

Insiders and Outsiders: Local Ethnic Politics and Public Good Provision *

Kaivan Munshi[†] Mark Rosenzweig[‡]

February 2016

Abstract

A large and growing literature has taken the position that ethnic politics plays a major role in both limiting the supply of public goods and distorting their allocation in many developing countries. We examine the role of local ethnic polities in supplying public goods within a framework that incorporates two aspects of ethnic groups: an inclusionary aspect associated with internal cooperation and an exclusionary aspect associated with the disregard for others. The inclusionary aspect of ethnic politics results in the selection of more competent political representatives who exert more effort, resulting in an increased supply of public goods that are non-excludable at the local level. The exclusionary aspect of ethnic politics results in the capture of targetable public resources by insiders; i.e. the representative's own group, at the expense of outsiders. Using newly available Indian data, covering all the major states over three election terms at the most local (ward) level, we provide empirical evidence that is consistent with both aspects of ethnic politics, with positive and negative consequences, respectively, for public good provision. Counterfactual simulations using structural estimates of the model quantify the impact of alternative policies that, based on our theory and the empirical results, would be expected to increase the supply of public goods when ethnic politics is salient.

*We are grateful to Toke Aidt and Soenje Reiche for helpful discussions. Brandon D'Souza provided outstanding research assistance. Bruno Gasperini graciously shared the code for the threshold test. Munshi acknowledges research support from the National Science Foundation through grant SES-0617847. We are responsible for any errors that may remain.

[†]University of Cambridge

[‡]Yale University

1 Introduction

There is a large and growing literature that is concerned with the negative effects of ethnic politics on the provision of public goods, economic growth, and development.¹ Ethnic affiliation, on linguistic, tribal, or caste lines, determines the selection of political leaders and the level and distribution of public resources throughout the developing world. One negative consequence of ethnic politics for economic development is that citizens tend to vote mechanically on ethnic lines when information about candidates is limited, resulting in the selection of incompetent or corrupt leaders (Banerjee and Pande 2010, Casey 2015). Both insiders – members of the leaders’ own ethnic group – and outsiders are hurt, as public resources are siphoned off to the leaders themselves. A second consequence of ethnic politics is that it distorts the allocation of public resources. Public resources are targeted along ethnic lines, rather than by efficiency criteria, and insiders benefit at the expense of outsiders (Besley, Pande, and Rao 2007, Bardhan and Mookherjee 2010, Anderson et al. 2015, Burgess et al. 2015).

Our approach to ethnic politics in this paper diverges from the received literature in three ways. First, we examine ethnic politics in the context of local elections, which determine a substantial fraction of total government expenditures around the world and where the problem of limited candidate information is less salient.² Second, we consider two fundamental characteristics that define ethnic groups – their ability to cooperate internally and their disregard for individuals outside their own group. In our framework, when individuals vote for candidates from their own ethnic group they do so because of the benefits they receive and not because of group identity *per se*. Third, we simultaneously consider the competence of elected officials as well as the level and distribution of two types of public resources: public goods that are non-excludable at the local level and targetable welfare transfers. In particular, we examine the role of caste groups in Indian local governments in a framework that incorporates the two fundamental characteristics of ethnic groups, the two types of public resources, and heterogeneity in self-interested politicians’ competence.

Local governments in India are responsible for the construction and maintenance of public infrastructure and for the identification of welfare recipients. Their decisions affect 65% of the total Indian population. Data from the 2006 Rural Economic Development Survey (REDS), which we use in the empirical analysis, documents the importance of caste in local Indian politics. Key informants were asked to list the various sources of financial and organizational support that elected representatives at the most local (ward) level received in each of the last three elections. As described in Table 1, caste is clearly the dominant source of support: 82 percent of the elected ward representatives received support from their caste inside the village and 29 percent received support from caste members outside the village. This support is reflected in electoral outcomes. We see in Table 2 that the elected representative is drawn from the largest caste 90% of the time when it accounts for a majority of the

¹This literature goes back to Easterley and Levine (1997). Recent contributions include Alesina and La Ferrara 2005, Miguel and Gugerty 2005, Padro i Miquel 2007, Burgess et al. 2015, and Alesina et al. forthcoming.

²There has been a worldwide trend towards political decentralization and democratization in recent decades. Local government expenditures account for approximately 25% of total government expenditures in the United States and Europe, and in Brazil approximately 15% of federal government revenue is disbursed to municipalities (Ferraz and Finan 2011). Local governments are responsible for the financing and the administration of the school system in the United States, and in rural India local governments construct and maintain public infrastructure and identify welfare recipients.

ward population, but only 60% of the time when it does not. The next largest caste is represented 23% of the time when no caste has a majority as opposed to just 6% of the time when the largest caste does have an absolute majority.

The main question we address is how the characteristics of caste groups – internal cooperation and the disregard for outsiders – affect the supply and distribution of public resources. A well known reason for the under-supply of public goods, going back to Olson (1965), is that individual volunteers will not internalize the benefits that other individuals derive from public goods, resulting in a sub-optimal level of public good provision. This collective action problem is perhaps most acute in local governments where non-professional politicians must volunteer for important tasks that involve the supply and distribution of substantial public resources. All ethnic communities, regardless of whether they are defined by kinship (tribe or caste), religion, or language, are associated with a high degree of social connectedness. This connectedness supports high levels of internal cooperation. A large and growing literature has documented the role played by community-based networks in supporting private economic activity when markets function imperfectly.³ We assess whether the same communities reduce the collective action problem in local democracies. The basic idea is that social connectedness makes local political leaders internalize the benefit derived from public goods by all members of their ethnic community – the insiders – not just themselves. This could be because members of the community can credibly commit to compensate them *ex post* for their efforts. This results in the selection of more competent leaders and greater effort, which leads, in turn, to greater public good provision.

A second fundamental characteristic of ethnic communities – the flip side of the in-group cooperation described above – is their disregard for others. Political leaders elected with the support of their ethnic group will thus ignore (fail to internalize) the benefit derived from locally provided public goods by outsiders, resulting in a level of public goods that is less than first-best. If public goods cannot be spatially targeted within the local constituency, then everyone is nevertheless better off relative to the benchmark where leaders only value their own benefit from public goods. However, local governments are often entrusted with tasks, such as the distribution of welfare transfers, that can be targeted at specific households. When ethnicity is salient, local political leaders will target welfare transfers to their own group to the extent possible, leaving outsiders strictly worse off.⁴

The theory that we develop incorporates the role played by ethnic communities in the selection of leaders and the supply of two types of public resources: non-excludable public goods and publicly financed welfare transfers to households. In our theory, each political constituency consists of multiple differently-sized ethnic groups. We begin by focusing exclusively on the provision of non-excludable public goods whose level depends on the amount of effort exerted by the elected representative. The

³Greif's (1993) analysis of the Maghribi traders' coalition and Greif, Milgrom, and Weingast's (1994) investigation of the medieval merchant guild highlight the role played by non-market institutions in solving commitment problems in the pre-modern economy. In the contemporary economy, a voluminous literature documents high levels of risk-sharing in informal mutual insurance arrangements throughout the world (e.g., Townsend 1994, Grimard 1997, Ligon, Thomas, and Worrall 2002, Fafchamps and Lund 2003, Mazzocco and Saini 2012, Angelucci, Di Giorgi, and Rasul 2015).

⁴Although the targeting of public goods to co-ethnics has been documented at higher levels of spatial aggregation; e.g. Burgess et al. (2015), we show that such targeting is absent in the local constituencies (wards) that we study. In our analysis, targeting is restricted to welfare transfers.

cost of effort is decreasing in the representative's ability, but more competent individuals also have superior outside options. Consistent with the internal cooperation within ethnic groups and the disregard for outsiders, we allow for *ex post* transfers within but not between groups. Each group will therefore put forward as its representative the individual who maximizes its collective benefit from the public good, net of his effort and opportunity cost. Because the representative of a larger group internalizes the benefit derived by a larger number of individuals, larger groups will put forward more competent representatives who, in turn, provide a higher level of the public good. If the public good is non-excludable, it follows that everyone in the constituency wants the representative of the largest group to be elected. However, this may not be the electoral outcome in practice because groups can free-ride on others. In particular, if a group that is sufficiently close in size puts forward its representative as a candidate, then the largest group may prefer not to contest the election since it will receive a high level of public goods without having to bear its own representative's effort and opportunity cost. Our first theoretical result is that the largest group will, nevertheless, put forward its representative with certainty when its share of the population in the constituency crosses a threshold level, and both insiders and outsiders will benefit from the increased provision of public goods.

As noted, local governments are often responsible for the distribution of welfare transfers. If the total amount of these transfers is fixed, then outsiders will be systematically crowded out, particularly when the representative belongs to a large ethnic group. Adding welfare transfers to our theory affects public good provision in two ways. First, the largest group in the constituency will now be less inclined to free-ride on smaller groups. Second, the representative of the largest group may no longer be preferred by everyone if the negative crowding-out effect associated with the welfare transfers dominates the positive public good effect. Adding welfare transfers to the representative's list of responsibilities will *increase* the level of public goods if the largest group's representative continues to be preferred by everyone, by shifting down the population-share threshold at which it comes to power with certainty. In contrast, if the negative crowding-out effect dominates, then the largest group will only come to power, with an accompanying increase in public good provision, when it has an absolute majority; i.e. when its share of the population exceeds 0.5. The coupling of public good provision and the distribution of welfare transfers now *reduces* the level of public goods.

We test the predictions of the theory with panel data from India across three elections, describing electoral outcomes at the most local (ward) level. This is an ideal setting in which to test theories of ethnic politics for a number of reasons. The caste system is arguably the most distinctive feature of Indian society. The exploitation, prejudice, and discrimination that are associated with the hierarchical structure of the caste system are well known and have been extensively documented. There is, however, another side to this system, which has to do with the solidarity and social connectedness *within* caste groups. Survey evidence indicates that over 95% of Indians continue to marry within their caste. Marriage ties built over many generations result in a high degree of social connectedness within castes, which span a wide area covering many villages, with a small but growing literature documenting the role played by caste communities in supporting economic cooperation.⁵ Within a village, castes

⁵Mutual insurance arrangements have historically been organized, and continue to be organized, around the caste in rural India (Caldwell, Reddy, and Caldwell 1986, Mazzocco and Saini 2012, Munshi and Rosenzweig 2016). When

tend to be spatially clustered. This increases social connectedness within castes, but also expands the already wide social distance between castes. The insider-outsider dichotomy that lies at the heart of our theory of ethnic politics is exemplified by caste relations within the village.

The village governments or *panchayats* that we study have been responsible for the provision of local public goods and the identification of welfare-program recipients since 1991, when a constitutional amendment devolved substantial political power to the local level. A *panchayat* is typically divided into 10-15 wards, each of which comprises approximately 70 households. Political decisions are jointly made by the *panchayat* president and the ward representatives and our assumption will be that the representative's ability and effort determine (in part) the level of public goods that his ward receives. If local caste communities are able to cooperate in the public sphere, as they have traditionally in private economic activity, then the representative's ability and effort will depend on the numerical strength of his caste in the ward. Our theory extends this result to describe the relationship between the population share of the most numerous caste in the ward and the supply of public goods.

A ward with a numerically dominant caste will be less heterogeneous, and there is a large literature documenting the negative correlation between ethnic heterogeneity and the demand for public goods; e.g. Alesina and La Ferrara (2005), Miguel and Gugerty (2005). To disentangle the effect of ethnic affiliation on the supply of public goods, which we are interested in, from this demand effect, we take advantage of the panel nature of our data and the system of set asides in Indian local governments that randomly reserves the ward representative's position for historically disadvantaged castes (and tribes) from one election to the next. This changes the set of castes that can put forward their preferred representative, while leaving the electorate, and the demand for public goods, fixed over time. Net of ward fixed effects, which capture all time invariant ward characteristics, the theory predicts that there should be a discontinuous improvement in the elected representative's competence, which we measure by his education, and an accompanying increase in public goods supplied to the ward when the population share of the most numerous *eligible* caste (which changes over time) crosses a threshold.

Formal statistical tests for a structural break provide support for this prediction of the theory, locating a threshold with a high degree of statistical confidence. Our estimates indicate that when the population share of any one eligible caste crosses this threshold, public goods rise by over 13 percent and the schooling of the elected representative increases by 54 percent. Our estimates also indicate, consistent with the model, that groups of castes do not operate as a single social unit, putting forward candidates who operates on their behalf. Castes that are similarly placed in the social hierarchy; e.g. SC, ST, OBC, have been known to form coalitions at higher levels of government. However, we find that when the population share of any one of these caste-groups crosses the same threshold, there is no increase in the supply of public goods.

A second key finding, based on the threshold's location, is that the most numerous eligible caste is only assured of being in power when it has an absolute majority; i.e. when its population share is at

urban jobs became available in the nineteenth century, with colonization and urbanization, these castes supported the migration of their members and the subsequent formation of urban labor market networks (Morris 1965, Chandravarker 1994, Munshi and Rosenzweig 2006). Castes continue to support the movement of their members into more remunerative occupations in the contemporary economy (Munshi 2011).

least 0.5, despite the fact that it supplies a substantially higher level of public goods. This indicates that the crowding out of the outsiders with respect to welfare transfers dominates the increased supply of public goods; a result that is consistent with a literature going back to Lizzeri and Persico (2001) that shows that electoral outcomes are affected more by targetable transfers than by non-excludable public goods. Providing empirical support for the negative crowding out effect, we find, using data from multiple election terms, that a household is less likely to benefit from programs targeted at disadvantaged households when it shifts from being an insider to an outsider; i.e. when its ward representative belongs to another caste. Moreover, this crowding out is substantially greater when a larger group is in power and is not observed for public goods.

Underlying the threshold test described above are two implications of the theory: (i) there is a discontinuous increase in the probability that the representative of the largest caste is elected at a threshold population share, and (ii) there is a continuous and positive relationship between the size of the representative's caste in the ward and his competence. As shown in Table 2, there is a 30 percentage-point increase in the probability that the largest group's representative is elected when its share of the population exceeds the 0.5 threshold. Providing support for the second implication, we find that representatives of larger castes are systematically drawn from higher up in the education distribution in their caste in the ward, despite the fact that this distribution does not vary with group size.

Education, which is typically treated as a proxy for ability in the economics literature, is positively associated with characteristics such as wealth and experience in managerial occupations that will directly determine the competence of local political representatives. However, given the high opportunity cost of their time, such individuals might also choose to participate less in the political system; indeed, more educated individuals are much less likely to vote at all levels of government in India (Eldersveld and Ahmed 1978). The fact that ward representatives are positively selected on education, particularly when they are drawn from larger groups, is indicative of the caste's ability to solve the individual incentive problem, and the broader collective action problem, that motivate our analysis.

The improved ability of the elected representative and the increased supply of public goods when the largest group comes to power is indicative of the positive – inclusionary – aspect of local ethnic politics. The fact that this group can only come to power when it has an absolute majority; i.e. at a population share of 0.5, is indicative of the negative – exclusionary – aspect. It follows from the theory that decoupling public good provision and the distribution of welfare transfers would shift the population share threshold at which the largest group comes to power down below 0.5, with an accompanying increase in the overall level of public goods. To quantify the benefit of this decoupling, and the efficiency and equity consequences of the caste-based reservation system at the ward level, we estimate the model structurally and perform counter-factual simulations. The results indicate that a combination of policies, as opposed to a single intervention, could substantially increase the supply of local public goods without adverse equity consequences when ethnic politics is salient.

2 Institutional Setting

2.1 Social Structure

The insider-outsider dichotomy is a feature of all traditional societies and is a key assumption in the literature on ethnic politics. The reason why we focus on Indian local governments is that this dichotomy is especially pronounced in Indian society, even at the local level, due to its special caste-based structure. The basic marriage rule in Hindu society is that individuals must match within their caste or *jati*. Muslims also follow this rule, matching within *biradaris* which are equivalent to *jatis*, while converts to Christianity continue to marry within their original *jatis*. Although the population census has not collected caste information since 1931, recent rounds of nationally representative surveys such as the 1999 Rural Economic Development Survey (REDS) and the 2005 India Human Development Survey (IHDS) report that over 95% of Indians continue to marry within their caste (or equivalent kinship community). Newly available genetic evidence indicates that these patterns of endogamous marriage have been in place for over 2,000 years, dividing the Indian population into 4,000 distinct genetic groups, each of which is a caste (Moorjani et al. 2013).

With 4,000 castes, and given India's population of one billion, each caste comprises 250,000 members on average. Social sanctions are needed to support cooperation in such a large group. These sanctions typically involve exclusion from social interactions, which will only be effective when interactions within the group are sufficiently frequent and important, and all castes are intolerant of outsiders. A single caste will thus not have a presence in all villages in the area that it covers. It will instead cluster in select villages. This shows up clearly in the latest, 2006, REDS round, which we use for much of the analysis in this paper. The mean number of castes per state is 64, while the mean number of castes per village is 12. With 340 households on average in a village, this implies that a caste will have about 30 households in the select villages where it locates. This is a large enough number to support meaningful local interactions (and accompanying social sanctions, when required).

While the structure of Indian society results in a high degree of social connectedness within castes, the hierarchical nature of the caste system often gives rise to adversarial relations between castes in the village. Even where castes co-exist in relative harmony, they rarely socialize. Figure 1 uses data from the 2006 REDS village census to describe the distribution of castes and the nature of transactions in a sample of villages drawn from the major Indian states. Focusing on the 91% of villages for which there is ward-level information that can be used for the analysis of ethnic politics that follows, each caste makes up on average 6% of the population of a village. Within the ward, which is a smaller spatial unit, the average caste's share of the population increases to 14%, indicative of the spatial clustering along caste lines that characterizes the Indian village.⁶

The REDS village census asks each household head to list the names of three individuals in the village that he would approach for food, in the event of a temporary shortfall, and for a short-term loan. If these individuals were approached without regard to caste affiliation, we would expect on average

⁶Castes will often live together on particular streets; indeed, streets in Indian villages are informally referred to by caste names. If we examined spatial clustering at the neighborhood level, the proportion of the population belonging to a single caste would likely increase much further.

that 6% of the individuals providing food transfers and loans would belong to the respondent's caste. What we see in Figure 1, based on the identity of the first listed individual, is that around 50% of food transfers and loans are from individuals belonging to the same caste. Munshi and Rosenzweig (2016) document the importance of cross-village loans and transfers in caste-based rural insurance networks. If we included caste loans from outside the village, which make up more than half of caste loans, the fraction of individuals within the caste that would be approached for a loan would increase well above 50%. Informal transactions are evidently concentrated within the caste in the Indian village.

Table 3 indicates that the role of the caste within the village goes beyond supporting informal transactions. REDS respondents were asked who they approached for help when they faced problems with respect to (i) the enforcement of rental and share contracts in agriculture, and (ii) the protection of crops against theft and destruction. Political organizations are the dominant source of support when individuals face contractual problems. However, individuals turn largely to their own caste for support when seeking protection against theft and destruction (from insiders or outsiders). Based on Figure 2 and Table 3, an individual who loses the support of his caste within the village is disadvantaged on multiple important dimensions. The threat of withdrawing this support can thus be used by the caste to sustain high levels of cooperation among its members. Importantly, as seen in both columns in Table 3, other castes, including the dominant caste in the village, play little role in informal transactions or protection, limiting the scope for cooperation between castes.

2.2 Local Politics

The 73rd Amendment of the Indian Constitution, passed in 1991, established a three-tier system of local governments or *panchayats* – at the village, block, and district level – with all seats to be filled by direct election. The village *panchayats*, which often cover multiple villages, were divided into 10-15 wards. *Panchayats* were given substantial power and resources, and regular elections for the position of *panchayat* president and for each ward representative have been held every five years in most states. The major responsibilities of the *panchayat* are to construct and maintain local infrastructure (e.g.; public buildings, water supply and sanitation, roads) and to identify targeted welfare recipients. Although public goods account for the bulk of local government expenditures, publicly funded transfers to individual households, including programs for households below the poverty line (BPL) and employment schemes account for 15% of total expenditures.⁷

The data that we use to relate the caste composition of the ward to public expenditures are unique in their geographic scope and detail. They are from the 2006 Rural Economic and Development Survey (REDS), the most recent round of a nationally representative survey of rural Indian households first carried out in 1968. The 2006 REDS covers 242 of the original 259 villages in 17 major states of India. We make use of three components of the survey data - the village census, the village inventory, and the household survey - for 13 states in which there were ward-based elections and complete data in all components.⁸ The census obtained information on all households in each of the sampled

⁷Although *panchayats* raise their own revenues, through land and water usage taxes, and benefit from specific central government programs, the state government is the major source of funding.

⁸The states are Andhra Pradesh, Bihar, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Karnataka, Kerala,

villages, by ward. This enables us to measure the population share of each caste in each ward.⁹ The village inventory was designed, in part, to specifically assess models of public goods delivery, collecting information on the characteristics of the elected ward representatives and public good provision, at the street level, in each ward in each of the last three *panchayat* election terms prior to the survey. The household survey, administered to a sample of households in each REDS village, records participation in programs intended for households below the poverty line (BPL) in each of the last three election terms.

The village inventory obtained information on whether new construction or maintenance of specific public goods *actually* took place on each street in the village for each term. These local public goods include drinking water, sanitation, improved roads, electricity, street lights, and public telephones as well as schools, health and family planning centers, and irrigation facilities. The survey was designed to permit the mapping of street-level information into wards so that public goods expenditures can be allocated to each ward, and its constituents, for each election term. Ninety-five percent of the wards have information for at least two election terms. Our analysis focuses on six goods for which the benefits have a significant local and spatial component; i.e., goods for which placement in the ward is desirable. The goods are: drinking water, sanitation, improved roads, electricity, street lights, and public telephones.¹⁰ As reported in Appendix Table A1, these six goods account for 67.5 percent of the *panchayat's* discretionary spending, which is over four times the amount it spends on schools and health facilities.¹¹ We assume that while the supply of these goods at the ward level is endogenously determined, they cannot be targeted more finely within the ward; i.e. at the street level, so as to favor their own group. Empirical support for this assumption will be provided below.

How are *panchayat* expenditures allocated? The council makes decisions collectively (the president does not have veto power) and so the level of public goods that a ward receives will depend on its representative's influence within the *panchayat* as well as his ability to ensure that the earmarked resources reach their destination.¹² In general, the supply of public goods to a ward will depend on the representative's ability and the effort he puts into the job. The central premise of our theory is that the caste can avoid the collective action problem, which arises because the representative only values his own benefit from the public good, by ensuring that its members can credibly commit to

Madhya Pradesh, Maharashtra, Orissa, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal. Punjab and Jharkhand did not have any ward-based elections and the election data are not available for Gujarat and Kerala.

⁹A caste group is any set of households within a village reporting the same (sub) caste name. Christian households provided their original caste names and Muslim households provided their equivalent *biradari* affiliation. Most Christians continue to marry within their original caste group. We counted Muslim households within a village that were without a formal *biradari* name as a unique caste. On average, there are seven wards per village, 67 households per ward, and six castes per ward.

¹⁰Public irrigation investments or school buildings, for example, are valued local public goods whose placement within the ward (defined by place of residence) may not be desirable.

¹¹Key informants in the village were asked to rank 12 issues, by importance, that came under the purview of the elected *panchayat*. Inadequate roads and drinking water were ranked 1 and 2, followed by health, schooling, sanitation, street lights and electrification. Note that the low spending on health and education and the relatively low level of importance assigned to these goods by the key informants reflects the fact that they are largely allocated at the state level and so fall outside the purview of the village *panchayat*.

¹²Key informants in the 2006 REDS were asked who in the *panchayat* decided the allocation of local revenue. Although 81% of informants reported that the president had a say, 93% said that it was, nevertheless, a joint decision of all *panchayat* members. In contrast, just 5% of respondents said that allocation decisions were determined by an influential caste group in the village.

compensating him *ex post* for his efforts. These *ex post* transfers can result in politicians appearing to be enriched by their position. However, the important distinction between caste-based transfers and personal corruption is that the entire constituency benefits from the increased level of public goods generated by these transfers and public resources are not diverted for personal use.

Ex post transfers within the caste are one (second-best) solution to the collective action problem. A second solution is to compensate the representative monetarily for his effort and opportunity cost. Ferraz and Finan (2011) provide evidence from Brazilian municipalities that higher wages attract wealthier and more educated candidates and increase legislative productivity. However, local politicians are rarely (adequately) compensated. Ferraz and Finan note that 98% of municipal legislators held a second job, and in our Indian local governments, the president is paid 50-60 dollars per month (less than the minimum wage) and the ward representatives earn even less. One reason why wage compensation, especially for local politicians, is not an effective solution to the collective action problem may be because high-powered incentive schemes, rather than fixed wages, are needed to elicit effort. It is difficult, however, to create clear and credible performance measures for many important political tasks (Besley 2004).

A third solution to the collective action problem is long-term incumbency. The members of the representative's caste can credibly commit to compensate him *ex post* for his efforts even when he is elected for a single term because they interact with him outside the political system on many other dimensions. If the representative could stay in office for a sufficiently long period of time, then the standard repeated game argument tells us that cooperation could even be sustained with socially unconnected outsiders. This solution is infeasible in Indian local governments because reservation for women and disadvantaged castes (and tribes) generates a high level of exogenous turnover in council seats. The rule followed by almost all Indian states is that seats are reserved in each election for three historically disadvantaged groups – Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Castes (OBC) – in proportion to their share of the population in each district. Within each of these categories, and in constituencies open to all castes in a given election, one-third of the seats are reserved for women. Seats are reserved randomly across wards and, for the position of the president, randomly across *panchayats* from one election to the next in each district. The only restriction is that no seat can be reserved for the same group across consecutive elections within a constituency (Besley, Pandey, and Rao 2007).

Given the negative priors that the electorate will have about female politicians and politicians drawn from historically disadvantaged groups, council representatives chosen in reserved elections have little chance of being subsequently reelected.¹³ The representatives with the greatest chance for re-election are men elected in unreserved seats. However, the probability that an unreserved election will be followed by another unreserved election within a ward is just 0.4.¹⁴ Assuming that leaders

¹³Chattopadhyay and Duflo (2004) note that not a single woman in their sample of reserved constituencies in the state of Rajasthan was elected in the subsequent term (without female reservation). Exposure can change these priors, but Beaman et al. (2009) find that it takes two reserved election terms before an increase in women elected in unreserved seats can be detected.

¹⁴In our sample of ward-terms, 60 percent were open to all castes (see Table 7 below). With one-third of the seats in all categories reserved for women, this implies that unreserved elections would occur 40 percent of the time.

in reserved seats are never reelected in the subsequent election, the maximum fraction of incumbent representatives that will be elected for an additional term is 0.16. Consistent with these low rates of re-election, only 14.8 percent of the ward representatives in our sample had held a *panchayat* position before. Without the possibility of reelection, and with limited financial compensation for ward representatives, the only available solution to the collective action problem is cooperation within the caste.

2.3 Selection of Ward Representatives

Based on the discussion above, a major task of the ward representative is to channel resources to his constituency and to subsequently ensure that the planned construction and maintenance of public goods actually takes place. We measure the ability of the elected ward representative by his years of schooling. Education provides many of the skills that are associated with political competence in Indian local governments. Table 4 shows that among household heads, there is a positive correlation between schooling, being a manager, and the size of landholdings, within the caste. Thus, the more educated are more likely to have managerial experience. Individuals who manage large enterprises, such as businessmen and farmers, will be particularly well-suited to manage public goods delivery, and we would expect larger landowners to be more influential in the *panchayat* council. The challenge is to get such individuals to agree to stand for election.

The 2006 REDS census indicates that more educated household heads were significantly less likely to have voted in the last *panchayat* election. This same result (reported in Appendix Table A2) is obtained for rural households in the last national election from the 2005 India Human Development Survey (IHDS), which covers a nationally representative sample of households, and from the 1985 Indian National Election Study. These results would seem to imply that these competent individuals would also be more reluctant to serve as ward representatives, especially if education is positively correlated with the opportunity cost of their time.

Table 5 provides descriptive evidence indicating that ward representatives are, to the contrary, positively selected on education. The village inventory in the 2006 REDS includes a special module that collected information on the education of the elected representative from each ward in each of the last three election terms. Educational attainment of the elected ward representative is measured in four categories – illiterate, primary graduate, secondary graduate, and post-secondary graduate – which we convert into years of schooling.¹⁵ The village census provides the years of schooling of each household head. This allows us to compute the educational attainment of the median individual in the ward, separately by reservation category. Since 95% of household heads are male, we report average educational attainment separately for male and female representatives. Male representatives have substantially higher schooling than the median household head in all reservation categories. Highlighting the positive selection that we document, even female representatives have higher schooling than the median household head in SC and OBC elections, despite the fact that female education is

¹⁵Years of schooling are imputed by assigning 4 years of schooling to primary graduates, 10 years of schooling to secondary graduates, and 14 years of schooling to post-secondary graduates.

much lower than male education in rural India.¹⁶ The positive selection on education documented in Table 5 indicates that some mechanism is in place to solve the collective action problem and incentivize more educated individuals to stand for election. The theory of ethnic politics that follows generates additional predictions linking this selection to the size of the representative's group in the ward.

3 The Theory

3.1 Ethnic Groups and Public Resources

N individuals, belonging to K ethnic groups, reside in a political constituency. Each ethnic group k consists of N_k individuals, such that $\sum_k N_k = N$ and the k subscript sorts groups by size; $N_{k-1} < N_k, \forall k$. An ethnic group is defined as a set of individuals who interact frequently with each other but very little with outsiders. Exclusion from these interactions is an effective sanctioning device, which allows groups to support high levels of internal cooperation. In the context of local politics, each group can, therefore, credibly commit to compensate its representative *ex post* for his effort, conditional on being elected, even if he holds office for a single term. The representative, in turn, will internalize the benefit that all his co-ethnics derive from the public resources he provides. In contrast, he places no weight on the benefit derived by outsiders because they cannot credibly commit to compensating him in the same way.

The residents of the constituency receive two types of public resources: non-excludable public goods and a targetable welfare transfer. All individuals derive the same utility from these public resources. The literature on ethnic politics has focused on the targeting of public goods to co-ethnics (the insiders) at the cost of outsiders. At the local level, which is the focus of our analysis, a large component of public goods are non-excludable. The analysis that follows will derive the relationship between the ethnic composition of the constituency and the ability (and effort) of the selected representative, which will, in turn, determine the supply of public goods that benefit both insiders and outsiders. It will, in addition, examine how these relationships change when targetable welfare transfers are included in the representative's list of responsibilities.

3.2 Representative's Effort and Candidate Ability

The constituency must select an unpaid representative from among its residents for a single term.¹⁷ We ignore the welfare transfers to begin with, so the representative's sole task is to supply public goods to his constituents. The level of public goods will depend on the amount of effort that he exerts. When choosing his level of effort, the representative will take account of the cost of effort, which is decreasing in his ability, and the benefit derived from public goods by his co-ethnics residing in the constituency. A representative with ability ω belonging to group k of size N_k , will thus choose

¹⁶We do not distinguish between male and female representatives when testing the theory below. Female reservation is independent of the reservation by caste that generates variation in the size of the largest eligible caste over time, and thus can be ignored in the analysis.

¹⁷The representative could receive a fixed compensation or a fixed rent from being in office without changing any of the results that follow. It is possible that the rent he extracts is increasing in his ability. This will not affect the results as long as the opportunity cost of time is increasing more steeply in ability than the rents, in which case the challenge of getting more able individuals to become representatives still remains.

effort, e , to maximize

$$N_k e^\beta - \frac{e}{\omega},$$

where e^β is the level of public goods received in the constituency. While more effort will obviously increase the supply of the goods, $\beta > 0$, we assume that the return to effort is decreasing at the margin, $\beta < 1$. In fact, we will need to place the stronger restriction that $\beta < 1/4$. We assume, in addition, that the level of public goods maps linearly into the utility derived from their consumption by each resident. We normalize so that this mapping is one-for-one. The representative's optimal level of effort can then be expressed as an increasing function of his ability and the size of his group,

$$e(\omega, N_k) = (\beta \omega N_k)^{\frac{1}{1-\beta}}. \quad (1)$$

The level of effort when ethnicity is salient is greater than the benchmark where the representative only cares about himself (N_k would be replaced by one in the preceding equation) but less than first-best (in which case N_k would be replaced by N).

As in the risk-sharing literature, social connectedness supports *ex post* transfers within the ethnic group, which thus operates as a unitary decision-making unit. It will put forward as its candidate the individual who maximizes its overall benefit from public goods net of his effort cost and the cost of the outside opportunities he foregoes, which we assume to be increasing linearly in his ability to be consistent with the specification of the effort cost. A group of size N_k will thus put forward the individual with ability ω that maximizes

$$N_k [e(\omega, N_k)]^\beta - \frac{e(\omega, N_k)}{\omega} - \alpha \omega. \quad (2)$$

Substituting the expression for $e(\omega, N_k)$ from equation (1) and then maximizing with respect to ω , the candidate's ability can be expressed as an increasing function of the size of his group,¹⁸

$$\omega(N_k) = \left[\frac{\beta N_k}{\alpha^{1-\beta}} \right]^{\frac{1}{1-2\beta}}. \quad (3)$$

Substituting this expression back in equation (1), his effort (conditional on being elected) is also increasing in the size of his group

$$e(N_k) = \left[\frac{(\beta N_k)^2}{\alpha} \right]^{\frac{1}{1-2\beta}}. \quad (4)$$

It follows that representatives of larger groups will have higher ability and supply a higher level of public goods. Substituting the expressions for ability and effort from equations (3) and (4) in (2), the net benefit from public good provision is

$$\left(\frac{\beta^2}{\alpha} \right)^{\frac{\beta}{1-2\beta}} N_k^{\frac{1}{1-2\beta}} (1 - 2\beta), \quad (5)$$

which is strictly positive. Each group would prefer to put forward its own representative to the default with no public good provision. The analysis that follows examines which group's representative will be selected in equilibrium.

¹⁸This result is independent of the ability distribution, which could vary across groups. All that we need for an interior solution is that the optimal ability level should lie within the support of the ability distribution in each group and that an individual with that level of ability is present in the group. This will be true if group sizes are large and there is sufficient heterogeneity in ability within groups.

3.3 Representative Selection and Public Good Provision

Elections are contestable. Each ethnic group in the constituency chooses whether or not to put its preferred candidate up for election.¹⁹ The decision to stand is accompanied by an entry cost, which is close to zero. The only role for this entry cost is to rule out equilibria in which candidates with no chance of winning stand for election. After all groups have simultaneously made their entry decision, the election takes place and the candidate with the most votes is selected to represent the constituency for a single term.

If all ethnic groups fielded their preferred representatives as candidates, then the representative of the largest group would always be elected. Once we allow groups to decide whether or not to field a candidate, however, this outcome will not necessarily be obtained. In particular, the largest group could free-ride on a smaller group (and avoid bearing the effort and opportunity cost of its own representative) if the level of public goods supplied by the other group's representative is sufficiently large.

Our first theoretical result describes the relationship between the size of the largest group in the constituency and the size of the elected representative's group, which, in turn, maps into the representative's ability and the supply of public goods. When deriving this result, we exogenously vary the size of the largest group, N_K , holding constant the population in the constituency, N . As N_K increases, the average size of all other groups must decrease to maintain the total population at a constant level. We make the slightly stronger assumption that the size of all other groups, N_j , $j \neq K$, is weakly declining as N_K increases. The analysis considers the entire range of N_K , or its population share equivalent, $S_K \equiv \frac{N_K}{N}$, starting from the case where all groups in the constituency are of equal size, $\frac{N}{K}$, and then increasing N_K up to the point where $N_K \rightarrow N$.

Although there is a continuous and positive relationship between the size of the elected representative's group and both the representative's ability and the supply of public goods, the relationship between the size of the largest group and these outcomes is characterized by a discontinuity:

Proposition 1. *There is a discrete increase in the ability of the selected representative and the supply of public goods when the population share of the largest group, S_K , reaches a threshold, S^* .*

To establish this result, we first characterize the relationship between the size of the largest group and the size of the elected representative's group, which is an endogenous outcome of the electoral process. Because he supplies a higher level of the public good than any other group's representative, the representative of the largest group will certainly be elected if his group puts him up as a candidate. His group will choose to do so if the *Incentive Condition* is satisfied; i.e. if it prefers him to the representative of any other group, despite having to bear his effort and opportunity cost.

For the incentive condition to be satisfied, the largest group must prefer its own representative to the next largest group's representative. If this is the case, then it will certainly prefer its own

¹⁹The electoral process is the same as the citizen-candidate models of Osborne and Slivinsky (1996) and Besley and Coate (1997), except that groups put up their preferred candidates.

representative to any other (smaller) group's representative. The required condition is

$$N_K e(N_K)^\beta - \frac{e(N_K)}{\omega(N_K)} - \alpha\omega(N_K) \geq N_{K-1} e(N_{K-1})^\beta. \quad (6)$$

Substituting from equations (3) and (4), the preceding inequality can be rewritten as,

$$N_K \left(\frac{\beta^2}{\alpha} \right)^{\frac{\beta}{1-2\beta}} \left[N_K^{\frac{2\beta}{1-2\beta}} (1 - 2\beta) - N_{K-1}^{\frac{2\beta}{1-2\beta}} \right] \geq 0. \quad (7)$$

If this condition is satisfied, there is a unique equilibrium in which the largest group puts forward its representative for election and no other group fields a candidate.²⁰

When will the incentive condition (IC) be satisfied? This will depend on the ethnic composition of the constituency; in particular, the population share of the largest group. If all groups in the constituency are of equal size, N/K , then the term in square brackets in inequality (7) will be negative (recall that $\beta \in (0, 1/4)$ by assumption). If the largest group accounts for almost the entire population, $N_K \rightarrow N$ and $N_{K-1} \rightarrow 0$, the term in square brackets will reverse sign. Holding constant the population of the constituency, the assumption is that the size of all other groups, including the next largest group, will (weakly) decline as N_K increases. By a continuity argument, there is thus a threshold N_K^* or, equivalently, a threshold population share, $S^* \equiv N_K^*/N$, at which the inequality is just satisfied. Above the threshold, the largest group will always put its representative up for election.

Below that threshold, inequality (7) is not satisfied and there will be multiple equilibria. Replacing N_{K-1} with a smaller sized group, there will be a group k for which inequality (7) is just satisfied. Any strategy profile in which a group of size $N_k \in (N_k, N_K]$ fields its representative, while all other groups stay out, will be an equilibrium.²¹ Because any of these groups could be selected in a given election, there will be a discontinuous increase in the size of the elected representative's group when the population share of the largest group just reaches the threshold, S^* .

Intra-group cooperation results in a positive and continuous relationship between a group's size and the ability of its preferred representative, from equation (3), as well as the amount of the public good he supplies, from equation (4). Because the size of the elected representative's group increases discontinuously at a population share threshold, S^* , this internal cooperation results in a discontinuous increase in both the elected representative's ability and the level of public goods at that threshold.

²⁰(i) The strategy profile in which no one contests, and public goods are not provided, is not an equilibrium. Any group would be better off by deviating and fielding its representative, who would generate a positive net benefit for the group – from expression (5) – once selected. (ii) Any strategy profile with multiple candidates is not an equilibrium. Given the cost of entry, smaller groups (who are sure to lose) would be better off not contesting. (iii) Any single-candidate strategy profile in which a group other than the largest group fields its representative is not an equilibrium. The largest group will always deviate and field its candidate if inequality (7) is satisfied. (iv) The proposed strategy profile is an equilibrium. The largest group will not deviate because it receives a positive net benefit from having its representative selected, which exceeds the default (with no public good provision) when no group fields a candidate. No other (smaller) group wants to deviate and put forward a candidate because it would certainly lose the contest, while having to bear the cost of entry.

²¹(i) Because inequality (7) is *not* satisfied for group k , the largest group, K , will not deviate from this equilibrium and put its candidate up for election. It follows that no group that is larger than k but smaller than K will want to deviate. (ii) No group smaller than k will deviate because it will lose the election. (iii) Group k will not deviate because it receives a positive net benefit from the public good.

3.4 Adding Welfare Transfers

Local government representatives are often entrusted with the task of distributing welfare transfers in addition to their primary function of supplying public goods. Let each beneficiary receive one unit of the transfer, which maps into its utility equivalent θ , and let the constituency receive a fixed T units of the transfer. Unlike public goods, which are non-excludable at the local level, these transfers can be targeted to the elected representative's ethnic group.²² The representative will first ensure that each member of his group receives the welfare transfer, allocating the remaining units (if any) to the outsiders in his constituency. We assume that $T \in (\frac{N}{K}, N)$, which implies that there is always some rationing of the welfare transfers ($T < N$) but that outsiders are not crowded out when all groups are of equal size ($T > \frac{N}{K}$).

It is straightforward to verify that the probability that an outsider will receive the welfare transfer, $\max\left(\frac{T-N_k}{N-N_k}, 0\right)$, is (weakly) decreasing in N_k . A given ethnic group will continue to put forward the same (most preferred) representative and that representative will continue to exert the same level of effort and supply the same level of public goods if elected. While a larger group's representative thus continues to supply a higher level of public goods, outsiders are worse off with respect to the welfare transfer when a larger group is in power. This changes which group's representative gets elected, and, by extension, the representative's ability and the supply of public goods in equilibrium.²³

When deriving Proposition 1, we noted that the largest group would want its representative to be elected if the incentive condition (IC) was satisfied. Once welfare transfers are introduced, the incentive condition will be easier to satisfy because free-riding on a smaller group is less attractive. As derived formally below, the population share at which the incentive condition just binds for the largest group will decline from S^* to a lower threshold, S^{**} . While the largest group now has a greater incentive to have its representative elected, an additional *Feasibility Condition* (FC) must also be satisfied to ensure that its representative is elected when his group does not have an absolute majority. For this condition to be satisfied, the largest group's candidate must be preferred to any other group's candidate by voters belonging to neither of those groups.

Without welfare transfers, the feasibility condition is always satisfied because the largest group's representative supplies a higher level of non-excludable public goods than any other group's representative. With welfare transfers, this need not be the case because the largest group's representative is the least preferred representative with respect to the delivery of welfare transfers to outsiders. We show that if the welfare transfers are relevant but do not completely dominate the public goods effect, that there is a unique S^F , such that the feasibility condition is satisfied for $S_K \equiv \frac{N_k}{N} \leq S^F$, but not for $S_K > S^F$.

In light of the preceding discussion, adding targetable welfare transfers to the representative's list

²²In practice, the transfers are intended for economically or socially disadvantaged individuals. This feature of the transfer mechanism could be easily incorporated in the model, without changing the results that follow, by assuming that a fraction of non-eligible insiders receive the transfers and that the transfers that remain are randomly allocated to eligible outsiders.

²³An ethnic group could put forward a candidate whose ability is higher than its most preferred representative as a way of getting elected and subsequently capturing the welfare transfers. This strategy is not credible if other members of the group, in particular the preferred representative, can function as proxies for the candidate once he is elected.

of responsibilities will change which group gets elected, and the accompanying supply of public goods in the following way: If the feasibility condition is satisfied for the largest group, $S^{**} < S^F$, then the threshold at which its representative gets selected will shift down from S^* to S^{**} . The reduced incentive to free-ride on other groups thus leads to an increase in the overall supply of the public good. In contrast, if the feasibility condition is not satisfied due to the crowding-out effect of the welfare transfers, $S^{**} > S^F$, then the representative of the largest group will only be elected when it has an absolute majority. The threshold will now shift up, resulting in an overall decline in the supply of the public good. We prove formally in the Appendix that

Proposition 2. *(a) If $S^{**} > 0.5$, or if $S^{**} < 0.5$ but the feasibility condition continues to be satisfied for the largest group; i.e. $S^{**} < S^F$, then there will be a discontinuous increase in the ability of the selected representative and the supply of the public good at a threshold population share, S^{**} , which is lower than the threshold without welfare transfers, S^* . (b) If $S^{**} < 0.5$ and the feasibility condition is not satisfied for the largest group; i.e. $S^{**} > S^F$, then the discontinuous increase in the representative's ability and the supply of the public good will only occur when the largest group has an absolute majority; i.e. at a population share of 0.5.*

The theory incorporates the positive and the negative aspects of ethnic politics. If the threshold at which there is a discontinuous increase in the elected representative's ability and the supply of public goods is located at 0.5, this will allow us to empirically identify both these aspects. As noted, the discontinuous gain when the population share of the largest group crosses a threshold is indicative of underlying cooperation within groups. The fact that the largest group's representative is only elected when it has an absolute majority, despite the fact that he will provide a higher level of non-excludable public goods, is indicative of targeting (of welfare transfers) to insiders, at the expense of outsiders.

4 Verifying Key Assumptions

The prediction that there will be a discontinuous increase in the ability of elected representatives, with an accompanying increase in public good provision, when the population share of the largest ethnic (caste) group crosses a threshold is based on the key assumption that the opportunity cost of the representative's time is increasing in his ability, which we measure by education, and that some public goods are non-excludable at the local level. The additional prediction that this threshold could be precisely at 0.5 if the negative targeting effect dominates the positive public goods effect relies crucially on the assumption that welfare transfers are targetable. We verify each of these assumptions below.

A1. Welfare transfers are targetable, while public goods are locally non-excludable:

Our theory highlights how the ability to target welfare transfers reduces the attractiveness of candidates from larger groups. To test directly for the existence of targeting, we examine the receipt of Below the Poverty Line (BPL) transfers by households in our sample villages. Information on whether a household received a BPL transfer, in each of the last three election terms, is obtained from the

detailed survey that is administered to a sample of households in each REDS village. These transfers are meant to be received by economically disadvantaged households, but it is well known and well documented (for example, by Besley, Pande, and Rao 2007, and Bardhan and Mookherjee 2010) that ineligible households who are politically connected can also benefit from them.

To identify BPL targeting in the ward along caste lines, we estimate the following equation using the survey data:

$$BPL_{ijt} = \eta_1 RC_{ijt} + \eta_2 RN_{jt} + Z_{ijt}\delta + \xi_{ijt}, \quad (8)$$

where the dependent variable is an indicator that takes the value one if household i receives a BPL transfer in ward j in term t ; RC_{ijt} is a binary variable indicating whether or not the ward representative in that term belongs to the household's caste; RN_{jt} measures the number of households belonging to the representative's caste in the ward. We also include a vector of additional regressors, Z_{ijt} – household fixed effects, reservation dummies, election-term dummies, and the election year. ξ_{ijt} is a mean-zero disturbance term. The sample for the current analysis is restricted to the 35% of ward-terms for which information on the ward representative's caste is available from the village inventory and can be matched to castes in the ward (based on the village census).

The conditional (fixed effects) logit model is used to estimate equation (8) because the mean of the dependent variable is far from 0.5.²⁴ The coefficient on the representative's caste in Table 6, Column 1 is positive and significant. Because household fixed effects are included in the regression, the interpretation of this result is that a household is more likely to receive a BPL transfer when it shifts from being an outsider to an insider; i.e. when the ward representative belongs to its own caste. This is directly indicative of targeting to insiders, at the expense of outsiders. However, the probability of receiving a BPL transfer does not appear to depend on the size of the representative's caste.²⁵

A major virtue of the test of our theory is that it is based on the population share of the largest eligible caste. This share changes exogenously within the ward across election terms due to random caste reservation. In contrast, the caste identity of the elected representative and the size of his group, which we include as regressors in Table 6 are outcomes of the electoral process. Thus, the estimates of η_1 and η_2 are not causal. Once household fixed effects are included, the threat to identification is that RC_{ijt} and RN_{jt} proxy for unobserved changes in the household's BPL status. The issue is whether the estimates constitute a valid test of targeting.

Suppose, for example, that castes with a greater fraction of households below the poverty line put more effort into getting their representative elected. Then belonging to the representative's caste is an indicator of adverse economic circumstances, and the η_1 coefficient will be biased upward if current economic conditions are omitted from the regression. Note that the over-representation of poorer

²⁴As with the ward-level regressions that we use to test the theory, the sample is restricted to wards with more than one caste and more than one street. Outlying ward-terms in the top 1 % of the *panchayat*-level public goods expenditure distribution are also excluded. Standard errors are clustered at the ward level.

²⁵This result is consistent with our theory. Let the number of insiders be N_I and the number of outsiders be N_O , where the total population of the ward, $N = N_I + N_O$. Assume that $N_I < T$, the total transfers, so the outsiders are not crowded out completely. Then the probability of receiving a transfer, $\frac{N_I}{N} \cdot 1 + \frac{N_O}{N} \frac{T-N_I}{N_O} = \frac{T}{N}$, which is independent of N_I .

castes will not be observed if there is no targeting to begin with, since the caste would not benefit from having its own representative elected. Thus, a spurious correlation will not be generated if there is no targeting. We are primarily interested in identifying the presence of caste-based targeting, and this is still accomplished by testing (and rejecting) the hypothesis that $\eta_1 = 0$ in equation (8).

If the entire ward faces a negative economic shock, then the demand for BPL transfers will increase in the electorate. The representative of a large group will be relatively unpopular in this environment because his group crowds out other castes. Representatives of large castes are thus less likely to be elected when (unobserved) economic conditions are unfavorable, biasing the η_2 coefficient downward. This explains the negative coefficient on the RN_{jt} variable in equation (8), although it is not statistically significant at conventional levels.

Our theory further indicates that for outsiders, the larger the size of the representative's caste, the lower should be the probability of receiving BPL transfers. To test this, we allow the size of the representative's caste to affect insiders and outsiders separately by adding an interaction term to equation (8):

$$BPL_{ijt} = \eta_1 RC_{ijt} + \eta_2 RN_{jt} + \eta_3 RC_{ijt} * RN_{jt} + Z_{ijt}\delta + \xi_{ijt}. \quad (9)$$

The η_2 coefficient now provides an estimate of how the size of the representative's caste affects outsiders and is expected to be negative. Consistent with the theory, and directly indicative of targeting to insiders at the expense of outsiders, this coefficient in Table 6, Column 2 is negative and significant (at the 10 percent level). The point estimates indicate that for every increase in the size of the representative's caste by 10 households, there is a decrease in the probability that an outsider receives the BPL transfer by 6 percentage points (25 percent). At the mean size of the representative's caste (37 households), being an insider increases the probability of receiving the transfer by 4 percentage points.

Unlike welfare transfers, the theory assumes that public goods are not targetable at the local level (within the ward). To test this assumption, we replace access to BPL transfers with public good provision as the dependent variable in equations (8) and (9). As noted, public goods are measured at the street level. We can thus see whether public goods are targeted to streets on which the representative's caste members are concentrated. Our measure of overall public good provision at the street level is the fraction of the six major public goods that are received in a given election term. The implicit assumption when constructing this aggregate statistic is that caste size, the source of forcing variation in our theory, affects the supply of each public good equally. Empirical support for this assumption will be provided below. Note that this does not imply that the average level of each public good is the same. Appendix Table A3 reports the fraction of households in the ward that received each public good, averaged across wards and election terms, by type of reservation. It is apparent that a large fraction of households benefited from expenditures on water, roads, and sanitation, while a much smaller fraction benefited, in any term, from expenditures on electricity, street lights, and public telephones.

Because the value of the aggregate public goods variable is the same for all households on a street, we test for targeting across streets rather than across households. The variable indicating whether the

representative belongs to the household's caste is thus replaced by the fraction of households on the street that belong to the representative's caste. The specification of the estimating equation and the sample that we use for estimation are otherwise unchanged.²⁶ Columns 3-4 report the estimates of the public goods equation. In contrast with what we observe for BPL transfers in Columns 1-2, there is no evidence that public goods are being targeted.

A2. Opportunity cost is increasing in the representative's ability: The key implication of the theory is that the representative's competence, measured by his education, and the accompanying level of public goods, are increasing in the size of his caste in the ward. This result is driven by two assumptions: (i) that the representative internalizes the utility derived from the public goods by all N_k members of his caste group, and (ii) that the opportunity cost of being the representative is greater for more able individuals ($\alpha > 0$).

If the first assumption is not satisfied, then it follows from equations (3) and (4) that the representative's ability and public good provision will be independent of group size. The earlier discussion on intra-caste cooperation provides support for this assumption. If the second assumption is not satisfied, then it follows from the same equations that the representative will always be the most competent member of his caste, regardless of whether or not there is intra-caste cooperation. Representatives of larger castes could be more competent, even if the ability distribution is independent of caste size, because there are more individuals to choose from at the very top of the distribution (where numbers are necessarily small).

To provide support for the assumption that $\alpha > 0$, we test the additional implication of our theory, which is that representatives of larger castes should be drawn from higher up in the education distribution of their caste in the ward. The alternative theory, which assumes $\alpha = 0$, would predict instead that the representative is always drawn from the top of the education distribution. As predicted by our theory, we see in Figure 2 that there is a positive and continuous relationship between the size of the representative's caste in the ward and his position in his caste's education distribution; starting from just above the median level of education and reaching as high as the 80th percentile (which is, nevertheless, quite far from the very top of the distribution).

5 Testing the Theory

5.1 Identification Strategy

The theory relates the ethnic composition of the constituency, which we measure by the caste composition of the ward, to the supply of public goods. However, ethnic (caste) composition could also determine the demand for public goods. To disentangle the supply effect from the demand effect, we

²⁶If we continued to use the binary variable indicating whether the representative belonged to the household's caste to measure targeting, then this variable would vary across households on the street, whereas the dependent variable measuring public good provision would remain the same. This would bias the coefficient on the indicator variable towards zero. Once the appropriate measure of targeting – the fraction of households on the street belonging to the representative's caste – is included as a regressor, there is no household-level variation within a street in the dependent variable or any of the regressors. We continue to estimate the public goods regressions in Columns 3-4 at the household level, with household fixed effects, to be consistent with the BPL regressions in Columns 1-2. Street-level regression results, with street fixed effects, provide qualitatively similar estimates (available from the authors).

take advantage of the panel data on elections and randomized caste reservation in Indian local governments, which exogenously changes the composition of castes that are eligible to stand for election in the ward from one election to the next. This variation in the (eligible) caste composition over time allows us to subsume all permanent characteristics of the ward, including the demand for public goods by the electorate, in a fixed effect.²⁷

The ward fixed effects will also capture the role played by any dominant caste in the ward. Anderson et al.'s (2015) theory of clientelist politics describes an environment in which a particular caste captures power, providing side-transfers to the rest of the electorate in return. Although the provision of public resources could vary across wards in this theory, depending on whether or not that caste accounts for a majority of the electorate, public resources would not vary over time within a ward, depending on the set of eligible castes, as implied by our theory. Anderson et al. note that clientelist politics will not be universally salient. Where it is salient in our data, the effect will be subsumed in the ward fixed effect.

The theory predicts that there will be a discontinuous increase in the ability of elected representatives and the supply of the public good when the population share of the largest ethnic group (caste) crosses a threshold level. Once we include ward fixed effects, the equivalent prediction is that there will be a discontinuous change when the population share of the most numerous *eligible* caste crossed a threshold level. To implement this test, there needs to be sufficient variation in the population share statistic within wards over election terms. As noted, ward elections are reserved for Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Castes (OBC) in proportion to the share of these groups in the population at the district level. Among the ward-terms in our sample, 11 percent were reserved for Scheduled Castes, 6 percent were reserved for Scheduled Tribes, and 23 percent were reserved for Other Backward Castes. Panel A of Table 7 describes the share of the most numerous eligible caste in the ward by type of election. These shares are generally quite large, even in reserved elections, reflecting the fact that neighborhoods in rural India are often dominated by a single caste. Nevertheless, there is substantial variation in population shares within and across reservation categories.

Panel B of Table 7 displays the fraction of ward-terms in which the share of the most numerous eligible caste exceeds alternative pre-specified cutoffs, by the type of reservation. Matching the descriptive statistics in Panel A, the proportion of elections in which the cutoff is exceeded is largest for unreserved elections, followed by elections in which the ward candidates are restricted to ST, OBC, and SC in that order, regardless of the cutoff that is specified. Just as the likelihood that any pre-specified cutoff is crossed varies across different reservation schemes in Table 7, there will be variation in the likelihood that the threshold will be crossed from one term to the next *within* a ward as the type of reservation changes. It is possible that the pool of potential candidates will be weaker in reserved

²⁷The implicit assumption when we include the ward fixed effect is that the population of the ward, or the electorate, remains essentially unchanged over time. This is a reasonable assumption given the unusually low spatial mobility that is characteristic of rural India. Munshi and Rosenzweig (2016), for example, report that permanent migration rates of men aged 20-30 out of their origin villages were as low as 8.7 percent in 1999. The corresponding rates for entire households would be much lower. Indeed, the census listing data indicate that since 1991 when the local electoral system was put in place, less than 3 percent of new households had migrated into the sample villages.

(lower caste) elections. All the specifications that follow consequently include a full set of reservation dummies in addition to ward fixed effects.

Ward representatives operate independently in our theory. In practice, the level of public goods received in a ward is determined by a collective decision-making process that involves all the ward representatives and the *panchayat* president. Although the political equilibrium in the ward continues to matter, the equilibrium in other wards will also be relevant to the extent that there is a fixed component to the total resources available in the *panchayat*. However, random and independent reservation in elections across wards and for the president's position implies that the political equilibrium in a ward, measured by the population share of the largest eligible caste, will be orthogonal to the equilibrium in other wards, which is subsumed in the error term.

5.2 Descriptive Evidence

The theory indicates that the elected representative's ability and the supply of the public good will increase discontinuously when the population share of the largest eligible caste in the ward crosses a threshold level. We semi-parametrically examine the relationship between the population share and each of the outcomes by estimating the following equation:

$$y_{jt} = \phi(S_{jt}) + X_{jt}\gamma + \zeta_{jt}, \quad (10)$$

where S_{jt} is the population share of the largest eligible caste in ward j and election term t ; y_{jt} is either the elected representative's ability, measured by educational attainment, or the supply of public goods, measured by the population weighted average of the fraction of the six major public goods received by each street in the ward in that term; X_{jt} is a vector of additional regressors including a full set of ward and reservation dummies, election-term dummies, and the election year; and ζ_{jt} is a mean-zero disturbance term.

We estimate the relationship between y_{jt} and S_{jt} in two steps. In the first step, we partial out the additional regressors, X_{jt} , by estimating an equation that replaces the $\phi(S_{jt})$ function with a vector of dummy variables. There are five dummy variables, partitioning $S_{jt} \in [0, 1]$ into six intervals, in the benchmark specification.²⁸ We discuss the robustness of the results to alternative specifications of the $\phi(S_{jt})$ function below.

The estimated γ vector from the first step can be used to construct the conditional outcome, $\tilde{y}_{jt} \equiv y_{jt} - (X_{jt} - \bar{X}_{jt})\hat{\gamma}$, where the sample average of each additional regressor, \bar{X}_{jt} , is included to preserve the mean of the outcome. In the second step, \tilde{y}_{jt} is regressed nonparametrically on S_{jt} :

$$\tilde{y}_{jt} = \phi(S_{jt}) + \zeta_{jt}. \quad (11)$$

Figure 3 reports nonparametric estimates of equation (11), using the two-step procedure, with the elected representative's education and the fraction of the six major public goods received in each

²⁸The dummy variables, $D_{1jt} - D_{5jt}$ are indicator variables that take the value one if $S_{jt} \in (0.2, 0.35], (0.35, 0.5], (0.5, 0.65], (0.8, 1.0]$, respectively. The reference category is $S_{jt} \in [0, 0.2]$.

ward-term as outcomes.²⁹ The representative's education and public good provision both increase (discontinuously) at a population share of 0.4, and continue to increase till 0.6, in Figure 3. Although the increase at the threshold is very steep, it is not vertical. We will see momentarily that this is an artifact of the estimation procedure, which uses a local polynomial smoother. Notice that the two variables track closely over the entire range of population shares, highlighting the close relationship between the representative's ability and public good provision that is an important feature of our theory.

5.3 Locating a Threshold

The descriptive evidence indicates that there is a discontinuous increase in the representative's ability and the supply of the public good when the population share of the most numerous caste reaches a threshold level. We now proceed to formally test for the presence of a threshold, place statistical bounds on its location, and estimate the gain in the representative's ability and the supply of public goods at the threshold. The gain at a threshold population share is most parsimoniously described by a threshold function. We follow Hansen (1999) to estimate the following equation at different hypothesized thresholds, S :

$$\tilde{y}_{jt} = \pi_1 + \pi_2 D_{jt} + \epsilon_{jt}, \quad (12)$$

where \tilde{y}_{jt} is the conditional outcome in ward j in term t ; D_{jt} is an indicator variable that takes the value one if the population share of the largest eligible caste exceeds S ; and ϵ_{jt} is a mean-zero disturbance term.

Suppose that the data-generating process for \tilde{y}_{jt} is characterized by a step function, as in equation (12), with the true threshold at \underline{S} . If we use Least Squares to estimate equation (12), and the range of hypothesized thresholds, $S \in [S_{min}, S_{max}]$, is wide enough to cover the true threshold, then the Residual Sum of Squares (RSS) will be minimized when the hypothesized threshold is precisely \underline{S} . However, when testing our theory, we are also interesting in determining whether or not the data-generating process is characterized by a threshold (step function). As described below, this can be done by examining the pattern of the Residual Sum of Squares across the entire range of hypothesized thresholds. When the data generating process is characterized by a step function, the RSS will increase steeply as the hypothesized threshold moves away from the value at which the RSS is minimized (this will be the true threshold). If the data generating process is characterized by a smoothly increasing function instead, there will still be a hypothesized threshold at which the RSS is minimized. However, the RSS will now change relatively little with other hypothesized thresholds.

Hansen (1999) formalizes this argument by constructing a likelihood ratio statistic,

$$LR(S) = \frac{RSS(S) - RSS(\hat{S})}{RSS(\hat{S})} \cdot N,$$

²⁹These estimates, and the tests of the theory that follow, restrict the sample of wards to those with more than one caste and more than one street (this will allow us to test, and reject, the spatial targeting of public goods below). Outlying ward-terms in the top 1% of the public goods expenditure distribution are also excluded. The theory implicitly assumes that there is a large number of individuals in each ethnic group, which allows it to select its optimal representative. Because the theory is consequently more accurately tested in large wards, we weight both the first-step and the second-step regressions by the ward population.

which is simply a normalization of the Residual Sum of Squares statistic, $RSS(S)$, based on its minimum value, $RSS(\hat{S})$, and the number of observations, N . The advantage of this statistic is that its asymptotic distribution can be derived. Thus, while the LR statistic will attain its minimum value, which is zero by construction, at the same assumed threshold, \hat{S} , where the RSS statistic is minimized, the additional advantage is that we can place bounds on the location of the true threshold with any pre-specified level of statistical confidence, say 95% or 99%. If the bounds are tight, the data generating process is likely to be characterized by a discontinuity at a threshold.

To illustrate how Hansen's LR statistic can be used to both identify the presence of a threshold and to pin down its location, we generated a data set that consists of three variables: the population share of the largest eligible caste in our ward-terms, a variable, V_S , that is constructed to be consistent with equation (12), with the threshold set at 0.5, and a variable, V_L , that is increasing linearly in the population share. The variables are constructed to have the same mean and we add a mean-zero noise term (with the same variance) to both variables. Figure 4 nonparametrically estimates the relationship between each variable and the population share. The V_S variable, which is constructed to increase discontinuously (and vertically) at a population share of 0.5, actually exhibits an increase from around 0.4 that continues until 0.6. This is precisely what we observed for the conditional outcomes in Figure 3, and is due to the local polynomial smoother that we use for the estimation and the noise we have added to the V_S variable. Nevertheless, there is clearly a discontinuous relationship between this variable and the population share that contrasts with the V_L variable, which is increasing continuously in the population share. The test that we describe below will pin down the threshold for the V_S variable at precisely 0.5.

Figure 5 plots the $LR(S)$ statistic for hypothesized thresholds, $S \in [0.25, 0.75]$ in increments of 0.001.³⁰ This statistic reaches its minimum value, which we noted is zero by construction, when the hypothesized threshold, S is equal to 0.5 with the V_S variable. It increases steeply as the hypothesized threshold moves away (on either side) from 0.5. In contrast, the $LR(S)$ statistic for the V_L variable changes very little over the entire range of hypothesized thresholds. The horizontal line in Figure 5, which is the 1% Critical Value, allows us to bound the location of the threshold for each variable with 99% confidence. Consistent with the preceding discussion, the bounds are very tight for the V_S variable, whereas the $LR(S)$ statistic always lies below the Critical Value line for the V_L variable.

We now proceed to apply the same procedure to the REDS data. We test formally for the presence of a threshold with the two outcomes – the representative's education and public good provision – after partialling out the additional regressors, as before. The benchmark specification uses equation (12) to locate a threshold, with the hypothesized thresholds ranging from 0.25 to 0.75 (in increments of 0.001). The Likelihood Ratio (LR) statistic corresponding to each hypothesized threshold is plotted in Figure 6, separately for each outcome. The LR statistic reaches its minimum value, which is zero by construction, at a hypothesized threshold of 0.5 for both outcomes. Variation in the LR statistic across the range of hypothesized thresholds in Figure 6 matches closely with the pattern for the V_S variable in Figure 5, which was constructed to increase vertically at a population share of 0.5. The

³⁰ Andrews (1993) shows that the search for a minimum value must be restricted to a subset of the unit interval for tests of a structural break to have sufficient power.

Critical Value (CV) line bounds the location of the threshold, and as in Figure 5, the true threshold is located very close to 0.5 with 99% confidence for both the representative's education and public good provision.

Recall that the $\phi(S_{jt})$ function in equation (10) is approximated by a vector of five dummy variables in step one of the two-step procedure (where we partial out the additional regressors in the equation). Figures 7 and 8 assess the robustness of the results to alternative specifications of the $\phi(S_{jt})$ function. One alternative specification approximates the $\phi(S_{jt})$ function by a vector of 13 dummy variables, partitioning $S_{jt} \in [0, 1]$ more finely into 14 intervals. A second specification approximates the $\phi(S_{jt})$ function by a quintic polynomial in S_{jt} . Partialling out the additional regressors in the first step, using each of these alternative specifications, Figure 7 reports the corresponding second-step relationships between (conditional) public good provision and the population share of the largest eligible caste. The estimated relationships are very similar to what we obtained with the benchmark specification for the $\phi(S_{jt})$ function in Figure 3. Threshold tests, reported in Figure 8, locate the threshold close to 0.5, with a high degree of precision, particularly for the specification where S_{jt} is partitioned more finely.

Although the theory indicates that there will be a discontinuous increase in the representative's competence and the supply of public goods at a threshold, it is not informative about the relationship between these outcomes and the population share of the largest eligible group below the threshold (where there are multiple equilibria). When we specify a step function in equation (12) to test for a threshold, we are implicitly assuming that there is no relationship between either outcome and the population share below the threshold. We are also ignoring the possibility that both the representative's competence and the supply of public goods could be increasing in population share above the threshold (where the largest group is always in power). Figure 9 assesses the robustness of the threshold test to a flexible specification of equation (12) that allows for a linear relationship between the outcome and the population share below the threshold, a mean-shift at the threshold, and a linear relationship (with a possibly different slope and sign) above the threshold. Including the benchmark test based on equation (12) for comparison, we see in Figure 9 that greater flexibility increases the precision of the threshold test, while maintaining the estimated location of the threshold at 0.5.

5.4 Gain at the Threshold

Based on our theory, intra-group cooperation generates a continuous and positive relationship between the size of the elected representative's caste and both the representative's competence (education) and the supply of public goods he would provide. This translates into a discontinuous increase in both outcomes at a threshold when the relationship is derived, instead, in terms of the population share of the largest (eligible) caste. Having precisely located the threshold at 0.5 for both the representative's education and the supply of public goods, we proceed to estimate the following equation for each outcome:

$$y_{jt} = \lambda_1 + \lambda_2 D_{jt} + X_{jt}\gamma + \zeta_{jt}, \quad (13)$$

where D_{jt} is an indicator variable that takes the value one if the population share of the largest eligible caste exceeds 0.5 and the remaining variables were defined earlier in equation (10).³¹

Columns 1 and 3 in Table 8 report estimation results with the supply of public goods and the elected representative's education as dependent variables. The mean-shift coefficient, λ_2 , is positive and significant for each outcome. Our estimates indicate that there is a 13 percent increase in the supply of the public good and a 54 percent increase in the educational attainment of the elected representative at the 0.5 threshold. As noted, this increase at a threshold population share is indicative of the internal cooperation within the caste that characterizes one side of our theory of local ethnic politics. In contrast, Banerjee and Pande's (2010) theory, which applies to higher levels of government where voters have limited information and vote mechanically on caste lines, would predict a *decline* in the elected leader's quality when a single caste had an absolute majority in the constituency.

The assumption in our model is that multiple castes cannot form coalitions. This is consistent with previous research documenting the inability of castes to trade with each other within the village (Anderson 2011) and the limited (often adversarial) local social interactions between castes. At higher levels of government, caste groups do sometimes form coalitions based on a common position in the caste hierarchy; e.g. SC, ST, OBC. If such coalitions could form locally, the caste-group would be the social unit within which cooperation was maintained. This would imply a discontinuous increase in the elected representative's competence and the supply of public goods when the population share of the caste-group exceeded 0.5. Columns 2 and 4 allow for this possibility by including an indicator variable that takes the value one if the representative's caste-group accounts for an absolute majority of the ward's population. In line with the assumption of our model, the coefficient on the caste-specific indicator variable, D_{jt} , is largely unchanged, while the coefficient on the caste-group variable is small and statistically insignificant.

The implicit assumption when constructing the aggregate public goods variable is that the effect of caste size on the supply of each public good is the same. Table 9 tests this assumption by separately estimating the mean-shift at the threshold for each of the six public goods. While the joint effect is highly significant, matching the result in Table 8, we cannot reject the hypothesis that the individual effects are equal.

In our theory, larger ethnic groups select more competent representatives, measured by their educational attainment, because a larger number of co-ethnics benefit from the non-excludable public good. An alternative explanation for the increase in the representative's competence when the largest group has an absolute majority is that larger groups are more educated on average. We assess the validity of this explanation by separately estimating the relationship between the 25th, 50th, and 75th percentiles of the education distribution for each caste in each ward, obtained from the household census, on the size of the caste in the ward. Nonparametric estimates of these relationships, reported in Figure 10, indicate that there is a mild decline in educational attainment with group size. Moreover, there is no evidence of a threshold in Figure 11. The sharp increase in the educational attainment of the ward representative that we observe in Figure 3, and formally identify as a structural break at a

³¹We account for the correlation in ζ_{jt} across wards in a given election term by clustering the residuals in equation (13) at the panchayat-term level.

threshold population share in Figure 6, is evidently due to the systematic selection of more competent representatives by larger groups (castes), as seen directly in Figure 2.

In our theory, public good provision increases discontinuously when the population share of the largest group crosses a threshold because there is a discrete increase in the size of the elected representative's caste, which translates into higher ability and increased effort. Given that the caste affiliation of the elected representative is only available for 35% of ward-terms, we do not have the statistical power to identify an increase in the probability that the representative of the largest caste is elected at a threshold. However, we saw in Table 2 that this representative was almost always elected (with 90% probability) above the 0.5 threshold, as predicted by the theory, and we also documented a 30 percentage point increase in the probability that this representative was elected when his caste had an absolute majority in the ward. While this change in the caste affiliation of the elected representative apparently translated into a sizeable change in public good provision, an important assumption of our theory is that the level of welfare transfers, in contrast, is constant. If the total amount of transfers is fixed, then we should observe no relationship between the fraction of the ward that receives BPL transfers and the population share of the largest eligible caste. Figure 12 uses the two-step procedure described above to show the relationship between the population share of the largest eligible caste and both BPL transfers and public good provision at the ward-term level (as a basis for comparison). We see in the figure that there is no relationship between BPL transfers and the population share in the range where the transfers could affect the supply of public goods; i.e. up to the point where the largest group has an absolute majority and comes to power with certainty. Formal tests for a threshold in Figure 13, moreover, find no evidence of a threshold for BPL transfers, in contrast to what we observe for public goods.

5.5 Structural Estimation and Counter-Factual Simulations

Having found evidence consistent with both the positive and the negative aspects of ethnic politics as incorporated in the model, we proceed to estimate its structural parameters. The structural estimates are used to (i) validate the assumption that $0 < \beta < 1/4$, (ii) quantify the increase in the supply of public goods due to the internal cooperation within castes, and (iii) assess the impact of specific policy interventions that, based on our theory, are expected to increase the supply of public goods.

From equation (4), the supply of public goods, $G \equiv e^\beta$, can be expressed as:

$$G = \left[\frac{(\beta RN)^2}{\alpha} \right]^{\frac{\beta}{1-2\beta}}, \quad (14)$$

where RN is the number of households in the representative's caste.

When estimating the parameters, α and β , we attempt to stay as close to the structure of the model as possible. However, one obvious dimension where the theory and data diverge is the assumption that all households in the ward receive the same supply of public goods. In the data, as noted, public good provision is measured at the street level and varies across streets in a ward. We account for this feature of the data in the structural estimation by assuming that the same effort by the elected representative

translates into different levels of public good provision on the ground, across streets and over time.³² Public good provision on street, s , in ward, j , and term, t , is thus denoted by $G_{s jt} = F_{s jt} \cdot G_{jt}$, where $F_{s jt}$ maps ward-level effort into street-level public goods.

Taking logs, and substituting the expression for $G_{s jt}$ in equation (14), with appropriate subscripts,

$$\log(G_{s jt}) = \frac{\beta}{1 - 2\beta} \log\left(\frac{\beta^2}{\alpha}\right) + \frac{2\beta}{1 - 2\beta} \log(RN_{jt}) + f_{sj} + \nu_{s jt}, \quad (15)$$

where the $F_{s jt}$ term is decomposed into a street-ward fixed effect, f_{sj} , and a mean-zero disturbance term, $\nu_{s jt}$.

The structural parameters, α and β , are evidently identified in the preceding equation and can be estimated using Nonlinear Least Squares.³³ The point estimates are reported in Table 10. Providing support for the parametric assumptions, we reject that $\beta < 0$ and that $\beta > 1/4$ with a high degree of statistical confidence. Consistent with the implication of our theory that larger groups supply more public goods, we can also reject that the coefficient on the size of the representative's caste, $\frac{2\beta}{1-2\beta}$, is equal to zero at conventional levels. Based on the point estimate of the β parameter, the elasticity of public goods with respect to the size of the representative's caste is 0.32. The α parameter, which measures the increase in outside options with ability, is also positive as expected, but less precisely estimated in Column 1. This parameter is jointly significantly greater than zero in Column 2, where we allow α to vary with caste status. It is substantially smaller for lower castes, presumably because the return to education and, hence, the opportunity cost of being a representative is lower for them. Notice, however, that the β parameter, which we use for the simulations below, remains very stable across the alternative specifications.

We use the structural estimates to conduct two counter-factual experiments. The first counter-factual experiment compares the benchmark supply of public goods with what would be obtained if (i) the caste was unable to ameliorate the collective action problem, and (ii) this problem was solved completely. To implement the first scenario, we replace the size of the representative's caste in the ward, RN_{jt} , by one (the representative), and then predict the supply of public goods from equation (15). In the second scenario, we replace RN_{jt} by the number of households in the ward. We plot the predicted supply for each counter-factual scenario, together with the benchmark (where the predicted supply is based on RN_{jt}) against the size of the representative's group in Figure 14. The estimates indicate that internal cooperation within the caste substantially increases the supply of public goods, and when a large caste is present in the ward, brings the supply close to first-best.

Our second counter-factual experiment quantifies the impact of specific policy interventions that we expect will increase the supply of public goods. The structural estimates and Figure 14 show that there is a sizeable gain in the supply of non-excludable public goods when the representative of a larger caste-group is elected. The negative crowding-out effect associated with the welfare transfers

³²Note that we found no evidence that street-level differences in the supply of public goods reflects targeting.

³³As with all the preceding estimates, we restrict the sample to wards with more than one caste and more than one street. Outlying ward-terms in the top 1 % of the public goods expenditure distribution (at the *panchayat* level) are also excluded. We weight by the number of households on the street to allow for heteroscedasticity in the residual term, $\nu_{s jt}$. Standard errors are clustered at the ward level.

must, therefore, be substantial to explain why the representative of the largest caste is only elected when his group has an absolute majority. The obvious policy solution to this problem would be to decouple public good provision and the distribution of welfare transfers, in which case the population share threshold at which the representative of the largest caste is surely elected would shift down.

In our data, the representative of the largest eligible caste is elected 90% of the time when it has an absolute majority and 60% of the time below the estimated population share threshold. When predicting the supply of public goods in the benchmark scenario for our policy experiment, we thus assume that the representative of the largest eligible caste is elected with certainty when his caste has an absolute majority. For ward-terms in which no caste has an absolute majority, the data indicate that the winning caste is 0.8 times the size of the largest caste on average. We use this fraction in the counter-factual simulations when the population share of the largest caste is below the relevant threshold. Given the predicted size of the elected representative's caste in each ward-term, the associated supply of public goods is obtained from equation (15).

In the decoupled scenario, the representative of the largest caste will be elected, regardless of whether or not it has an absolute majority, if the following condition is satisfied from equation (7):

$$\left(\frac{N_K}{N_{K-1}}\right) \geq \left(\frac{1}{1-2\beta}\right)^{\frac{1-2\beta}{2\beta}}.$$

Based on the point estimate of the β parameter, the term on the right hand side of the preceding inequality is 2.38. To map this value into the corresponding value for the threshold population share, we estimated the relationship between $\frac{N_K}{N}$ and $\frac{N_K}{N_{K-1}}$ across all ward-terms. Based on the estimated relationship, a value of 2.38 for $\frac{N_K}{N_{K-1}}$ corresponds to a value of 0.37 for $\frac{N_K}{N}$. This is the value that we use for the population share threshold in the counter-factual decoupled scenario. The predicted supply of public goods is the same as in the benchmark scenario for all ward-terms where the population share of the largest eligible caste is less than 0.37 or greater than 0.5. The only change is in the 15% of ward-terms where the population share is in the 0.37-0.5 range, where the largest eligible caste is now predicted to be in power.

Although the elasticity of public goods with respect to the size of the representative's caste is quite large, we note that decoupling only increases the average size of the elected representative's caste, by 20%, in 15% of the ward-terms that are affected by the decoupling. We thus do not expect the impact of this policy intervention to be dramatic. A policy that could potentially have a bigger overall impact on the supply of public goods would be to roll back the current system of caste-based reservation at the ward level (but not necessarily at other levels). By restricting the set of caste-groups that are eligible to stand for election, reservation mechanically lowers the size of the largest caste that can put up a candidate for election. In our data, as shown in Table 7, the decline in the average size of the largest eligible caste due to reservation is substantial.

In a system where ward representatives are responsible for the distribution of welfare transfers, caste-based reservation is equity enhancing because it ensures that groups who tend to be relatively small, and who would otherwise be excluded from the welfare transfers, have access to them in at least a fraction of ward-terms. Because representatives of small groups capture a smaller fraction of the

total transfers, they would also be preferred from a social welfare perspective. This advantage of the reservation system must be weighed against the reduced supply of public goods due to the reduction in the average caste-size of elected representatives. To quantify this reduction, we replace the set of eligible castes in each ward-term with the full set of castes in the ward (who would be eligible in the absence of reservation). We then repeat the simulations described above, with the benchmark scenario and the decoupled scenario.

The simulated distribution of public goods is reported in Figure 15 for the benchmark case and for three counter-factual scenarios: (i) with reservation, but decoupled; (ii) without reservation, but retaining the dual role for the representative; and (iii) without reservation and decoupled. It is evident that the biggest increase in the supply of public goods is achieved by de-reservation. For example, 39% of wards across all elections receive less than one-third of total public goods in the benchmark simulation in which caste reservation is in place and representatives serve the dual role of supplying public goods and distributing welfare transfers. This fraction of under-served wards declines to 35% when welfare transfers are removed from the representative’s list of responsibilities. The fraction of under-served wards declines further to 26% and 20%, respectively, when caste reservation is removed, with and without the coupling of public goods and welfare transfers. The advantage of the decoupling is that it is relatively simple to implement, while generating a sizeable 9 percent decrease in the fraction of under-served wards, when the reservation system is in place. The advantage of the de-reservation is that the effects are much larger. However, there are equity consequences if the decoupling policy is not applied simultaneously.

6 Conclusion

The undersupply of public goods is a hallmark of underdevelopment. A large and growing literature has taken the position that ethnic politics plays a major role in both limiting the supply of public goods and distorting their allocation in many developing countries. This literature has identified the cost of group identity in higher-level elections characterized by asymmetric information about candidate quality. In our theory of ethnic politics at the local level, where information asymmetries are less relevant, there are two aspects of ethnic groups that affect public goods provision: an inclusionary aspect associated with internal cooperation and an exclusionary aspect associated with the disregard for outsiders. The inclusionary aspect of ethnic politics results in the selection of more competent political representatives who exert more effort, resulting in an increased supply of public goods that are non-excludable at the local level. This inclusionary effect is hypothesized to be particularly strong in large groups because their representatives will internalize the benefits derived from the public goods by a larger number of individuals. The exclusionary aspect of ethnic politics, in contrast, results in the capture of targetable public resources by insiders; i.e. the representative’s own group, at the expense of outsiders. This crowding out will be especially severe when the representative belongs to a large group.

Given the two aspects of ethnic politics, our theory shows how the availability of public resources that can be targeted to individual households by politicians affects the equilibrium level of non-

excludable public goods. On the one hand, all groups, including the largest group that might not otherwise put up a candidate, have a greater incentive to have their own representatives elected. On the other hand, representatives from larger ethnic groups may, despite providing a greater supply of non-excludable goods, attract less electoral support from outsiders who are crowded out of the targetable resources. If the former effect dominates, representatives from the largest group will be more likely to be elected, resulting in an increase in the supply of public goods. If the latter effect dominates, these representatives will only be elected if they have a majority share of the population, reducing the overall supply of public goods.

Using newly available Indian data, covering all the major states over three election terms at the most local (ward) level, we provide empirical evidence that is consistent with the existence of both sides of ethnic politics. Ward representatives are responsible for the supply of non-excludable public goods, as well as the distribution of welfare transfers that can be targeted to specific households. We show that there is a positive relationship between the size of the representative's caste – the relevant ethnic group in India – and the supply of public goods, which is consistent with the inclusionary aspect of ethnic politics. Nevertheless, we find that the representative of the largest caste in the ward is only elected when his group has an absolute majority. The exclusion of outsiders from the welfare transfers apparently dominates the benefit they derive from the increased supply of public goods when that individual is elected. Additional analysis using panel data from a representative sample of households over three election terms finds that ward representatives do indeed target welfare transfers to their own group, at the expense of outsiders. And, consistent with the theory, this crowding out is especially severe when the representative belongs to a large group.

Our theory and empirical findings imply that the supply of local public goods in rural India could be increased by decoupling public good provision and the distribution of welfare transfers. This would result in more competent representatives drawn from larger castes being elected even when their caste groups do not have an absolute majority. The supply of public goods could also be increased by rolling back the caste-based reservation system at the ward level. This would result in an increase in the size distribution of eligible castes and, therefore, an increase in the average size of the elected representative's caste. However, de-reservation, without decoupling public good provision and the distribution of welfare transfers, will have adverse equity consequences because outsiders will be disproportionately excluded from receiving welfare transfers when the representative belongs to a larger caste. Counter-factual simulations based on structural estimates of the model indicate that the gains in the supply of public goods from decoupling are modest, whereas the gains from removing the reservation system are more substantial. The preferred policy in rural India, with regard to both the supply of public goods and the distribution of welfare transfers, would thus be de-reservation together with decoupling. However, in other contexts, as emphasized by the theory, decoupling could lead to a decline in the supply of public goods by reducing the incentive of more competent individuals to run for office.

7 Appendix: Proof of Proposition 2

To prove the proposition, we begin by establishing that the population share at which the incentive condition (IC) is just satisfied for the largest group declines when welfare transfers are added to the representative's list of responsibilities; i.e. $S^{**} < S^*$. Inequality (7) characterizes the condition under which the IC is satisfied with respect to the next largest group when welfare transfers are absent. The reason why we focused on the next to largest group is because the IC condition was satisfied with respect to any other (smaller) group if it was satisfied for that group. Once welfare transfers are introduced, this is not necessarily true and the IC condition must be examined with respect to each group. The largest group will prefer its representative to the representative of another group $j \neq K$ if

$$N_K \left(\frac{\beta^2}{\alpha} \right)^{\frac{\beta}{1-2\beta}} \left[N_K^{\frac{2\beta}{1-2\beta}} (1 - 2\beta) - N_j^{\frac{2\beta}{1-2\beta}} \right] + \theta N_K \left[\min \left(1, \frac{T}{N_K} \right) - \left(\frac{T - N_j}{N - N_j} \right) \right] \geq 0. \quad (16)$$

The second term in square brackets in inequality (16) characterizes the gain to the largest group, with respect to the welfare transfers, from having its representative selected. When its representative is selected, each member of the largest group receives the transfer with probability one when $N_K \leq T$ and probability $\frac{T}{N_K}$ when $N_K > T$. When the representative belongs to another group j , the probability that a member of the largest group receives a transfer, $\frac{T - N_j}{N - N_j}$, is less than $\frac{T}{N}$, which, in turn, is less than $\frac{T}{N_K}$. Note that group j never captures all of the transfers, unlike group K , because we have assumed that $T > \frac{N}{K}$ and $N_j \in (0, \frac{N}{K}]$. The second term in square brackets is thus always positive.

Applying the same argument that we used to prove Proposition 1, the first term in square brackets is negative when all groups are of equal size, $N_K = N_j = \frac{N}{K}$, positive when $N_K \rightarrow N$ and $N_j \rightarrow 0$, and increasing in N_K (as N_j declines). This implies that there is a group size $N_K^*(N_j)$ at which the first term in square brackets is exactly zero. With welfare transfers, once the positive second term in square brackets is introduced, it follows that the group size at which inequality (16) is just satisfied, $N_K^{**}(N_j)$, must be less than $N_K^*(N_j), \forall j$. The incentive condition must be satisfied for all j ; $N_K^* = \max_j N_K^*(N_j)$, $N_K^{**} = \max_j N_K^{**}(N_j)$. Without welfare transfers, the incentive condition for the largest group is most difficult to satisfy with respect to the next largest group; $\max_j N_K^*(N_j) = N_K^*(N_{K-1})$. With welfare transfers, the additional second term in square brackets in inequality (16) is increasing in N_j and so $\max_j N_K^{**}(N_j)$ is not necessarily $N_K^{**}(N_{K-1})$. Nevertheless, as long as $N_K^{**}(N_j) < N_K^*(N_j), \forall j$, it follows that $N_K^{**} \equiv \max_j N_K^{**}(N_j) < N_K^* \equiv \max_j N_K^*(N_j)$. In terms of the equivalent population shares, this implies that $S^{**} < S^*$.

If $S^{**} \geq 0.5$, the feasibility condition is irrelevant. The representative of the largest group will be selected, regardless of his outside support, when its population share reaches S^{**} because it has an absolute majority. If $S^{**} < 0.5$, then the feasibility condition must also be satisfied to ensure that the representative of the largest group is elected. The required condition is that the representative of the largest group must be preferred to the representative of any other group, $j \neq K$, by an individual belonging to neither of those groups,

$$\left\{ \left(\frac{\beta^2}{\alpha} \right)^{\frac{\beta}{1-2\beta}} \left[N_K^{\frac{2\beta}{1-2\beta}} - N_j^{\frac{2\beta}{1-2\beta}} \right] \right\} - \left\{ \theta \left[\frac{T - N_j}{N - N_j} - \max \left(\frac{T - N_K}{N - N_K}, 0 \right) \right] \right\} \geq 0. \quad (17)$$

The first term in curly brackets in the preceding inequality, which we denote by PG , characterizes the advantage of selecting the largest group's representative with respect to the supply of public goods. PG is always positive because representatives of larger groups supply more public goods. The second term in curly brackets, which we denote by WT , characterizes the accompanying disadvantage with respect to the receipt of welfare transfers. The probability that an outsider will receive a transfer when group j 's representative is in power, $\frac{T-N_j}{N-N_j}$, is decreasing in N_j . It follows that WT is always positive because $N_j < N_K, \forall j$. Note that for $N_K \geq T$, outsiders receive no transfers when the representative belongs to the largest group. In that case, $WT = \theta \frac{T-N_j}{N-N_j} > 0$.

If $PG \geq WT$, then the feasibility condition is satisfied. The discussion that follows examines this condition as the size of the largest group, N_K , varies from $\frac{N}{K}$ to N . We retain the assumption that the size of all other groups, $N_j, j \neq K$, declines as N_K increases, to keep the population N constant; i.e. $\frac{dN_j}{dN_K} < 0$ and $-\sum_j \frac{dN_j}{dN_K} = 1$. The properties of the PG and WT functions can then be derived as follows:

$$\begin{aligned} PG'(N_K(N_j)) &= \left(\frac{\beta^2}{\alpha}\right)^{\frac{\beta}{1-2\beta}} \frac{2\beta}{1-2\beta} \left[\left(\frac{1}{N_K}\right)^{\frac{1-4\beta}{1-2\beta}} - \left(\frac{1}{N_j}\right)^{\frac{1-4\beta}{1-2\beta}} \frac{dN_j}{dN_K} \right] > 0 \\ PG''(N_K(N_j)) &= -\left(\frac{\beta^2}{\alpha}\right)^{\frac{\beta}{1-2\beta}} \frac{2\beta}{1-2\beta} \frac{1-4\beta}{1-2\beta} \left[\left(\frac{1}{N_K}\right)^{\frac{2(1-3\beta)}{1-2\beta}} - \left(\frac{1}{N_j}\right)^{\frac{2(1-3\beta)}{1-2\beta}} \left(\frac{dN_j}{dN_K}\right)^2 \right] \\ PG''(N_K(N_j)) \leq 0 \quad &\text{for } \frac{N_K}{N_j} \leq \left(\frac{-1}{\frac{dN_j}{dN_K}}\right)^{\frac{1-2\beta}{1-3\beta}} \\ PG''(N_K(N_j)) > 0 \quad &\text{for } \frac{N_K}{N_j} > \left(\frac{-1}{\frac{dN_j}{dN_K}}\right)^{\frac{1-2\beta}{1-3\beta}} \end{aligned}$$

For $N_K \leq T$:

$$\begin{aligned} WT'(N_K(N_j)) &= \theta(N-T) \left[\frac{1}{(N-N_K)^2} - \frac{\frac{dN_j}{dN_K}}{(N-N_j)^2} \right] > 0 \\ WT''(N_K(N_j)) &= 2\theta(N-T) \left[\frac{1}{(N-N_K)^3} - \frac{\left(\frac{dN_j}{dN_K}\right)^2}{(N-N_j)^3} \right] > 0 \end{aligned}$$

For $N_K > T$:

$$\begin{aligned} WT'(N_K(N_j)) &= -\theta(N-T) \frac{\frac{dN_j}{dN_K}}{(N-N_j)^2} > 0 \\ WT''(N_K(N_j)) &= -2\theta(N-T) \frac{\left(\frac{dN_j}{dN_K}\right)^2}{(N-N_j)^3} < 0. \end{aligned}$$

It is straightforward to verify, from inequality (17), that $PG = WT = 0$ when all groups are of equal size; i.e. $N_K = N_j = \frac{N}{K}$. As shown above, both PG and WT are monotonically increasing in $N_K(N_j)$. Given the assumption that $\beta \in (0, 1/4)$, PG is initially a concave function of $N_K(N_j)$ and, subsequently, past a point of inflection $N^I(N_j)$, a convex function of $N_K(N_j)$. In contrast, WT is a convex function of $N_K(N_j)$ for $N_K \leq T$ and a concave function of $N_K(N_j)$ for $N_K > T$.

Because $PG = WT = 0$ when all groups are of equal size, and PG is initially concave while WT is convex, the following conditions ensure, in the most natural way, that the two functions cross at least once.

C1. $PG'(N_K(N_j)) > WT'(N_K(N_j))$ as $N_K(N_j) \rightarrow \frac{N}{K}$.

C2. $WT(N_K(N_j)) > PG(N_K(N_j))$ as $N_K(N_j) \rightarrow N$.

We can establish, in addition, that there is a unique point of intersection, $N^F(N_j)$. The proof of uniqueness is by contradiction. Suppose that there are multiple points of intersection. Given that the PG function starts above the WT function as $N_K(N_j) \rightarrow \frac{N}{K}$ and ends below it at $N_K(N_j) \rightarrow N$, there must then be at least three points of intersection. This implies that there must be *at least three* values of $N_K(N_j)$ for which $PG'(N_K(N_j)) = WT'(N_K(N_j))$. Given the properties of the PG and WT functions, as described above, $PG'(N_K(N_j))$ is monotonically decreasing in $N_K(N_j)$ for $N_K(N_j) \leq N^I(N_j)$ and monotonically increasing in $N_K(N_j)$ for $N_K(N_j) > N^I(N_j)$. In contrast, $WT'(N_K(N_j))$ is monotonically increasing in $N_K(N_j)$ for $N_K(N_j) \leq T$ and monotonically decreasing in $N_K(N_j)$ for $N_K(N_j) > T$. Based on this characterization of $PG'(N_K(N_j))$, $WT'(N_K(N_j))$, it is evident that these functions can intersect at most twice; i.e. there are *at most two* values of $N_K(N_j)$ for which $PG'(N_K(N_j)) = WT'(N_K(N_j))$, which is a contradiction.

Based on the preceding discussion, $PG \geq WT$ and the feasibility condition is satisfied for $N_K(N_j) \leq N^F(N_j)$. $PG < WT$ and the feasibility condition is not satisfied for $N_K(N_j) > N^F(N_j)$. The feasibility condition must be satisfied with respect to all j , hence, $N^F = \min_j N^F(N_j)$. The equivalent population share threshold is $S^F \equiv \frac{N^F}{N}$. If $S^{**} < S^F$, then the threshold at which the representative's ability and the supply of public goods increase discontinuously is located at S^{**} . If $S^{**} > S^F$, then the threshold will be located at 0.5; i.e. only when the largest group has an absolute majority, because the feasibility condition (for all j) will not be satisfied for any $S_K > S^F$.

References

- [1] Alesina, Alberto and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance," *Journal of Economic Literature* 43(3): 762-800.
- [2] Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. Forthcoming. "Ethnic Inequality," *Journal of Political Economy*.
- [3] Anderson, Siwan. 2011. "Caste as an Impediment to Trade," *American Economic Journal: Applied* 3(1):239-263.
- [4] Anderson, Siwan, Patrick Francois, and Ashok Kotwal. 2015. "Clientilism in Indian Villages," *American Economic Review* 105(6): 1780-1816.
- [5] Andrews, Donald W.K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* 61(4): 821-856.
- [6] Angelucci, Manuela, Giacomo De Giorgi, and Imran Rasul. 2015. "Resource Pooling within Family Networks: Insurance and Investment," University of Michigan, typescript.
- [7] Banerjee, Abhijit and Rohini Pande. 2010. "Parochial Politics: Ethnic Preferences and Politician Corruption," Harvard University, Kennedy School of Government, typescript.
- [8] Bardhan, Pranab, Dilip Mookherjee, and Monica Parra Torrado. 2010. "Impact of Political Reservations in West Bengal Local Governments on Anti-Poverty Targeting," *Journal of Globalization and Development* 1(1): 1-36.
- [9] Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Prejudice?" *Quarterly Journal of Economics* 124(4): 1497-1540.
- [10] Besley, Timothy. 2004. "Paying Politicians: Theory and Evidence," *Journal of the European Economic Association* 2(2-3): 193-215.
- [11] Besley, Timothy and Stephen Coate. 1997. "An Economic Model of Representative Democracy," *Quarterly Journal of Economics* 112(1): 85-114.
- [12] Besley, Timothy, Rohini Pande, and Vijayendra Rao. 2007. "Just Rewards? Local Politics and Public Resource Allocation in South India," Harvard University, Kennedy School of Government, typescript.
- [13] Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padro i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya," *American Economic Review* 105(6): 1817-1851.

- [14] Caldwell, John C., P.H. Reddy, Pat Caldwell. 1986. "Periodic High Risk as a Cause of Fertility Decline in a Changing Rural Environment: Survival Strategies in the 1980-1983 South Indian Drought," *Economic Development and Cultural Change* pp. 677-701.
- [15] Casey, Katherine. 2015. "Crossing Party Lines: The Effects of Information on Redistributive Politics," *American Economic Review* 105(8): 2410-2448.
- [16] Chandavarkar, Rajnarayan. 1994. *The Origins of Industrial Capitalism in India: Business strategies and the working classes in Bombay, 1900-1940*. Cambridge: Cambridge University Press.
- [17] Chattopadhyay, Raghabendra and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica* 72(5): 1409-1443.
- [18] Easterley, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics* 112(4): 1203-1250.
- [19] Eldersveld, Samuel J. and Bashiruddin Ahmed. 1978. *Citizen and Politics: Mass Political Behavior in India*. Chicago: The University of Chicago Press.
- [20] Fafchamps, Marcel and Susan Lund. 2003. "Risk Sharing Networks in Rural Philippines," *Journal of Development Economics* 71: 261-287.
- [21] Ferraz, Claudio and Frederico Finan. 2011. "Electoral Accountability and Corruption: Evidence from the Audit Reports of Local Governments," *American Economic Review*, 101: 1274-1311.
- [22] Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: the Maghribi Traders' Coalition," *American Economic Review* 83(3):525-48.
- [23] Greif, Avner, Paul Milgrom, and Barry R. Weingast. 1994. "Coordination, Commitment, and Enforcement: The Case of the Merchant Guild," *Journal of Political Economy* 102(4): 745-776.
- [24] Grimard, Franque. 1997. "Household Consumption Smoothing through Ethnic Ties: Evidence from Cote d'Ivoire," *Journal of Development Economics* 53(3): 319-422.
- [25] Hansen, Bruce E. 1999. "Threshold effects in non-dynamic panels: Estimation, testing, and inference," *Journal of Econometrics* 93: 345-368.
- [26] Ligon, Ethan, Jonathan P. Thomas and Tim Worrall. 2002. "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies," *Review of Economic Studies* 69: 209-244.
- [27] Lizzeri, Alessandro and Nicola Persico. 2001. "The Provision of Public Goods Under Alternative Electoral Incentives," *American Economic Review* 91(1): 226-239.
- [28] Mazzocco, Maurizio and Shiv Saini. 2012. "Testing Efficient Risk-Sharing with Heterogeneous Risk Preferences," *American Economic Review* 102(1): 428-468.

- [29] Miguel, Edward and Mary Kay Gugerty. 2005. “Ethnic Diversity, Social Sanctions, and Public Goods in Kenya,” *Journal of Public Economics* 89: 2325-2368.
- [30] Moorjani, Priya, Kumarasamy Thangaraj, Nick Patterson, Mark Lipson, Po-Ru Loh, Periyasamy Govindaraj, Bonnie Berger, David Reich, and Laiji Singh. 2013. “Genetic Evidence for Recent Population Mixture in India,” *American Journal of Human Genetics* 93(3):422-438.
- [31] Morris, Morris David. 1965. *The Emergence of an Industrial Labor Force in India: A Study of the Bombay Cotton Mills, 1854-1947*. Berkeley: University of California Press.
- [32] Munshi, Kaivan and Mark Rosenzweig. 2006. “Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy,” *American Economic Review* 96(4):1225-1252.
- [33] Munshi, Kaivan. 2011. “Strength in Numbers: Networks as a Solution to Occupational Traps,” *Review of Economic Studies* 78:1069-1101.
- [34] Munshi, Kaivan and Mark Rosenzweig. Forthcoming. “Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap,” *American Economic Review*.
- [35] Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*, Cambridge, Massachusetts: Harvard University Press.
- [36] Osborne, Martin and Al Slivinski. 1996. “A Model of Political Competition with Citizen-Candidates,” *Quarterly Journal of Economics* 111(1): 65-96.
- [37] Padro i Miquel, Gerard. 2007. “The Control of Politicians in Divided Societies: The Politics of Fear,” *Review of Economic Studies* 74(4): 1259-1274.
- [38] Townsend, Robert. 1994. “Risk and Insurance in Village India,” *Econometrica* 62(3): 171-84.

Figure 1: The Importance of Caste within the Village

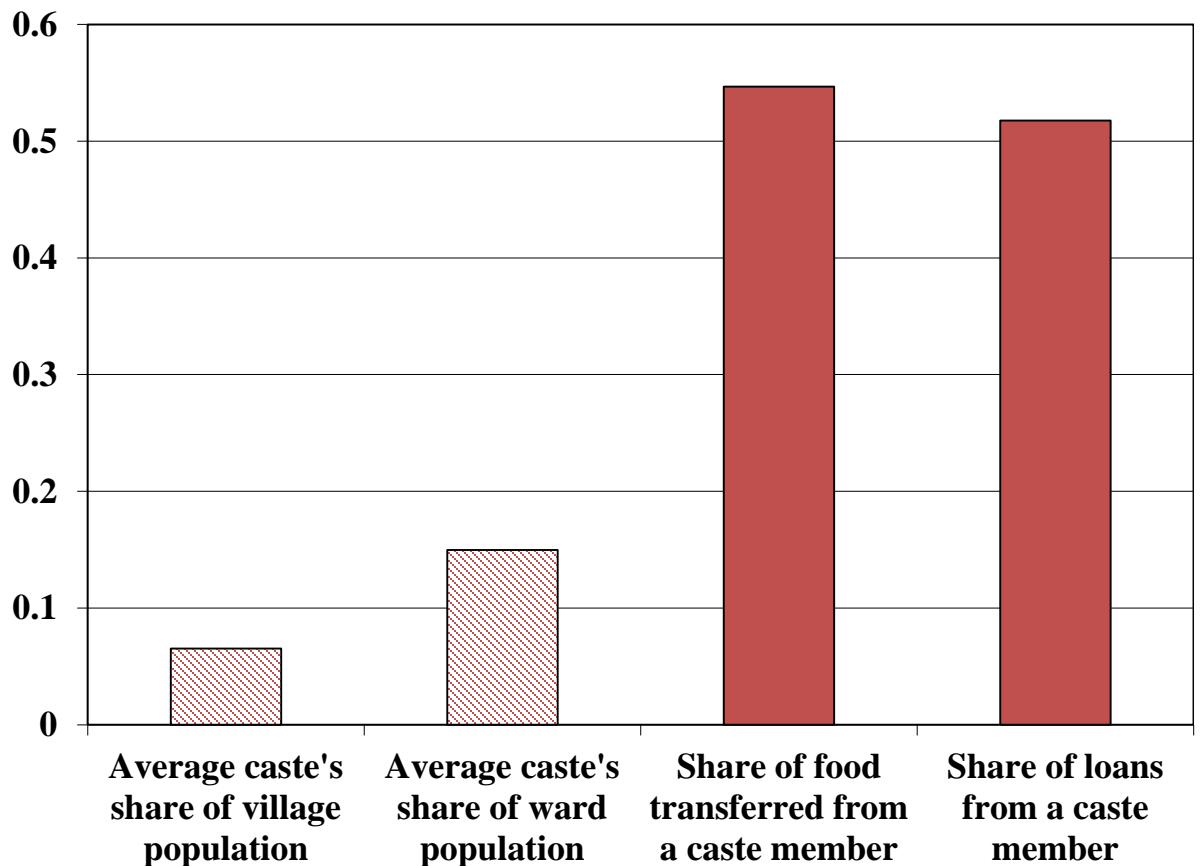
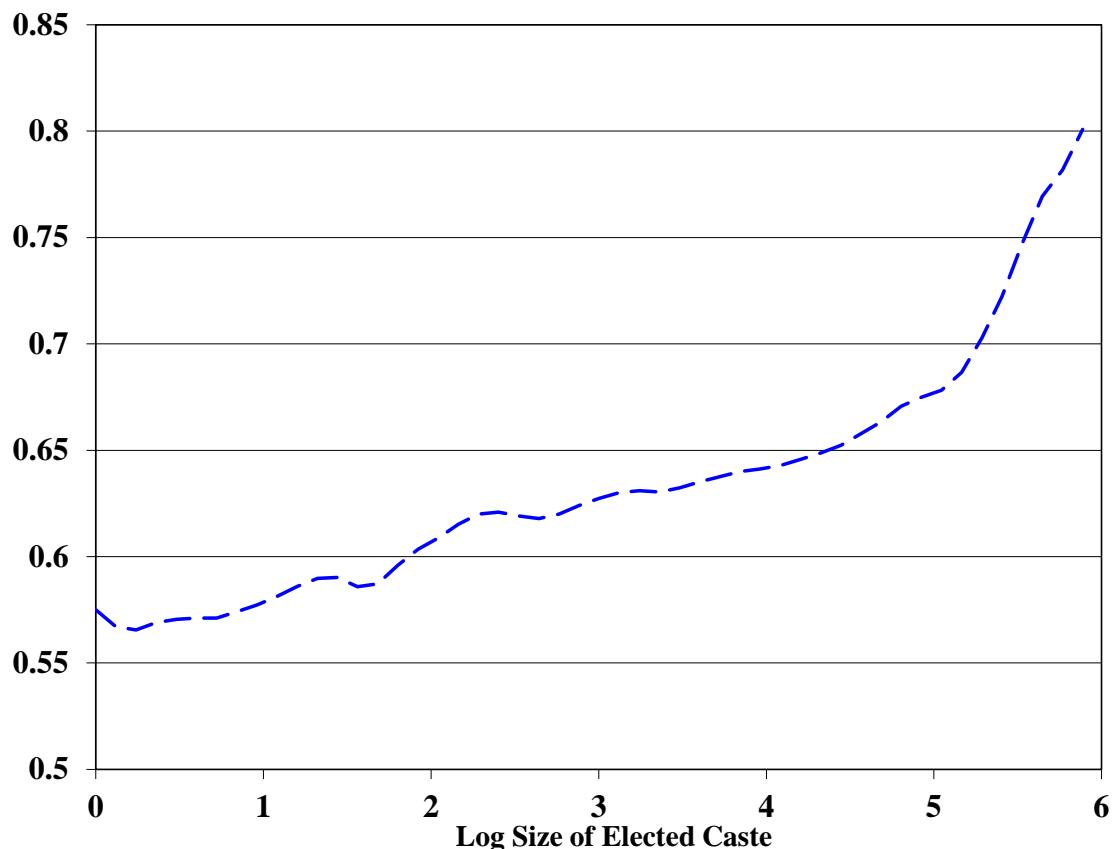


Figure 2: Proportion of the Caste with Schooling Below that of its Representative, by the Log of Caste Size



**Figure 3: Public Goods and the Elected Representative's Schooling (Years),
by the Population Share of the Largest Eligible Caste**

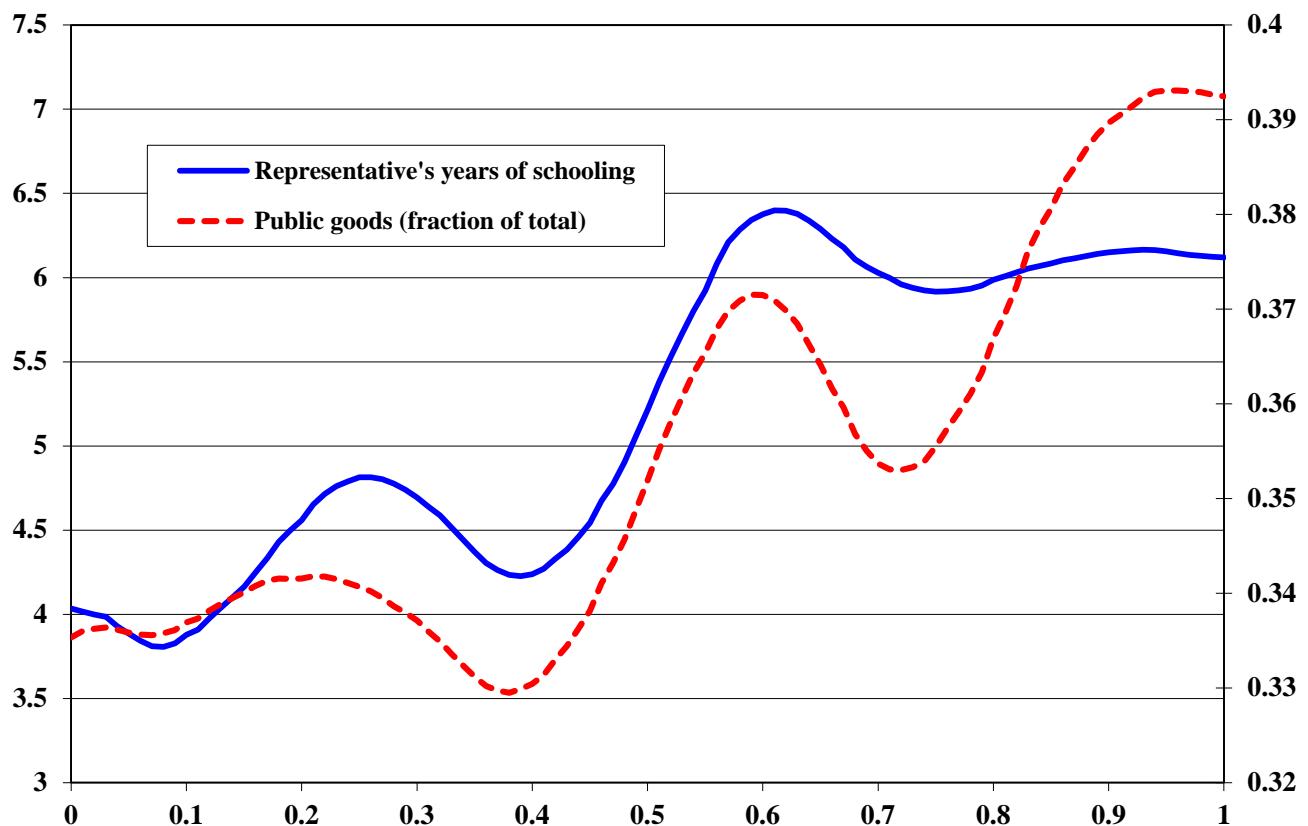


Figure 4: Simulated Variables V_S and V_L , by the Population Share of the Largest Caste

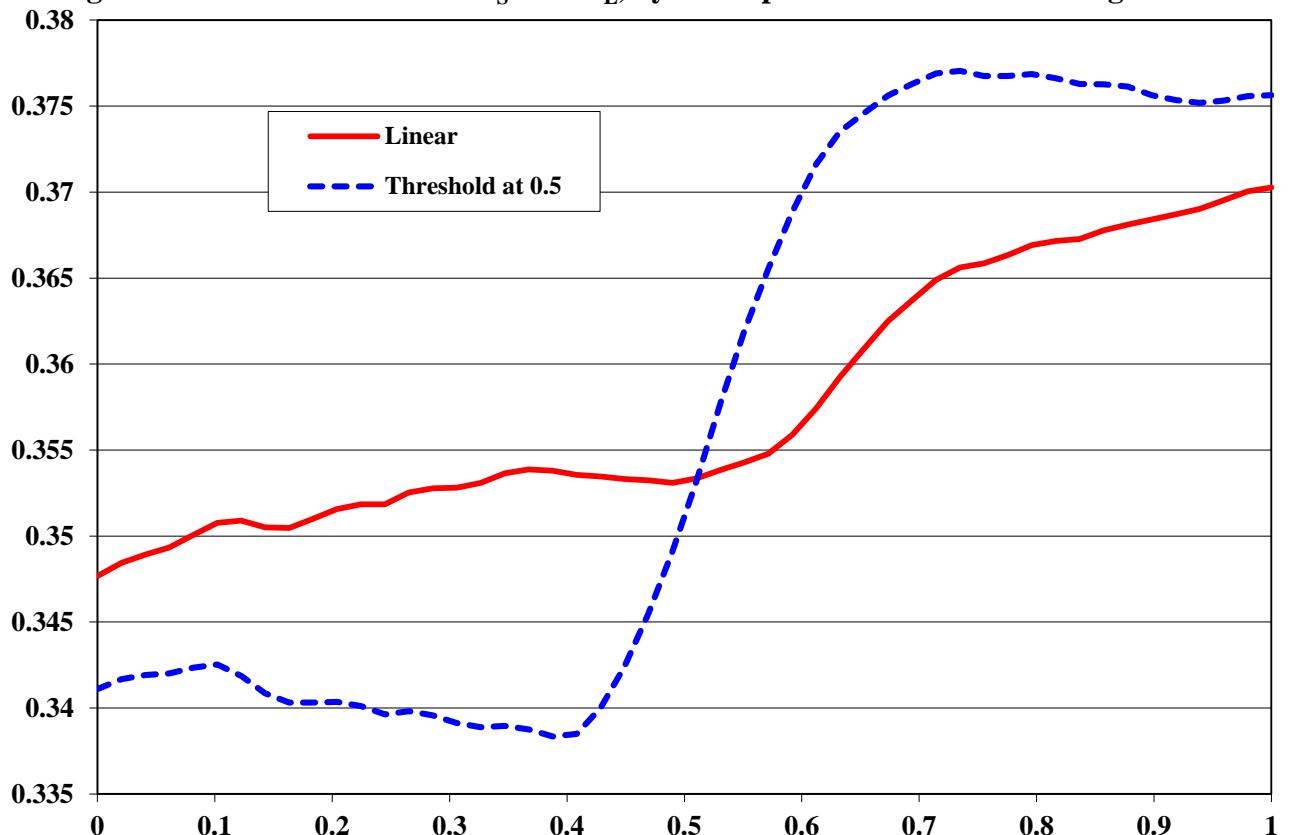


Figure 5: Threshold Tests on Simulated Data: Step versus Linear Function Likelihood Ratio by Hypothesized Threshold

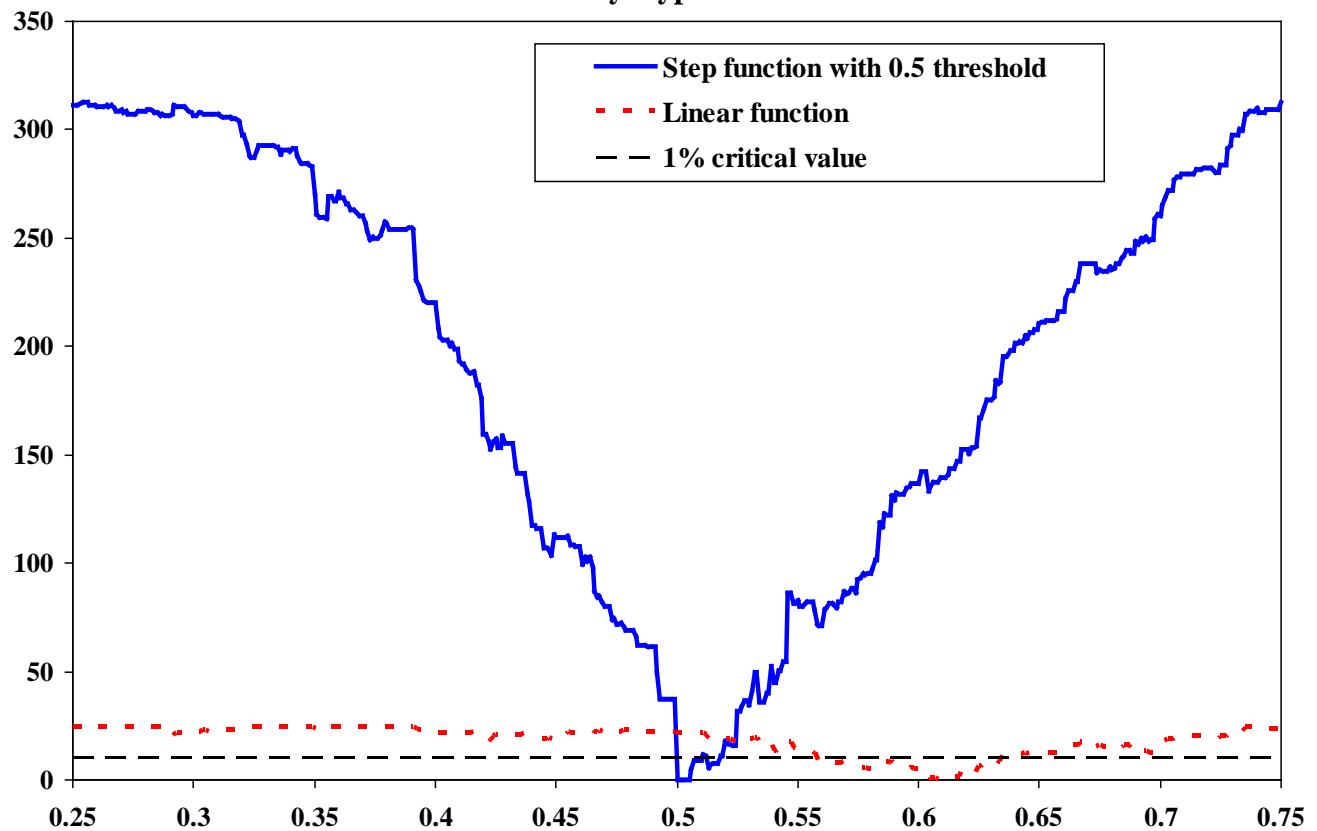
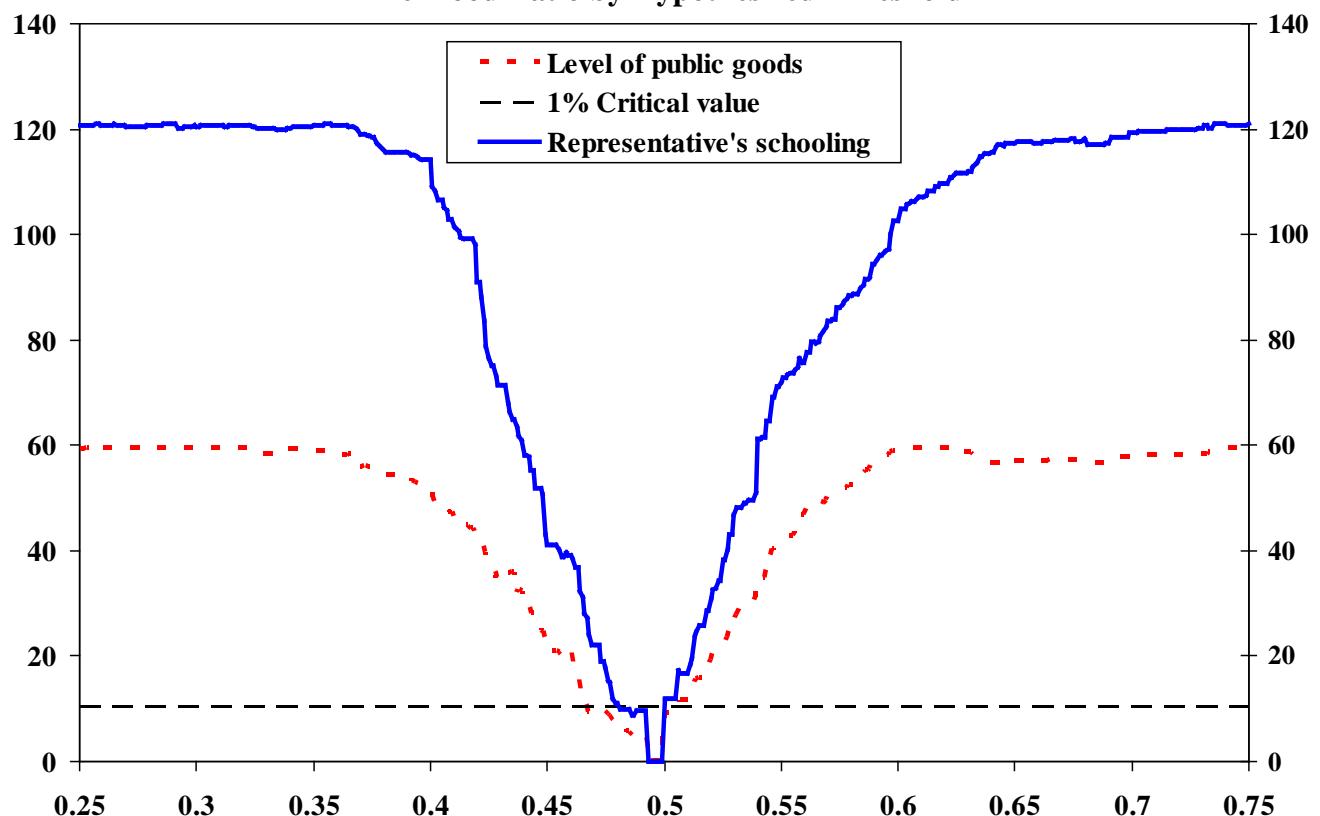
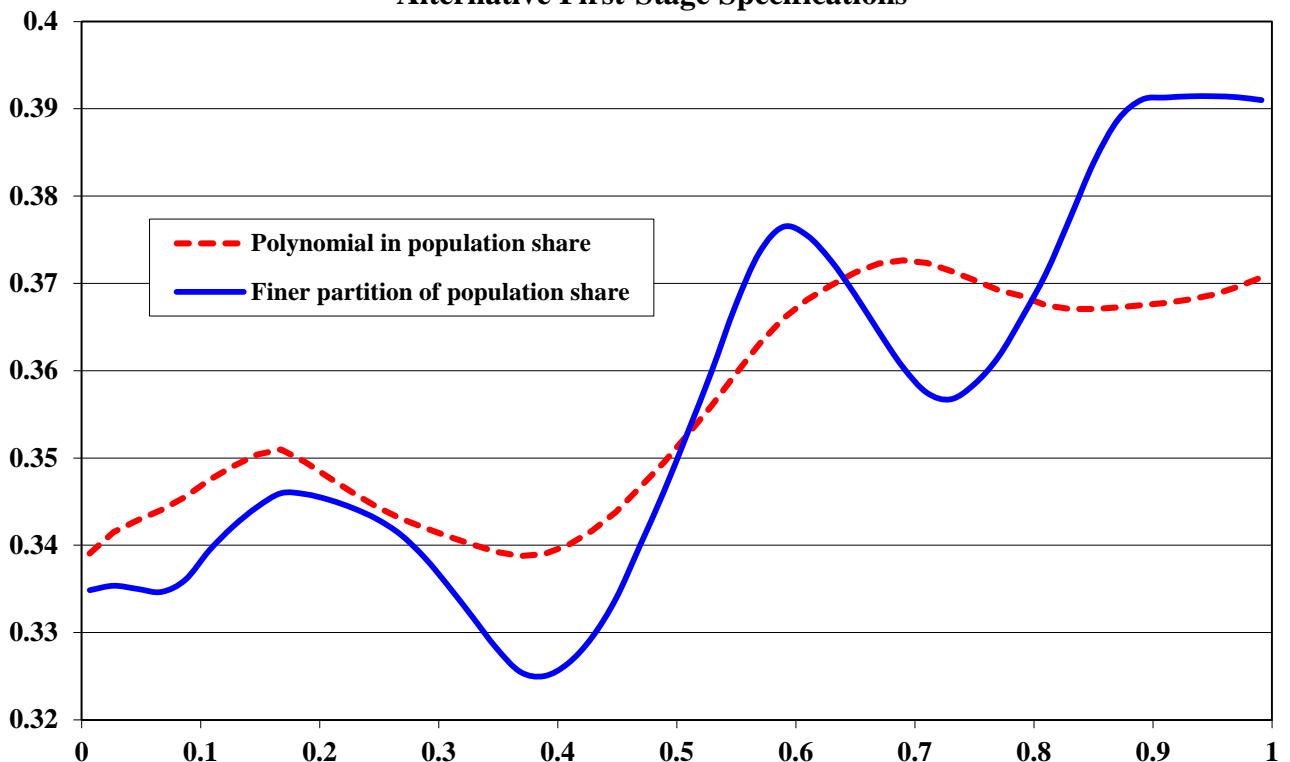


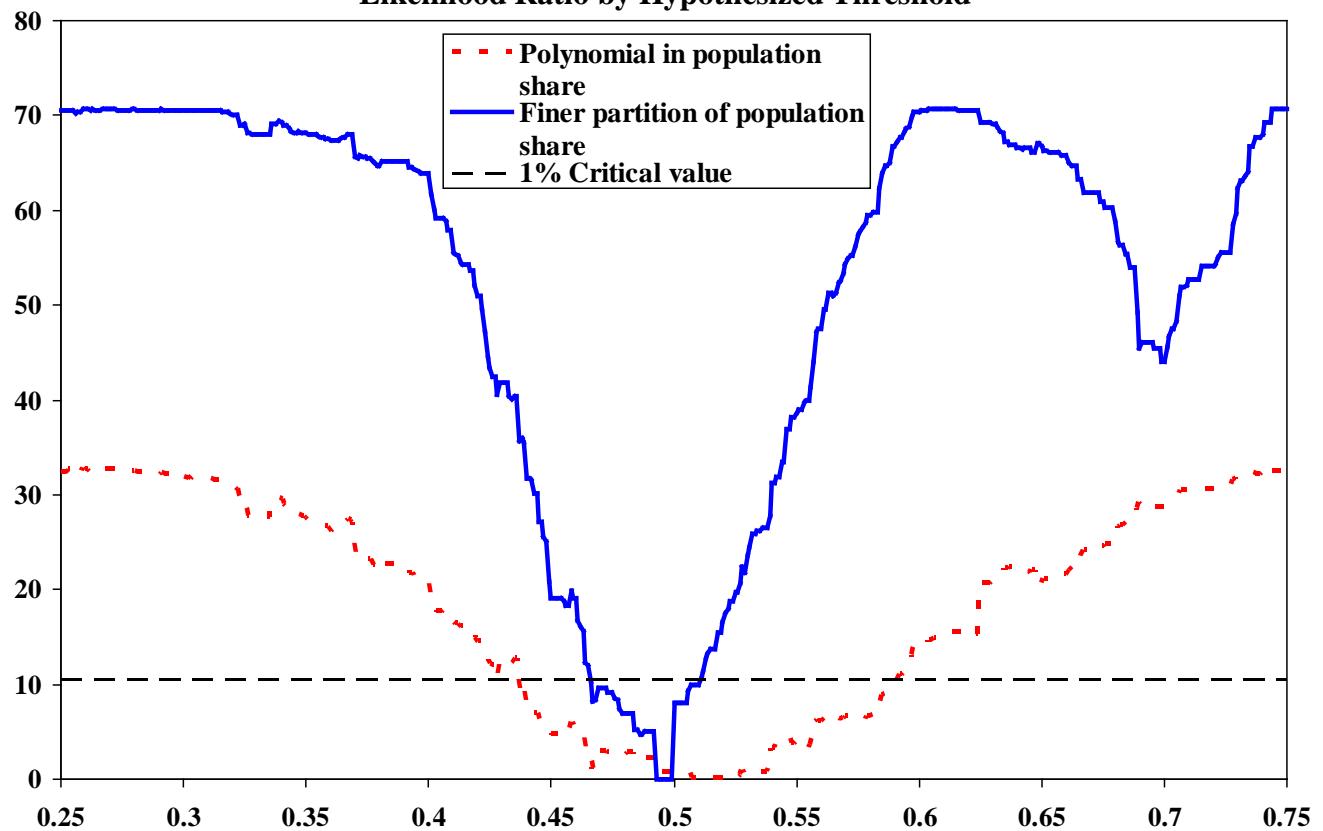
Figure 6: Threshold Tests: Public Goods and the Elected Representative's Schooling Likelihood Ratio by Hypothesized Threshold



**Figure 7: Public Goods, by the Population Share of the Largest Caste:
Alternative First-Stage Specifications**



**Figure 8: Threshold Tests: Alternative First-Stage Specifications
Likelihood Ratio by Hypothesized Threshold**



**Figure 9: Threshold Tests: Alternative Threshold Functions
Likelihood Ratio by Hypothesized Threshold**

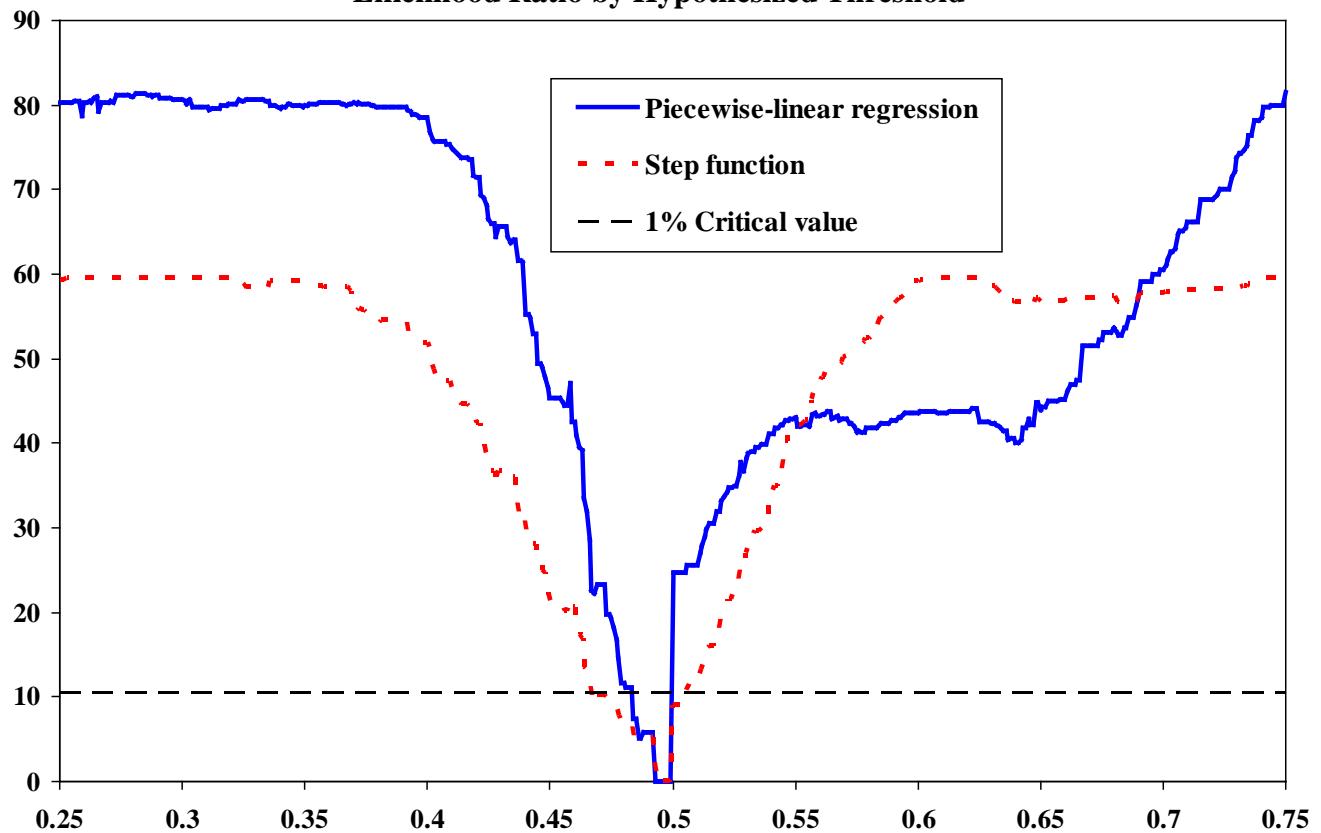


Figure 10: Caste Population Schooling (Years) Distribution, by Caste Population Share

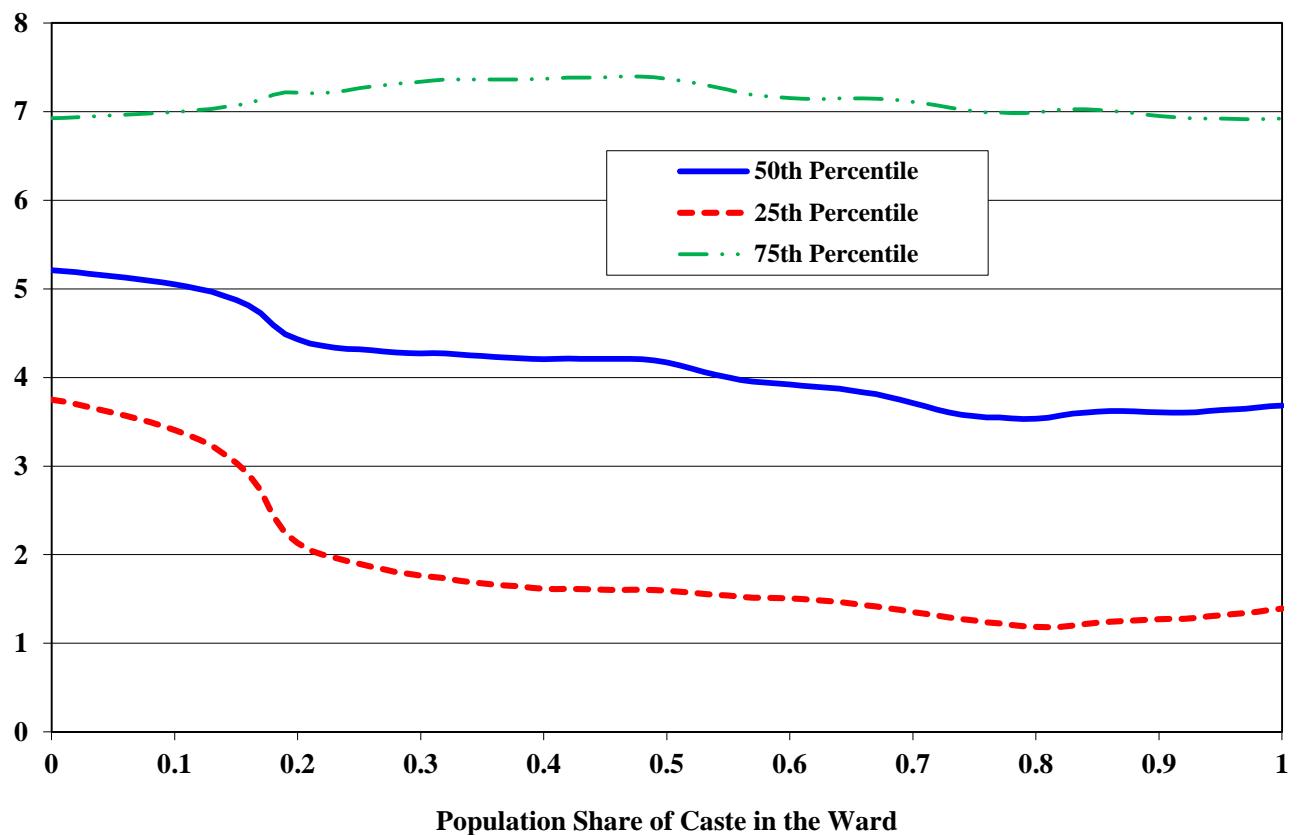


Figure 11: Threshold Tests: Education Distribution in the Caste Likelihood Ratio by Hypothesized Threshold

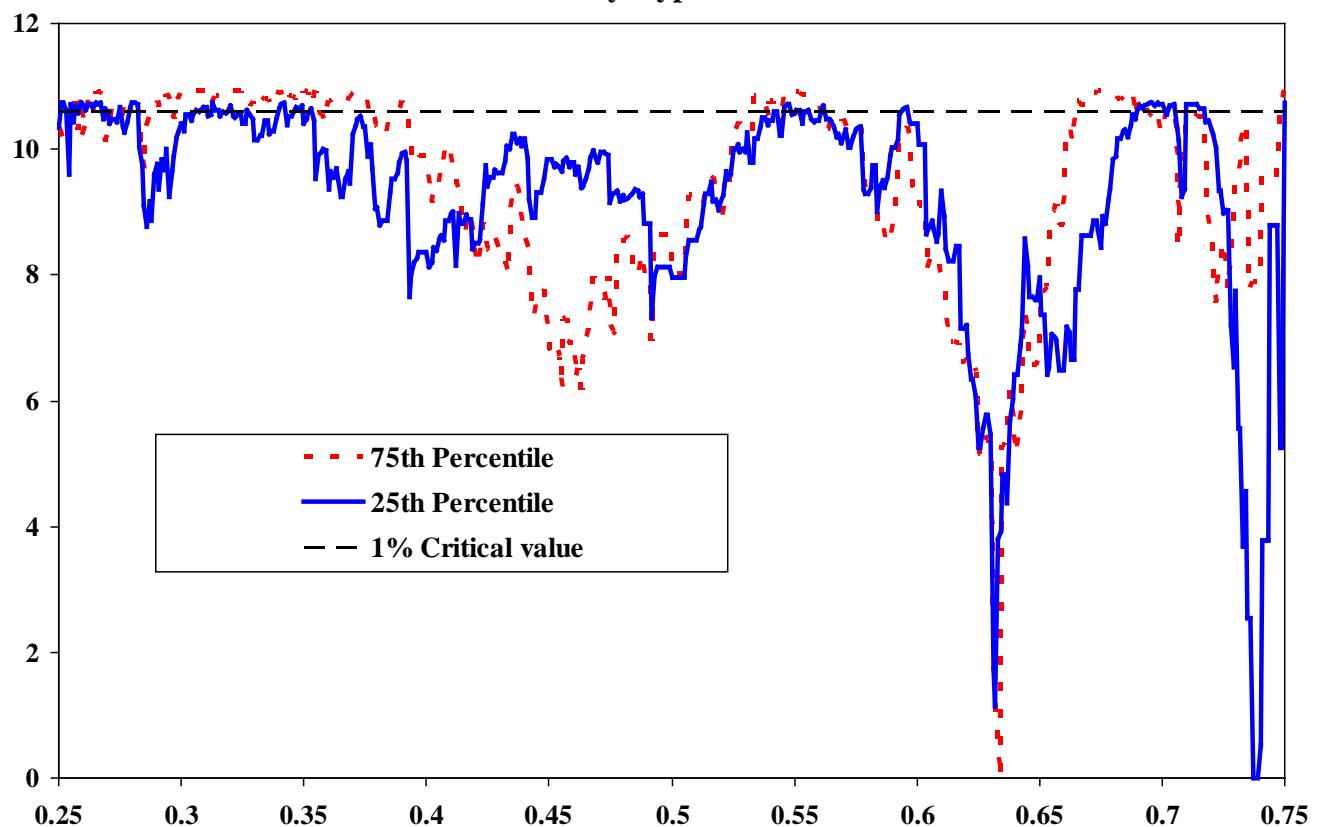


Figure 12: Public Goods and the Share of Households Receiving BPL Transfers, by the Population Share of the Largest Eligible Caste

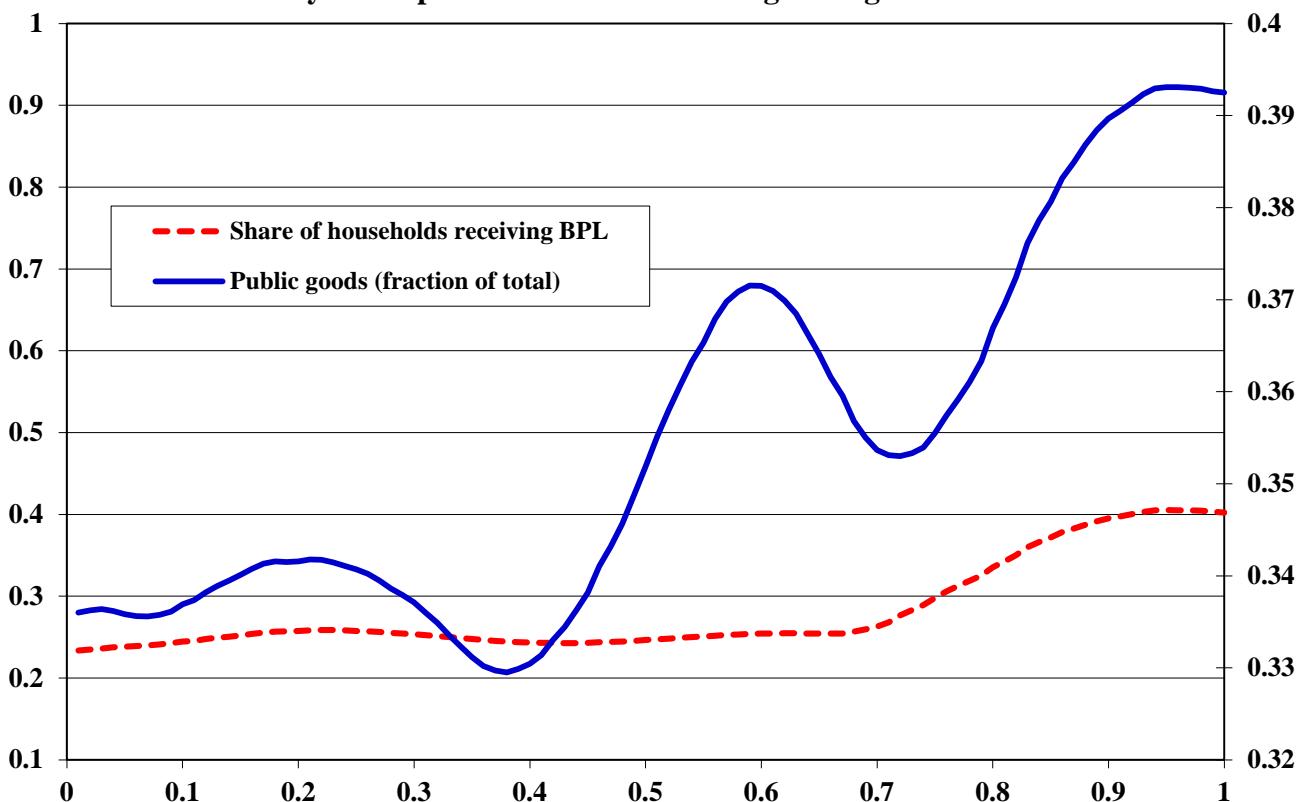
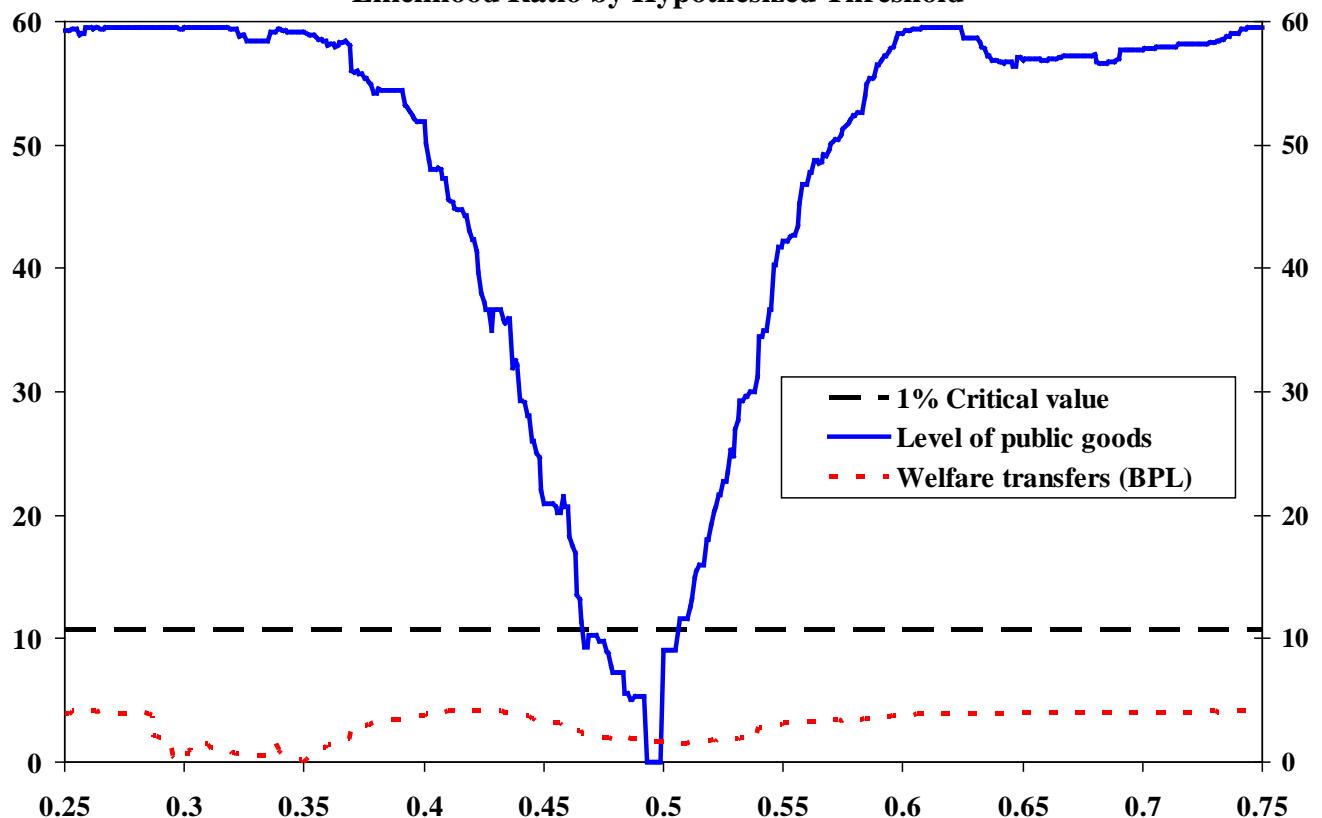
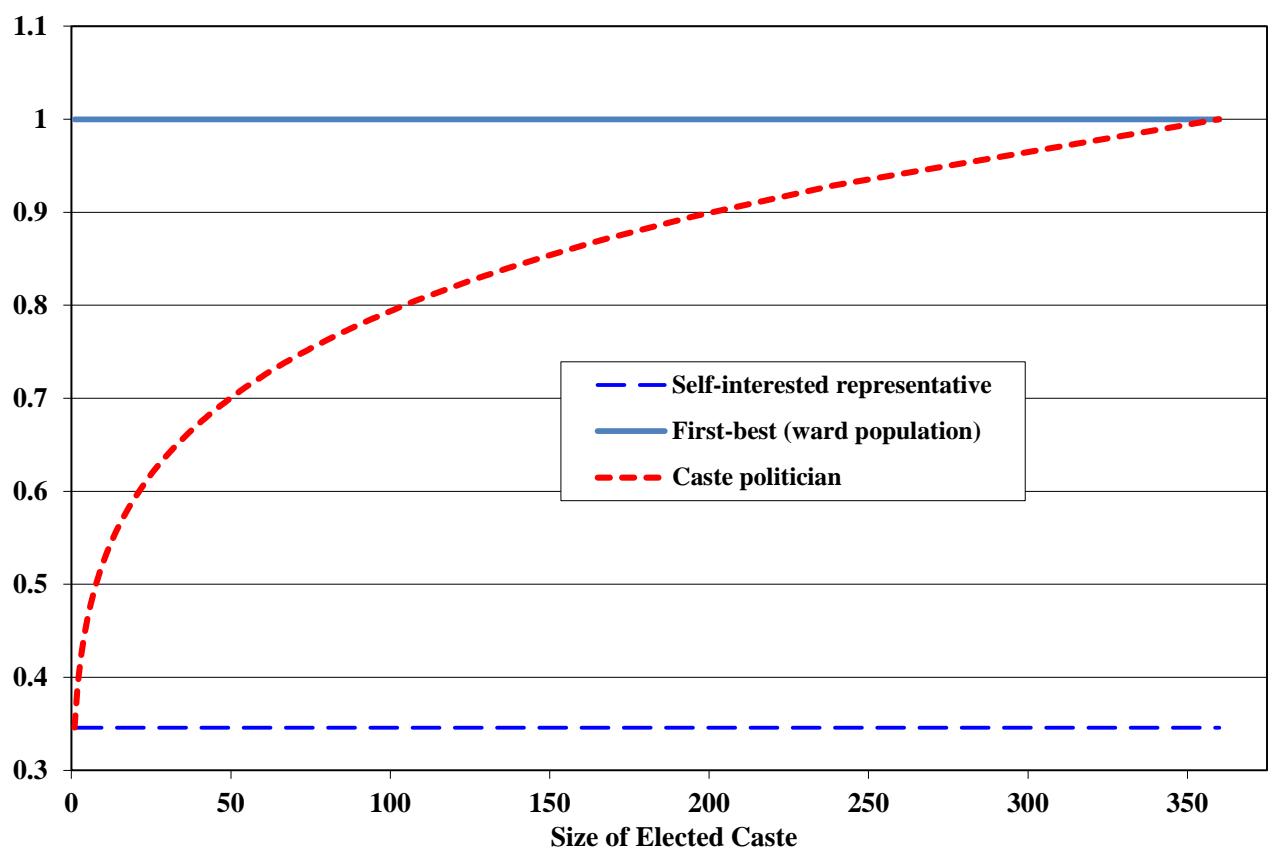


Figure 13: Threshold Tests: Public Goods and Welfare Transfers (BPL)
Likelihood Ratio by Hypothesized Threshold



**Figure 14: Public Goods Level, by the Size of the Elected Caste
Relative to Two Benchmarks**



**Figure 15: Counterfactual-Simulation Cumulative Distributions:
Fraction of Public Goods Received by a Ward**

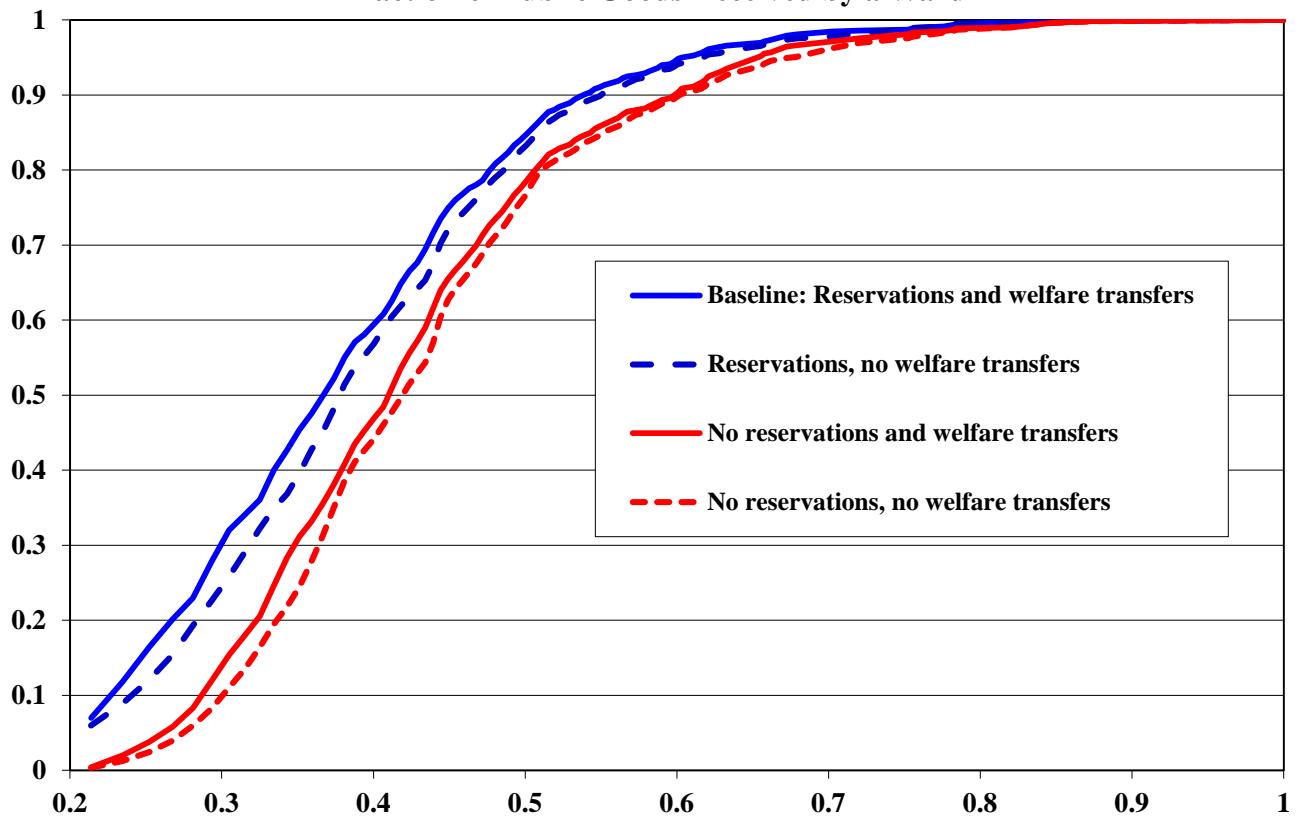


Table 1: Sources of Support for Ward Representatives

| Source of support (%): | Within village | Outside village |
|--------------------------|-------------------|--------------------|
| | (1) | (2) |
| From caste | 82 | 29 |
| From religion | 28 | 13 |
| From wealthy individuals | 38 | -- |
| From a political party | -- | 41 |

Source: Rural Economic Development Survey, 2006.

The statistics are computed over the last three election terms in each ward.

Each statistic reflects the percent of representatives who received support from a given source.

**Table 2: Probability the Winning Representative is from the Largest and Second-Largest Castes,
Last Three Panchayat Ward Elections Before 2007
(Sources: REDS 2006 Listing and 2007 Village Survey)**

| | Winner from Largest Caste | Winner from Second-Largest Caste |
|---|---------------------------|----------------------------------|
| Population share of the largest eligible caste is $\geq .5$ | .892 | .063 |
| Population share of the largest eligible caste is $< .5$ | .606 | .225 |
| <i>t</i> -test statistic, difference | 11.8 | 8.03 |
| N | 1,114 | 1,114 |

**Table 3: Which Local Groups Respondents Went to When They Faced a Problem,
by Problem Type (Percent)**

| Problem/source for solution | Enforcement of rental and share contracts | Protection of crops against theft and destruction |
|-------------------------------------|---|---|
| Own caste | 2.3 | 64.8 |
| Political organization | 88.7 | 12.5 |
| Dominant caste in village | 0 | 3.4 |
| Own religious group | 6.0 | 11.4 |
| Dominant religious group in village | 2.3 | 4.6 |

**Table 4: Schooling, Managerial Experience, Land Wealth and Voting Behavior,
Male Heads 25-60
(Source: 2006 REDS Listing Data)**

| Variable | Managerial Occupation | Land Wealth | Voted in Last Panchayat Election |
|----------------------|--------------------------|----------------|-------------------------------------|
| Estimation Procedure | Conditional Logit | Fixed Effects | Conditional Logit |
| Schooling | .0261 (.00758) | 8658 (3447) | -.0167 (.00464) |
| Age | .109 (.0167) | 6268 (9257) | .0338 (.0136) |
| Age squared | -.00083 (.000175) | 14.6 (70.1) | -.000026 (.000011) |
| N | 76,428 | 76,428 | 76,428 |

Standard errors in parentheses clustered at the village level. All specification include fixed effects for ward/caste.

Table 5: Mean Years of Schooling of Elected Representatives and Median Household Heads in the Wards, by Type of Election

| Election type: | Open (1) | SC (2) | ST (3) | OBC (4) |
|---------------------------|----------------|----------------|----------------|----------------|
| Median household head | 4.46 (3.76) | 3.53 (3.59) | 3.30 (3.65) | 4.03 (3.36) |
| Male representatives | 7.42 (4.43) | 6.01 (4.49) | 5.30 (3.99) | 7.05 (4.30) |
| Female representatives | 3.23 (3.83) | 5.78 (4.39) | 2.22 (2.05) | 4.72 (4.17) |

Median household schooling, by reservation category in each ward, is computed from the village household census. The mean (standard deviation) across all wards is reported in the table.

Representative's education is obtained for the last three terms in each ward.

Gender reservation is independent of caste reservation.

Table 6: Household Fixed Effects Estimates: Targeting of Public Resources

| Dependent variable | Household receives BPL transfers | | Public goods placed on household's street | |
|--|----------------------------------|----------------------|---|------------------------|
| | (1) | (2) | (3) | (4) |
| Household belongs to the representative's caste | 1.377** (0.577) | 0.213 (0.752) | -- | -- |
| Fraction of households belonging to representative's caste on street | - | - | -0.00915 (0.0573) | -0.0451 (0.106) |
| Size of representative's caste | -0.0192 (0.0190) | -0.0308* (0.0158) | 0.000463 (0.000957) | -0.000268 (0.00152) |
| Interaction ^a | -- | 0.053*** (0.0201) | -- | 0.00151 (0.00264) |
| Mean of dependent variable | 0.258 (0.438) | 0.258 (0.438) | 0.343 (0.228) | 0.343 (0.228) |
| N | 1387 | 1387 | 1387 | 1387 |

* p < 0.1, ** p < 0.05, *** p < 0.01.

Mean size of representative's caste is 37; standard deviation is 41.

Public goods are measured by the fraction of the six major public goods received on the household's street in each election term.

All specifications include household fixed effects, reservation dummies, term dummies, and election year.

^aIn column two the interaction variable is the representative's caste size X household belongs to the representative's caste. In column four the interaction variable is representative's caste size X the fraction of households in the representative's caste on the household's street.

Columns 1 and 2 are estimated using conditional logit.

Standard errors are clustered at the ward level.

Table 7: Population Share of the Largest Eligible Caste, by Election Type

| Type of election: | Open (1) | SC (2) | ST (3) | OBC (4) |
|---|-------------|-----------|-----------|------------|
| <u>Panel A: Distribution of shares</u> | | | | |
| 25 percentile | 0.42 | 0.14 | 0.16 | 0.20 |
| 50 percentile | 0.60 | 0.33 | 0.57 | 0.41 |
| 75 percentile | 0.85 | 0.65 | 0.95 | 0.69 |
| <u>Panel B: Fraction of ward-terms where largest share exceeds</u> | | | | |
| 0.25 | 0.97 | 0.48 | 0.64 | 0.69 |
| 0.5 | 0.66 | 0.29 | 0.50 | 0.44 |
| 0.75 | 0.36 | 0.18 | 0.41 | 0.21 |
| Fraction of ward-terms | 0.60 | 0.11 | 0.06 | 0.23 |

SC=scheduled caste, ST=scheduled tribe, OBC=other backward caste.

Population shares are measured in each ward-term.

Information on reservation status and population shares in each ward are obtained over three election terms.

Table 8: Change in Public Goods Levels and Representative's Schooling at the Population Share Threshold of 0.5

| Dependent variable | Public goods | | Representative's schooling (years) | |
|---|--------------------|--------------------|------------------------------------|---------------------|
| | (1) | (2) | | |
| Mean-shift at threshold for largest eligible single caste | 0.042** (0.021) | 0.058** (0.022) | 2.264*** (0.802) | 2.461*** (0.845) |
| Mean shift at threshold for the eligible caste group | - | -0.020 (0.018) | - | -0.0249 (0.672) |
| Threshold location | 0.50 | 0.50 | 0.50 | 0.50 |
| Mean of dependent variable below the threshold | 0.335 [0.081] | 0.335 [0.081] | 4.157 [3.034] | 4.157 [3.034] |
| N | 1666 | 1666 | 1591 | 1591 |

* p < 0.1, ** p < 0.05, *** p < 0.01.

All specifications include ward fixed effects, reservation dummies, term dummies, and election year.

Public goods are measured as the fraction of the six major goods that were received in each ward-term.

Each equation is estimated at the population share threshold derived from the Likelihood Ratio test.

Standard errors are clustered at the village/term level.

Table 9: Change in the Expenditure Probability at the Population Share Threshold of 0.5 for Six Public Goods
(Sources: REDS 2006 Listing and 2007 Village Survey)

| Public Good | Sanitation | Street Lights | Clean Water | Road Improvement | Communication | Electricity |
|---|------------------|------------------|------------------|------------------|------------------|------------------|
| Mean-shift at threshold | .0873 (.0328) | .0773 (.0198) | .0534 (.0303) | .0242 (.0343) | .0273 (.0171) | .0248 (.0244) |
| Number of observations | | | | 1666 | | |
| H_0 : no joint effect at the threshold, $\chi^2(6)$ [p] | | | | 26.7 [.0002] | | |
| H_0 : all effects are equal, $\chi^2(5)$ [p] | | | | 8.53 [.130] | | |
| Mean of dependent variable below the threshold | .483 | .366 | .748 | .609 | .222 | .369 |

All specifications include ward fixed effects, reservation dummies, term dummies, and election year.

Table 10: Structural Parameter Estimates

| Dependent variable | Log public goods | |
|--|----------------------|----------------------|
| Parameter | | |
| α | 0.0148 (0.0128) | - |
| α Upper castes | - | 0.945 (.746) |
| α Lower castes | | 0.0151 (.0144) |
| β | 0.122*** (0.0450) | 0.122*** (0.0450) |
| $H_0: \beta < 0, p$ | .004 | .004 |
| $H_0: \beta > 0.25, p$ | .000 | .000 |
| $H_0: 2\beta/(1 - 2\beta) = 0, p$ | .041 | .041 |
| $H_0: \alpha$ Upper, α Lower $< 0, p$ | - | .043 |
| $H_0: \alpha$ Upper = α Lower, p | - | .217 |
| N | 1957 | 1957 |

* p < 0.1, ** p < 0.05, *** p < 0.01.

Public goods measured by the log of the fraction of the six major public goods received on the household's street in each election term
Standard errors clustered at the ward level.

**Appendix Table A1: Budget Shares by Discretionary Expenditure Item
for the Complete Panchayat 5-Year Term,
(Previous to Most Recent Panchayat)**

| Expenditure Item | Expenditure Share (Percent) |
|------------------------------------|-----------------------------|
| Road improvement | 25.2 |
| Drinking water | 20.0 |
| Street lighting | 8.9 |
| School improvement | 8.0 |
| Sanitation and sewage | 7.2 |
| Ceremonies | 6.5 |
| Health facilities | 5.9 |
| Electrification | 4.5 |
| Natural Resource Management | 4.0 |
| Irrigation | 3.4 |
| Access to local government schemes | 2.3 |
| Employment schemes | 2.2 |
| Communication | 1.7 |
| Credit and input subsidies | 0.2 |

**Appendix Table A2: ML Logit Estimates:
Probability Respondent Voted in the Last *Lok Sabha* Election,
Rural Male Household Heads Aged 21 - 50, by Survey and Year**

| Variable/Survey Source | 1985 Indian National Election Study | 2005 India Human Development Survey |
|----------------------------|--|--|
| Secondary school and above | -1.12*** (0.286) | -0.135*** (0.0669) |
| Age group 36-50 | 0.164 (0.367) | 0.462 (0.0680) |
| Percent voting | 94.6 | 85.5 |
| N | 358 | 8,990 |

Standard errors in parentheses clustered at the state level. ***Significant at the .001 level.

Appendix Table A3: Fraction of Households Receiving Public Goods Each Term

| Type of election: | Open | SC | ST | OBC |
|-------------------|----------------|----------------|----------------|----------------|
| | (1) | (2) | (3) | (4) |
| Water | 0.69 (0.40) | 0.73 (0.39) | 0.78 (0.71) | 0.72 (0.39) |
| Sanitation | 0.42 (0.46) | 0.42 (0.46) | 0.55 (0.47) | 0.42 (0.46) |
| Roads | 0.69 (0.41) | 0.72 (0.40) | 0.74 (0.41) | 0.73 (0.39) |
| Telephones | 0.07 (0.24) | 0.12 (0.30) | 0.08 (0.25) | 0.10 (0.28) |
| Electricity | 0.14 (0.34) | 0.20 (0.38) | 0.17 (0.36) | 0.20 (0.38) |
| Street lighting | 0.16 (0.36) | 0.19 (0.38) | 0.19 (0.39) | 0.22 (0.40) |

Means and standard deviations (in parentheses).

SC=scheduled caste, ST=scheduled tribe, OBC=other backward caste.

Statistics are based on the last three terms in each ward.