important factor behind the shift in caste structures after the reorganization of the political boundaries of states in 1956. While it is difficult to pinpoint exactly how this might have worked, our results suggest that caste-based incentives – because of affirmative action access, the effect of social movements, political imperatives to seek more power within village society, in addition to individual incentives, may have played a part in the process. And these effects show systematic variation across states.

Conclusion:

Our goal in this paper is to demonstrate that caste structures – more specifically the type and number of castes within a particular region – are not primordially given. They are a function of political processes. As Bayly (1999) points out, over the centuries, shifts in political control have resulted in shifts in the caste system because of changes in systems of patronage and allegiance. Dirks (2002) specifically looking at British colonial rule makes a compelling case that the British propensity for measurement and administrative control forced standardized categories onto a hitherto fluid system that in turn had important effects on political mobilization – essentially

creating the modern caste system. This paper takes this argument a step further, demonstrating that these changes have continued in the post-independence period – processes as diverse as caste-based social movements, affirmative action – particularly the processes of listing and identifying marginal groups to give them differential access to public programs, state and village level political competition, and other economic and social changes within states, have caused caste structures to nurture and evolve within state boundaries.

We demonstrate this by using a natural experiment, namely the reorganization of state boundaries along linguistic lines in 1956, that shows the following: comparing villages matched on language (which is the basis of kinship groups and hence closely correlated to caste), we find that the names and the number of castes on the same side of state boundary are more likely to be similar than the names and the number of castes in matched villages across the boundary. To cross-check the validity of the instrument we examine landlessness data from these villages in 1951 – since the degree of landlessness is highly correlated with the presence of low castes and hence should have a large influence on caste structures. We find that matched villages on the same side and opposite sides of the 1956 border are no different in the 1951 incidence of landlessness, but are significantly different in landlessness measured in 2002.

We then, more speculatively, try to understand what processes may be driving the shift in caste structures by examining the determinants of caste "broadening," i.e.: the prevalence of caste names at the village and individual level that allow for more categories of sub-castes to fit within them. Examining the village data we find that caste broadening is more prevalent in the states of Karnataka and Tamil Nadu than in Kerala, which is consistent with the ethnographic and historical literature on the nature of caste-based social movements in these states. We find similar patterns at the individual level – the choice of a broad caste name is more prevalent in Karnataka, and among educated and older individuals. However, individuals from families with land are less likely to choose a broad identity as are individuals who belong to villages with more upper caste dominance. Scheduled castes are more likely to choose broad identities in Tamil Nadu, while Backward Castes are less likely to choose them in all states except Kerala. Finally, landed families are more likely to choose broad identities in Andhra Pradesh. Thus, individual choices in broad caste names also vary in systematic ways across states. While it is difficult to pin-point exactly how this might have worked, our results suggest that social movements, and their consequent impact on state administrative technologies (e.g. methods of caste identification),

state policies (e.g. affirmative action), and local politics may have had an important impact on shaping caste structures within these states over the last fifty years.

Therefore, our results show that political change can affect social structures and categories. In particular, they demonstrate that the Indian caste system is not a persistent and stable institution, set in stone for several millennia, as suggested by a large literature across many disciplines, but a malleable institution susceptible to political and economic influences. This suggests that some caution must be applied in treating ethnic categories as exogenous variables. An argument for exogeneity should be made on the basis of whether the dependent variable of interest has a faster rate of change than the social diversity that the analyst is attempting to correlate with it.

-------xx-------

References

Acemoglu, Daron, Simon Johnson, James A. Robinson, "The Colonial Origins of Comparative Development : An Empirical Investigation, " *American Economic Review*, Vol. 91, Pp: 1369-401, 2001

Acemoglu, Daron, Simon Johnson, James A. Robinson, "Institutions as a Fundamental Cause of Long-Run Growth," Chapter 6 in Philipe Aghion and Steven N. Durlauf (editors) *Handbook of Economic Growth*, Vol 1A, Elsivier BV, 2005

Akerlof, George A. "The Economics of Caste and of the Rat Race and Other Woeful Tales," *The Quarterly Journal of Economics*, Vol. 90, November 1976, Pp; 599-617

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *The Quarterly Journal of Economics*, MIT Press, vol. 115(3), pages 715-753, August 2000

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, Vol. XLII, Pp: 762-800, September 2005

Appadurai, Arjun "Is Homo Heirarchicus?," *American Ethnologist*, Vol. 13, No. 4. , November 1986, Pp: 745-761.

Banerjee, Abhijit, and Rohini Somanathan, "The Political Economy of Public Goods: Some Evidence from India," *Journal of Development Economics*, 2006

Banerjee, Abhijit and Lakshmi Iyer, "History, Institutions and Economic Performance: the Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95, no. 4 (September 2005): 1190-1213.

Bhagavan, Manu, *Sovereign Spheres: Princes, Education and Empire in Colonial India*, Oxford University Press, New Delhi, 2003

Bayly, Susan, *Caste, Society and Politics in India: From the Eighteenth Century to the Modern Age*, Cambridge University Press, Cambridge, 1999

Bloch, Francis, and Vijayendra Rao, "Statistical Discrimination and Social Assimilation," *Economics Bulletin*, 10(2), 2001, Pp: 1-5

Caselli, Francesco and Wilbur John Coleman, "On the Theory of Ethnic Conflict," mimeo, March 2006

Cohn, Bernard S., "The Census, Social Structure, and Objectification in South Asia," Chapter 10 in *An Anthropologist among the Historians and Other Essays*, Oxford University Press, Delhi, 1987

Conlon, Frank F., "Caste by Association: The Gauda Sarasvata Brahmana Unification Movement," *The Journal of Asian Studies*, Vol. 33, No. 3, Pp: 351-365, May 1974

Das, Veena, *Structure and Cognition: Aspects of Hindu Caste and Ritual*, Second Edition, Oxford University Press, Delhi, 1982

Deshpande, Ashwini, "Caste at Birth? Redifining Disparity in India," *Review of Development Economics*, 5(1), 2001, 130-144

Deshpande, Ashwini, "Does Caste Still Define Disparity: A Look At Inequality in Kerala, India," *American Economic Review Papers and Proceedings*, Vol. 90, No.2 , Pp: 322-325, May 2000

Dumont, Louis. *Homo Hierarchicus: The Caste System and its Implications* (Originally published in French in 1966), University of Chicago Press, Chicago, 1980

Dirks, Nicholas B., *Castes of Mind: Colonialism and the Making of Modern India*, Princeton University Press – NJ, and Permanent Black - Delhi, 2002

Dubois, Abbé J. A., *Hindu Manners, Customs, and Ceremonies*, (translated by Henry K. Beuchamp), Clarendon Press, Oxford, 1906 (first published in French in 1806)

Engerman, Stanley L. and Kenneth L. Sokoloff, "Institutions and Non-Institutional Explanations of Economic Differences" NBER Working Paper 9989, September 2003

Freitas, Kripa, "The Indian Caste System as a Means of Contract Enforcement," mimeo, Northwestern University, October 2006

Government of India, *Report on the States Reorganization Commission (Pan*, Manager of Publications, Government of India, Delhi, 1955

Government of India, *Census of India 1951*

Harrison, Lawrence E, and Samuel Huntington, *Culture Matters: How Values Shape Human Progress*, Basic Books, 2001

Heller, Patrick, *The Labor of Development: Workers And The Transformation of Capitalism In Kerala, India*, Cornell University Press, Ithaca, 1999

Herring, Ronald L., *Agrarian Reform for a Liberal Pattern of Society? Karnataka's Land Policy and Poverty Alleviation*, in G, Kadekodi, R. Kanbur, V. Rao (editors), *Development in Karnataka: Challenges of Governance, Equity and Empowerment*, Academic Foundation Press, New Delhi, 2007

Hoff, Karla and Priyanka Pandey, "Belief Systems and Durable Inequalities: An Empirical Investigation of Caste," World Bank Policy Research Working Paper 3351, June 2004

Khare, R.S., *Caste, Hierarchy and Individualism: Indian Critiques of Dumont's Contributions*, Oxford University Press, Delhi, 2006

Kijama, Yoko, "Caste and Tribe Inequality: Evidence from India, 1983-1999," *Economic Development and Cultural Change*, 54(2), January 2006, Pp: 369-404

Kumar, Dharma, "Caste and Landlessness in South India," *Comparative Studies in Society and History*, Vol. 4, No. 3, Pp: 337-363, April 1962

Kumar, Dharma, *Land and Caste in South India*, Manohar Publishers, New Delhi, 1992

La Ferrara, Eliana, "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana," *American Economic Review*, 93(5), 2002, Pp: 1730-51

Lal, Deepak, *The Hindu Equilibrium: Volume 1: Cultural Stability and Economic Stagnation c. 1500-1980*, Oxford, Clarendon Press, 1989

Leonard, Karen and Susan Weller, "Declining Sub-caste Endogamy in India: The Hyderabad Kayasths," *American Ethnologist*, Vol. 7, No. 3, Pp: 504-517, August 1980

Olson, Mancur, *Rise and Decline of Nations: Economic Growth, Stagflation and Social Rigidities,* Yale University Press, New Haven, 1982

Munshi, Kaivan and Mark R. Rosenzweig, "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy." *American Economic Review* 96(4):1225-1252, September 2006

Munshi, Kaivan and Mark R. Rosenzweig, "Why Is Mobility In India So Low: Social Insurance, Inequality and Growth, mimeo, July 2005

Miguel, Edward, "Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania," *World Politics*, 2004, 56 (3), 327-362

Miguel, Edwards and Daniel N. Posner, "Sources of Ethnic Identification in Africa," mimeo, January 2006

Mill, James, *The History of British India*, James Maddon, London, 1820

Omvedt, Gail, *Dalits and the Democratic Revolution: Dr. Ambedkar and the Dalit Movement in Colonial India,* Sage Publications, Delhi, 1994

Pandian, J., "Political Emblems of Caste Identity: An Interpretation of Tamil Caste Titles," *Anthropological Quarterly*, Vol. 56. No. 4, Pp: 190-197, October 1983

Rao, Vijayendra and Michael Walton, "Culture and Public Action: Relationality, Equality of Agency and Development," Chapter 1 in (V. Rao and M. Walton – eds.) *Culture and Public Action*, Stanford University Press, Stanford, 2004a

Rao, Vijayendra and Michael Walton, editors, *Culture and Public Action,* Stanford University Press, Stanford, 2004b

Risley, H.H., *The People of British India,* W. Thacker, London, 1915

Sahlins, Peter, *Boundaries: The Making of France and Spain in the Pyrenenees*, University of California Press, Berkeley, 1989

Said, Edward W. *Orientalism.* New York: Pantheon Books, 1978

Srinivas, MN, *Social Change in Modern India*, University of California Press, Berkely, 1966

Srinivasulu, K, "Caste, Class and Social Articulation in Andhra Pradesh: Mapping Differential Regional Trajectories," Working Paper 179, Overseas Development Institute, London, September 2002

Sudha, S., S. Khanna, S. Irudaya Rajan, Roma Srivatsava, "Traditions in Transformation: Gender Bias among the Niars of Kerala," Paper presented at the Seminar on Female Deficit in India: Trends and Perspectives, Singapore, December 2005

Trautman, Thomas R, *Dravidian Kinship*, Cambridge University Press, Cambridge, 1981

Weber, Max, *The Religion of India*, Free Press, 1996

Map 1

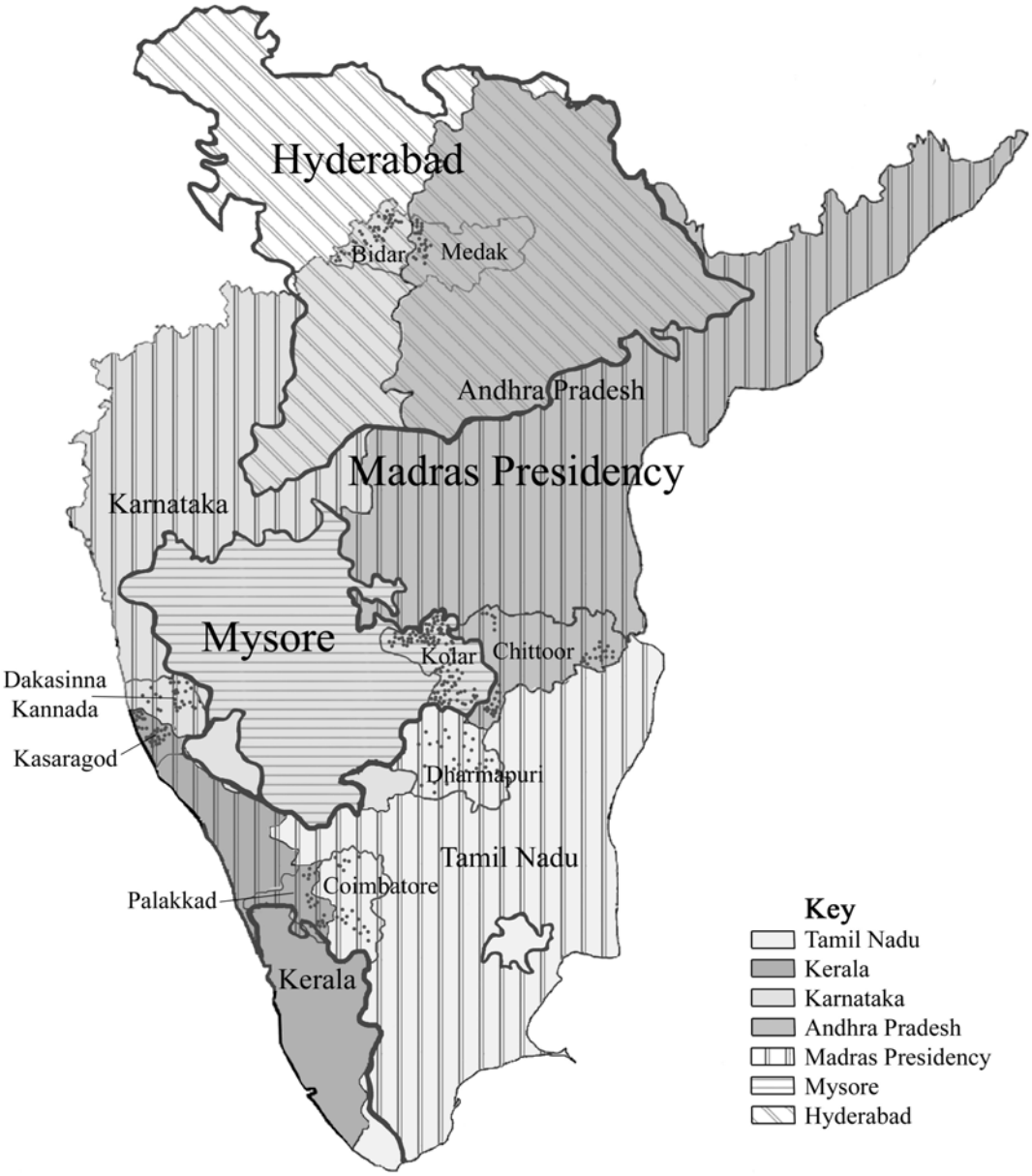
Hyderabad
Bidar
Medak
Andhra Pradesh
Madras Presidency
Karnataka
Mysore
Kolar
Chittoor
Dakasinna Kannada
Kasaragod
Dharmapuri
Tamil Nadu
Coimbatore
Palakkad
Kerala
Key
Tamil Nadu
Kerala
Karnataka
Andhra Pradesh
Madras Presidency
Mysore
Hyderabad

Table 1. Summary of block matching

Border	Total Nr of block pairs	Linguistic Distance				
		Mean	Median	Min	Max	SD
Andhra Pradesh - Karnataka	225	0.4660	0.4627	0.1480	0.6150	0.0855
Karnataka - Kerala	16	0.4747	0.4342	0.1579	0.8470	0.2142
Kerala - Tamil Nadu	45	0.7405	0.7373	0.6297	0.9935	0.1279

Table1b. Actual block matching

	Linguistic Distance
Andhra Pradesh - Tamil Nadu	
pair 1	0.1570
pair 2	0.2910
pair 3	0.3356
Andhra Pradesh - Karnataka	
pair 1	0.1480
pair 2	0.2025
pair 3	0.3170
Karnataka - Kerala	
pair 1	0.1579
pair 2	0.2598
pair 3	0.3392
Kerala - Tamil Nadu	
pair 1	0.6297
pair 2	0.6939

Table 2a: Summary of linguistic distance between villages

	Mean	Std. Dev.	Freq.
AP-TN border	0.1221	0.1309	70
AP-KA border	0.1165	0.1074	68
KA-KE border	0.1513	0.1748	74
KE-TN border	0.2659	0.2652	74
Total	0.1655	0.1908	286

Table 2b: Summary of caste overlap between villages

	all castes		top 5 castes		
	Mean	Std. Dev.	Mean	Std. Dev.	Freq.
AP-TN border	0.1889	0.3341	0.1502	0.3078	70
AP-KA border	0.4716	0.3004	0.3827	0.2938	68
KA-KE border	0.5896	0.2318	0.5114	0.2521	74
KE-TN border	0.3639	0.3393	0.2728	0.3503	74
Total	0.4051	0.3366	0.3306	0.3301	286

Table 3: Summary of Number of village level variables

			nr of castes		population		frac b/w caste GE(1)		frac b/w caste GE(2)		fractionalization		polarization	
State	District	Freq.	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Andhra Pradesh	CHITTOOR	9	11.56	3.05	1300	721	0.34	0.19	0.26	0.16	0.66	0.21	0.64	0.19
	MEDAK	11	14.55	4.06	1947	1786	0.43	0.12	0.40	0.19	0.81	0.08	0.53	0.14
Karnataka	BIDAR	23	13.48	7.24	1947	1786	0.37	0.15	0.30	0.15	0.66	0.16	0.64	0.11
	DAKASINNA KANNADA	13	12.85	4.67	4546	3148	0.35	0.15	0.32	0.15	0.78	0.07	0.58	0.13
Kerala	KASARAGOD	24	12.42	2.92	2160	380	0.27	0.17	0.20	0.16	0.75	0.14	0.61	0.13
	PALAKKAD	24	11.63	3.27	2106	480	0.33	0.17	0.21	0.17	0.73	0.14	0.62	0.13
Tamil Nadu	COIMBATORE	13	13.08	4.21	1403	1135	0.39	0.22	0.26	0.21	0.66	0.15	0.67	0.15
	DHARMAPURI	26	4.23	2.63	787	485	0.14	0.20	0.13	0.22	0.31	0.29	0.40	0.30
	Total	143	11.17	5.38	1893	1604	0.31	0.19	0.24	0.19	0.65	0.24	0.58	0.19

		Fraction landless 1951			Fraction landless 2002		
State	District	Freq.	Mean	Std. Dev	Freq.	Mean	Std. Dev
Andhra Pradesh	CHITTOOR	7	0.45	0.23	9	0.49	0.14
	MEDAK	9	0.47	0.12	11	0.17	0.17
Karnataka	BIDAR	19	0.43	0.20	23	0.22	0.15
	DAKASINNA KANNADA	10	0.40	0.08	13	0.15	0.13
Kerala	KASARAGOD	21	0.55	0.13	24	0.37	0.16
	PALAKKAD	24	0.84	0.04	24	0.69	0.22
Tamil Nadu	COIMBATORE	10	0.71	0.14	13	0.73	0.18
	DHARMAPURI	21	0.27	0.19	26	0.28	0.27
	Total	121	0.53	0.24	143	0.39	0.28

Table 4a: Difference in difference in fraction landless

	within state difference	between states difference	D in D
1951	0.1717	0.1757	0.0090
	(0.1438)	(0.1337)	(0.0304)
	55	63	
2002	0.1292	0.1975	0.0703
	(0.1119)	(0.1460)	(0.0340)**
	55	63	
Table 4b:			
D in D in D			0.0642
			(0.0342)*

Note:

1)Cell size indicated at bottom of each cell

2) The D in D estimations include border fixed effects, standard errors clustered by 2nd village in pair

3) The D in D in D estimation includes village pair fixed effects, standard errors clustered by village pair

4) * 10% significance, ** 5% significance

5) The one sided test, D in D in D < 0 can be rejected at 5% significance

Table 5. Ethnic diversity measures : difference in difference based on best matched villages

	within state	between states	D in D	
a. Absolute differences in village level variables				
Nr castes	3.1818	6.0350	2.8531	
	(3.0573)	(4.4995)	(0.7467)	***
Fractionalization (ELF)	0.1554	0.2280	0.0726	
	(0.1564)	(0.2325)	(0.0334)	**
Polarization (RQ)	0.1563	0.1906	0.0344	
	(0.1442)	(0.1605)	(0.0246)	
Fraction b/w caste inequality (GE 1)	0.1555	0.2332	0.0777	
	(0.1681)	(0.1633)	(0.0349)	**
Fraction b/w caste inequality (GE 2)	0.1617	0.2371	0.0754	
	(0.1795)	(0.1488)	(0.0315)	**
b. Caste name matching metric				
Caste overlap (all castes)	0.6157	0.1945	-0.4212	
	(0.2991)	(0.2204)	(0.0439)	***
Caste overlap (top 5 castes)	0.5015	0.1598	-0.3417	
	(0.3269)	(0.2303)	(0.0459)	***
N	143	143	286	

Note:
1) The D in D estimations include border fixed effects
2) D in D standard error clustered by 2nd village in the pair in parenthesis
3) * 10% significance, ** 5% significance, *** 1% significance

Table 6: Summary of broad caste names at village level

State	District	Fraction broad names	Freq
Andhra Pradesh	CHITTOOR	0.08	248
	MEDAK	0.09	313
Karnataka	BIDAR	0.17	663
	DAKASINNA KANNADA	0.16	347
Kerala	KASARAGOD	0.11	634
	PALAKKAD	0.08	416
Tamil Nadu	COIMBATORE	0.12	352
	DHARMAPURI	0.11	222
	Total	0.12	3195

Table 7: Summary of household level variables

		Broad caste name	female	age		education		Landed	household size	
	N			mean	sd	mean	sd		mean	sd
state										
Andhra Pradesh	416	0.1130	0.4952	34.6178	11.5575	3.3534	4.4383	0.6394	6.0889	2.8368
Karnataka	732	0.2514	0.4973	34.4549	12.5379	5.4303	4.7581	0.6940	5.8593	2.4873
Kerala	996	0.1365	0.4920	44.1647	14.1631	5.5070	4.3172	0.3333	5.6135	2.6108
Tamil Nadu	806	0.1141	0.4975	36.6948	11.0024	3.7878	4.4194	0.5372	4.4677	1.8221
Total	2950	0.1556	0.4953	38.3681	13.2811	4.7146	4.5650	0.5217	5.4285	2.4987

Table 8. Correlates of caste level broad caste name

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Andhra Pradesh	-0.033	0.012	-0.027	-0.018	-0.032	-0.031	-0.032
	(0.021)	(0.035)	(0.024)	(0.025)	(0.021)	(0.021)	(0.021)
Karnataka	0.125***	0.140***	0.131***	0.126***	0.126***	0.126***	0.125***
	(0.020)	(0.039)	(0.022)	(0.023)	(0.019)	(0.019)	(0.020)
Tamil Nadu	-0.013	-0.074**	0.052***	0.074***	-0.011	-0.011	-0.012
	(0.015)	(0.031)	(0.017)	(0.018)	(0.015)	(0.015)	(0.015)
SC/ST	-0.120***	-0.217***	-0.118***	-0.120***	-0.121***	-0.121***	-0.146***
	(0.021)	(0.024)	(0.021)	(0.021)	(0.021)	(0.021)	(0.030)
BC/OBC	-0.143***	-0.106***	-0.151***	-0.155***	-0.147***	-0.146***	-0.147***
	(0.015)	(0.024)	(0.015)	(0.014)	(0.015)	(0.015)	(0.022)
Fraction Land Owned by Caste in Village	-0.158**	-0.165**	0.125	-0.182**	0.042	-0.164**	-0.148*
	(0.077)	(0.078)	(0.114)	(0.077)	(0.091)	(0.076)	(0.078)
Fraction of caste in total village population	0.426***	0.420***	0.478***	0.901***	0.340***	0.571***	0.417***
	(0.101)	(0.101)	(0.103)	(0.116)	(0.103)	(0.111)	(0.102)
Fraction land owned by upper castes	-0.030	-0.031	-0.032	-0.034	0.032	0.041	-0.056
	(0.024)	(0.024)	(0.023)	(0.023)	(0.027)	(0.028)	(0.044)
AP*SC/ST		0.043					
		(0.048)					
AP*BC/OBC		-0.117***					
		(0.033)					
KA*SC/ST		0.172***					
		(0.056)					
KA*BC/OBC		-0.093**					
		(0.040)					
TN*SC/ST		0.201***					
		(0.053)					
TN*BC/OBC		0.061					
		(0.044)					
AP*Land Fraction			-0.305**				
			(0.121)				
KA*Land Fraction			-0.190				
			(0.126)				
TN*Land Fraction			-0.655***				
			(0.097)				
AP*Population Fraction				-0.517***			
				(0.145)			
KA*Population Fraction				-0.161			
				(0.161)			
TN*Population Fraction				-0.889***			
				(0.111)			
Frac. upper caste land*Land Frac.					-0.585***		
					(0.096)		
Frac. upper caste land*Pop Frac.						-0.673***	
						(0.130)	
Frac. upper caste land*SC/ST							0.099
							(0.073)
Frac. upper caste land*BC/OBC							0.015
							(0.055)
Adj. R-squared	0.066	0.080	0.082	0.089	0.073	0.072	0.066
N	3190	3190	3190	3190	3190	3190	3190

Notes: 1)Regression includes pair fixed effects
2)Total number of households and total land area included but not reported
3)Standard errors clustered at village level
4)* 10% significance; ** 5% significance; *** 1% significance

Table 9: Correlates of individual broad caste name

	(1)	(2)	(3)
SC/ST	0.074	-0.038	0.073
	(0.050)	(0.026)	(0.048)
BC/OBC	-0.022	0.162***	-0.018
	(0.034)	(0.048)	(0.034)
Fraction upper caste land	-0.238***	-0.239***	-0.234***
	(0.053)	(0.051)	(0.053)
female	0.013	0.012	0.012
	(0.010)	(0.010)	(0.010)
politician	-0.034	-0.023	-0.036
	(0.032)	(0.030)	(0.032)
Age	0.001*	0.001*	0.001*
	(0.001)	(0.001)	(0.001)
Education	0.004*	0.003	0.004**
	(0.002)	(0.002)	(0.002)
Respondent landed	-0.034*	-0.025	-0.039
	(0.017)	(0.016)	(0.024)
Household size	0.007**	0.005**	0.007**
	(0.003)	(0.003)	(0.003)
AP	-0.024	0.232***	-0.118**
	(0.053)	(0.080)	(0.054)
KA	0.126**	0.378***	0.088*
	(0.049)	(0.073)	(0.052)
TN	0.043	0.053	0.071
	(0.038)	(0.052)	(0.045)
AP * SC/ST		-0.118	
		(0.084)	
KA * SC/ST		0.010	
		(0.084)	
TN * SC/ST		0.441***	
		(0.100)	
AP * BC/OBC		-0.385***	
		(0.080)	
KA * BC/OBC		-0.466***	
		(0.073)	
TN * BC/OBC		-0.151***	
		(0.056)	
AP * Respondent landed			0.122***
			(0.043)
KA * Respondent landed			0.053
			(0.045)
TN * Respondent landed			-0.097*
			(0.055)
adjR-squared	0.077	0.186	0.085
N	2950	2950	2950

APPENDIX A

Caste Name	Broad	state
ADHI DRAVIDAR	1	AP
AGNIKULA KSHATRIYA	0	AP
ARE KAPPU	0	AP
ARE KATIKA	0	AP
ATEENDRA	0	AP
BALIJA	0	AP
BANDA	0	AP
BANJARA	0	AP
BOYA	0	AP
BUDGA JANGAM	0	AP
BYAGARA	0	AP
CHAKALI	0	AP
CHENGUNDAR	0	AP
CHRISTIAN - MALA	0	AP
DARZI	1	AP
DOMMARA - SC	0	AP
DRADIDULU	0	AP
GADHIGUDU	0	AP
GANDLA	0	AP
GANLA SEILU	0	AP
GOUD	0	AP
GOWDA/OBC	0	AP
HATAGAR	0	AP
IDIGA	0	AP
JAIN	1	AP
JANGAM	0	AP
JANGAM (BEG)	0	AP
JOGI	0	AP
KAMARI	0	AP
KAMDALI	0	AP
KAMMA	0	AP
KAPU	0	AP
KHATIKA - OBC	0	AP
KSHATRIYA	1	AP
KUMMARA	0	AP
KURUMA	0	AP
LINGAYAT	1	AP
LINGAYAT - BALIJA	0	AP
LINGAYATH - JANGAM	0	AP
MADIGA	0	AP
MADRAS	1	AP
MALA	0	AP
MANGALI	0	AP
MARATHA (CASTE)	0	AP
MARATHA (LANG)	0	AP
MARWADI	0	AP
MOCHI	1	AP
MUDALIAR	0	AP
MUDIRAJ	0	AP
MUNNURU KAPU	0	AP
NAIDU	0	AP
PADMASALI	0	AP
RAJU	0	AP
RATHODE	0	AP
REDDY	0	AP
SALOLLU	0	AP
SANGUNDHA MUDALIYAR	0	AP
TAMMALI	0	AP
TANGAM	0	AP
TELAGA	0	AP
THUTAVALU	0	AP
VADRANGI	0	AP
VALMIKI	0	AP
VELALLA	0	AP
VELAMA	0	AP
VISHWAKARMA - KAMMALA	0	AP
YADAVA	0	AP
YANADI	0	AP
YATAGIRI	0	AP
YERUKULA	0	AP

Caste Name	Broad	state	Caste Name	Broad	state	Caste Name	Broad	state
ADAPPAD	0	KA	KAVER	0	KA	WOO	0	KA
ADI KARNATAKA	1	KA	KOLI	0	KA			
ADLEYA	0	KA	KOMAK	0	KA			
AGASA	0	KA	KOMATGAR	0	KA			
AGER	0	KA	KOMTI	0	KA			
AITHARI	0	KA	KONKANI	0	KA			
BADAGA	0	KA	KORACHA	0	KA			
BAGWAN/MUSLIM	0	KA	KORAGA	0	KA			
BALIJA	0	KA	KORAMA	0	KA			
BANDAGARA	0	KA	KOTTARI	0	KA			
BARBER	1	KA	KUBBALIGA	0	KA			
BEDA JANGAM	0	KA	KUCHIYA	0	KA			
BESTA	0	KA	KULAL	0	KA			
BHATRAJU	0	KA	KULKIL	0	KA			
BHATTA BRAHMINS	0	KA	KUMBARA	0	KA			
BHOVI	0	KA	KUMRAGA	0	KA			
BILLAL	0	KA	KUNTARU	0	KA			
BILLAVA	0	KA	KURUBA	0	KA			
BOYA	0	KA	KURUJARU	0	KA			
BRAHMIN	0	KA	KURULARY	0	KA			
BRAHMINS	0	KA	KURUTARA	0	KA			
BUDBUDKI	0	KA	KURWARY	0	KA			
BUNT	0	KA	LANIGA	0	KA			
BYARI (M)	0	KA	LATBARA	0	KA			
CHILBADA	0	KA	LATVARA	0	KA			
CHRISTIAN	0	KA	LINGAYAT	1	KA			
DAGABARU	0	KA	LINGAYAT - REDDY	0	KA			
DARZI	0	KA	LOORGIES	0	KA			
DASARI	0	KA	MADIGA	0	KA			
DEHARI	0	KA	MAGATHA	0	KA			
DEVADIGA	0	KA	MAGAVALA	0	KA			
DEVANGA	0	KA	MALAIKUDI	0	KA			
DOMMARA	0	KA	MALIVALDRU	0	KA			
EKAMARA	0	KA	MALLKADI	0	KA			
FCHANA	0	KA	MARATAMAH	0	KA			
GAIFIS	0	KA	MARATHA	0	KA			
GANDA	0	KA	MARATI	0	KA			
GANDHI	0	KA	MUSLIM	1	KA			
GANIGA	0	KA	MUSLIM - BAGWAN	0	KA			
GASAYEE	0	KA	MUSLIM - BYARI	0	KA			
GOLLA	0	KA	MUSLIM - SHAFI	0	KA			
GONDALIGA	0	KA	MUSLIM - SHEIK	0	KA			
GONIGA	0	KA	MUSLIM - SYED	0	KA			
GOSAVI	0	KA	MUSLIM-HAJAM	0	KA			
GOWDA	0	KA	NADAPU	0	KA			
GOWNDER	0	KA	NAIK	0	KA			
GUNDA	0	KA	NAJAMA	0	KA			
HAKARAJADARA	0	KA	NAYAK	0	KA			
HALGA	0	KA	NEIKAR	0	KA			
HARIJAN	1	KA	NYAR	0	KA			
HATAPADA LINGAYAT	0	KA	PABAYA	0	KA			
HEGARA	0	KA	PANJEE	0	KA			
HEGDE	0	KA	PARAVAN - SC	0	KA			
HEGGADE	0	KA	RAI	0	KA			
HELAVA	0	KA	RAJANTHA	0	KA			
HOLEYA	0	KA	RAJPUT	0	KA			
HOTHIYAR	0	KA	RAMMA	0	KA			
INORA	0	KA	REDDY	0	KA			
JADAR	0	KA	SAMAGARA	0	KA			
JAIN	0	KA	SAPALIGA	0	KA			
JANGAM	0	KA	SAPHELIGA	0	KA			
JERANA	0	KA	SCHEDULED TRIBE	1	KA			
KABBALIGA	0	KA	SHEREGARA	0	KA			
KADAIYAN	0	KA	SHETTY	0	KA			
KALAL	0	KA	SUDAGADA	0	KA			
KAMATTI	0	KA	SUNNI/MUSLIM	0	KA			
KAMMARA	0	KA	SWATI	0	KA			
KANGAARE	0	KA	TENENT - REDDY	1	KA			
KATBRUGENI	0	KA	UMISARAN	0	KA			
KATTALIGA	0	KA	UPPARA	0	KA			

Caste Name	Broad	state
ADIYAN	0	KE
AGARAM VELAN CHETTIAR	0	KE
AGASA	0	KE
AJEERPEJI	0	KE
AMBATTAN/BARBER	0	KE
ARYA	0	KE
AYITHAR	0	KE
BAHULA	0	KE
BAIDYA	0	KE
BAKUDA	0	KE
BALLAL	0	KE
BANDARI	0	KE
BECHUDA	0	KE
BELILAPADA	0	KE
BHAI	0	KE
BHATT	0	KE
BHATTA BRAHMIN	0	KE
BOVIS MOGEYAR	0	KE
BOWIS MOGEYAR	0	KE
BOYAN	0	KE
BRAHMIN	0	KE
BRAHMIN - GSB	0	KE
BRAHMIN - NAMBUTHIRI	0	KE
BRAHMIN-IYER	0	KE
BUNT	0	KE
CHAKKILIYAN	0	KE
CHAKKYAR NAMBIAR	0	KE
CHARIYAN	0	KE
CHAVALAKKARAN	0	KE
CHEMMAN	0	KE
CHERUMAN	0	KE
CHETTIAR	0	KE
CHETTIES	0	KE
CHETTY	0	KE
CHOPPEN	0	KE
CHRISTAIN - SYRIAN	0	KE
CHRISTIAN	0	KE
CHRISTIAN - CATHOLIC	0	KE
CHRISTIAN - JACOBITES	0	KE
CHRISTIAN - ORTHODOX	0	KE
DEMBISAN	0	KE
DEVANGA	0	KE
DHEEVARA	0	KE
ERAVALLAN	0	KE
EZHUTACHAN	0	KE
GANIKA	0	KE
GATTI	0	KE
GOUNDER	0	KE
HOLEYA	0	KE
IZHAVA	0	KE
IZHAVATHI	0	KE
JAMAYATH	0	KE
JHATTAN	0	KE
JOGI	0	KE
KADAR	0	KE
KADIYAN	0	KE
KAIKOLAN	0	KE
KAKODA	0	KE
KAMMALAN	0	KE
KAMMALAVARGAM	0	KE
KANAKKAN	0	KE
KANANKAN	0	KE
KANIYAN	0	KE
KANNADA	1	KE
KARGPPAN	0	KE
KARNADAK	1	KE
KARNATAK	1	KE
KAVARA	0	KE
KINDARI JOGI	0	KE
KOLLAN	0	KE
KOTEYAR	0	KE
KUDUBI	0	KE
KURAVAN	0	KE
KURUKKAL	0	KE
KURUPPACHAN	0	KE
KUSAVAN	0	KE
MADILA	0	KE
MAGIYAU	0	KE
MAIKAN	0	KE
MAIVYANI	0	KE
MALAYAN	0	KE
MALLANAHA	0	KE
MANATHAN	0	KE
MANNAN	0	KE
MARAR	0	KE
MARATHUVAR	1	KE
MARATI	1	KE
MARAVAN	0	KE
MAVILAN	0	KE
MODYA	0	KE
MOGAVIRAR	0	KE
MOGER	0	KE
MOOLYA	0	KE
MOOTHAN	0	KE
MUPPAR	0	KE
MUSLIM	1	KE
MUSLIM - ROWTHER	0	KE
MUSLIM - SUNNI	0	KE
MUSLIM/SHAFI	0	KE
MUTAN	0	KE
MUTHENARY	0	KE
NAIKER	0	KE
NAIKKAN	0	KE
NAIR	0	KE
NALIKA	0	KE
NALKADAYA	0	KE
NANKEE	0	KE
NARSANNA	0	KE
NARTI	0	KE
NAYAR	0	KE
PADAKANYA	0	KE
PALIYAN	0	KE
PANAN	0	KE
PANCHI	0	KE
PANJI	0	KE
PARAYAN	0	KE
PATHIYAN	0	KE
PATTAGI	0	KE
PATTAK	0	KE
PATTALI	0	KE
PISHARADI	0	KE
POTTER CASTE	0	KE
PRAGER	0	KE
PULAYAN	0	KE
PULLUVAN	0	KE
RADMINI	0	KE
RULAL	0	KE
SAIVA PILLAI	0	KE
SALIA	0	KE
SAVALAKARAN	0	KE
SETTER	0	KE
SETTU	0	KE
SHAFI	0	KE
SHANTHI	0	KE
SHARODI	0	KE
SHAVALAR	0	KE
SHYDAER	0	KE
SUDRA	0	KE
TAMILIAN	1	KE
TARAKAN	0	KE
THANDAN	0	KE
THATTAN	0	KE
THIMMA	0	KE
VADUKAN	0	KE
VADYAKAR	0	KE
VALLUVAN	0	KE
VANIYAN	0	KE
VANIYAR	0	KE
VANNAN	0	KE
VARIYAR	0	KE
VELALLA	0	KE
VELAN	0	KE
VELUTHEDATHU NAIR	0	KE
VETTUVAN	0	KE
VISHWAKARMA	0	KE
YADAVA	0	KE

Caste Name	Broad	state
24 MANAI CHETTY	0	TN
ADHI DRAVIDAR	0	TN
AMBALAKARAR	0	TN
ANDI PANDARAM	0	TN
ARUNTHATHIYAR	0	TN
BAJJALLI KOTTOI	0	TN
BEEMAN KOTTOI	0	TN
BIILAVA	0	TN
BONDIL	0	TN
BOOMATHAN PATTI	0	TN
CHATTADA SRIVAISHNAVA	0	TN
CHETTIYAR	0	TN
CHINNA GOLLAPATTY	0	TN
CHINNAKARIYAN KOTTOI	0	TN
CHINNATHU PALLAM	0	TN
CHRISTIAN - SC	1	TN
CHRISTIANS	1	TN
CHRSITIAN (NADAR)	0	TN
CHRSITIAN - CATHOLIC	0	TN
DADAN KOTTOI	0	TN
DEVENGA	0	TN
DHOBI	1	TN
DOMBARA	0	TN
DORKOUNDER KOTTOI	0	TN
ELAVANDI	0	TN
EX V M KOTTOI	0	TN
GOUDU	0	TN
GOUNDER	0	TN
GOWDA	0	TN
GURKHA	0	TN
IRULAR	0	TN
IRUPATHINALU MANE TELEGU CHETTY	0	TN
KAIKOLAN	0	TN
KAMBALATTAN	0	TN
KAMMA	0	TN
KANNADIGA	1	TN
KAVARA	0	TN
KERALA CHRISTIANS	1	TN
KOKKANDI KOTTDI	0	TN
KONGU VELLALA	0	TN
KRAIYAR	0	TN
KUDUMBAN	0	TN
KUMMARI	0	TN
KURAVAN	0	TN
KURUMBA	0	TN
LABBAI (M)	0	TN
MADARI (M)	0	TN
MALAYALI	1	TN
MALAYALI (KERALITE)	0	TN
MARAVAR	0	TN
MARUTHUVAR	0	TN
MOTTU KOTTDI	0	TN
MOTTU KOTTOI	0	TN
MUDALIAR	0	TN
MUPAR	0	TN
MUSLIM - RAWTHER	0	TN
MUTHURAJA	0	TN
NADAR	0	TN
NAGARATHA	0	TN
NAICKER	0	TN
NAIDU	0	TN
NATTAR KOTTOI	0	TN
NONDI KOTTOI	0	TN
ODDAR	0	TN
OLAPALLI	0	TN
PALLAR[SC]	0	TN
PAMBAKARAR	0	TN
PANDARAM	0	TN
PANNADI	0	TN
PARA KOTTDI	0	TN

Caste Name	Broad	state
PARAYAN	0	TN
PATHAN (M)	1	TN
PATTARIYAR	0	TN
PERIYA GOLLPATTY	0	TN
PERIYAKARIYAN KOTTOI	0	TN
PULUVAR	0	TN
REDDIAR	0	TN
SAVULU KOTTDI	0	TN
SEMBADAVAN	0	TN
SRI LANKAN	1	TN
TENANT LANDHOLDER	1	TN
TEWAR	0	TN
THIRUPPATHI VOOTTU KOTTOI	0	TN
THORAIYAR	0	TN
THOTTI	0	TN
UDAIYAR	0	TN
UPPILIAN	0	TN
VADUGAN	0	TN
VALLUVAR	0	TN
VALMIKI	0	TN
VALTER	0	TN
VANNAR	0	TN
VANNIAR	0	TN
VANNIYAKULA KSHATRIYA	0	TN
VELLALA	0	TN
VELLALLA	0	TN
VISHWAKARMA	0	TN
VYSYA	0	TN
YEDAVA	0	TN
YEGALI	0	TN