Ethnic Diversity and the Under-Supply of Local Public Goods *

Kaivan Munshi[†] Mark Rosenzweig[‡]

September 2018

Abstract

We extend the citizen-candidate model by allowing for cooperation within, but not across, ethnic groups to examine the relationship between ethnic diversity, the distribution of welfare transfers, and the supply of public goods in representative democracies. We test the hypothesis that cooperation is restricted to the representative's ethnic group (which is the caste or *jati* in India) and, hence, that the supply of public goods is increasing in its size, using newly available data over multiple election terms at the most local (ward) level. We find support for both the hypothesis of within-group cooperation as well as for competition between groups for targetable welfare transfers. Counterfactual simulations using structural estimates of the model quantify the under-supply of public goods due to group-specific cooperation, indicating that ethnic diversity (which reduces group-size on average) significantly reduces the supply of local public goods in India. Additional counter-factual simulations examine the impact of policies, including caste-based reservation, that would be expected (based on the model) to change the supply of public goods by altering the size of the groups that come to power.

*We are grateful to many seminar participants for their constructive comments. 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 and the National Institutes of Health through grant R01-HD046940. We are responsible for any errors that may remain.

[†]University of Cambridge

[‡]Yale University

1 Introduction

The under-provision of local public goods is a common and persistent feature of developing economies. Many explanations have been proposed for this phenomenon. We focus in this paper on one of them – the well documented negative relationship between ethnic diversity and public good provision. Although this relationship is by now an established empirical fact, the specific mechanisms that underlie this relationship are less well understood and the empirical evidence less strong. Within the economics literature, two mechanisms have been proposed: First, ethnic diversity is associated with greater heterogeneity in preferences for different types of public goods. Voters' preferences are thus further away, on average, from chosen policies, resulting in a reduced *demand* for public goods (Alesina, Baqir, and Easterly 1999). Second, ethnic diversity is associated with smaller groups on average. If the supply of a non-excludable public good is increasing in group size because there are more (cooperating) individuals available to contribute, then ethnic diversity will be associated with a reduced *supply* of public goods (Miguel and Gugerty 2005).

Miguel and Gugerty's pioneering analysis of how ethnic diversity affects the supply of public goods makes the assumption, for which there is extensive anthropological evidence, that individuals are able to cooperate within but not across groups. Their analysis, consistent with the anthropological literature, is concerned with a collective action problem in which many individuals must contribute towards the supply of a public good. However, public goods, even in developing economies, are largely provided by governments. Over the past decades there has been a major policy shift towards the decentralized allocation of public goods using local democratic processes (World Bank 2004). A key premise of decentralization is that adequately-compensated political representatives will be accountable to local citizens, with the democratic process allowing the electorate to vote out under-performing representatives (Seabright 1996). However, standard political incentives are inadequate in many developing countries (see, for example, Ferraz and Finan 2010, 2011). A key question is whether in the absence of these incentives, the same ethnic ties that are used to solve the collective action problem could be used to elicit effort by elected political representatives.[1]

In this paper, we extend the citizen-candidate models of Osborne and Slivinsky (1996) and Besley and Coate (1997) to allow (as in Miguel and Gugerty) for cooperation within, but not across, ethnic groups. Our model delivers predictions for the relationship between ethnic diversity, the distribution of welfare transfers, and the supply of public goods in representative democracies. We test the hypothesis that cooperation is restricted to the representative's ethnic group (which is the caste or *jati* in India) and, hence, that the supply of public goods is increasing in its size, using newly available data over multiple election terms at the most local (ward) level. We find support for both the hypothesis of within-group cooperation as well as for competition between groups for targetable welfare transfers. Counterfactual simulations using structural estimates of the model quantify the under-supply of public

[1]The anthropological literature describes how tight-knit peasant communities cooperate to solve the collective action problem that is associated with the tragedy of the commons; e.g. Scott (1976), Hayami and Kikuchi (1982), Wade (1988), Ostrom (1990). However, with the exception of Tsai (2007), who documents cooperation within lineage groups in Chinese local governments, we are unaware of previous research that makes the logical extension to representative democracies.

goods due to group-specific cooperation, which is substantial, indicating that ethnic diversity (by reducing group size) significantly reduces the supply of local public goods in India. Additional counterfactual simulations examine the impact of policies, including caste-based reservation, that would be expected (based on the model) to change the supply of public goods by altering the size of the groups that come to power.

Our model, as noted, incorporates and integrates many features of existing models of public goods provision. Although Miguel and Gugerty's model and our model describe different institutional environments, they share a number of implications. In Miguel and Gugerty, many individuals must make a fixed per capita contribution, whereas in our model the political representative chooses the level of effort, which translates into the level of public goods. Nevertheless, both models predict that larger groups will supply more public goods. Free-riding by larger groups on smaller groups, to avoid bearing the cost of supplying the public good, is a feature of both models, and both imply that there will be a discrete increase in public good provision when the population share of the largest group crosses a threshold level (at which it stops free-riding for sure). A distinguishing element of our model is that the representative is responsible for the supply of public goods as well as the distribution of welfare transfers. The latter responsibility can make large groups relatively unpopular with the electorate, which is another reason why the elected representative may not always be drawn from the largest group. Besley et al. (2004) also incorporate the dual tasks of supplying public goods and distributing welfare transfers in their analysis. However, they do not endogenize the competence of the elected political representative and the associated supply of public goods, nor do they link the two tasks by showing how the additional responsibility of distributing welfare transfers affects the supply of public goods by changing who is elected.

We are able to test the central hypothesis of the supply-side mechanism, which is that the supply of public goods is increasing in group size, using data on Indian local governments that we have collected. These data are unique in their scope and detail. The 2006 REDS data cover a large sample of wards – the most local level of government – across the major Indian states over three election terms. The natural ethnic group around which cooperative political arrangements are organized in India is the caste or *jati*. Social connections within the caste have been used for centuries to facilitate private economic activity such as mutual insurance; e.g. Mazzocco and Saini (2012), Munshi and Rosenzweig (2016). The issue is whether the same connections can be used to support cooperation between political representatives and their castes in an environment where the standard political incentive mechanisms are less relevant. We have collected information on the caste and education of the elected representative, expenditures on both the maintenance and new investments in six major non-excludable public goods at the street level (which can be mapped to the ward level), and the receipt of welfare transfers by specific households in each election term. In addition, the data provide the caste of every household in each ward.

We first provide evidence that ward representatives target welfare transfers to individual households in their own caste based on our household panel data: if a household belongs to the caste of the elected representative it is significantly more likely to receive a transfer compared to when that same household's ward representative belongs to another caste. However, there is no evidence that any of the six public goods are targeted at the street level, consistent with the assumption that they are non-excludable within wards. Our additional evidence on group-specific cooperation and the resulting group-size effect on public goods supply is also consistent with the assumption that the caste or *jati* is the relevant social unit within which cooperative arrangements are organized in India. Although we find that there is a positive and significant relationship between the size of the elected representative's caste in the ward and both the representative's education, which we use to measure his competence, and the supply of the six public goods, the number of ward residents that belong to the representative's broader caste grouping; i.e. Scheduled Caste, Scheduled Tribe, or Other Backward Caste, has no bearing on the supply of public goods. And belonging only to a representative's caste grouping but not to his caste has no effect on a household's likelihood of receiving a welfare transfer.

Identification of the supply-side channel linking ethnic diversity to public good provision poses two challenges. First, there must be exogenous variation in the size of the representative's group, which our model predicts will determine the supply of public goods. Second, the supply-side channel must be distinguished from the demand-side channel linking heterogeneity in preferences in more diverse populations to the under-provision of public goods. Our research design addresses both challenges by taking advantage of (i) the panel data on elections, which allow us to subsume the demand-side channel in a ward fixed effect, and (ii) exogenously changing set asides based on broad caste groupings in Indian local governments, which alter the size of the largest *eligible* caste across terms in the same ward. We find, consistent with group-specific cooperation as implied by the model, that there is a significant jump-up in the supply of local public goods and the education of the elected representative when the population share of the largest eligible caste crosses a threshold.

Our research design improves on the empirical analysis in Miguel and Gugerty (2005) in a number of ways. Miguel and Gugerty estimate a negative cross-sectional relationship between public good provision and ethnic diversity, measured either by fractionalization or the population share of the largest group. However, their claim that omitted variable bias can be avoided because current ethnic diversity is strongly determined by historical ethnic diversity, which, in turn, was largely accidentally determined, has obvious limitations. In particular, historical diversity could have been determined by, or could be correlated with, factors that continue to have an effect on public good provision today (just as historical ethnic diversity does).[2] With just 84 observations, Miguel and Gugerty also cannot test the prediction of their model, and our model, which is that there will be a discontinuous increase in public good provision when the population share of the largest group crosses a threshold level. In addition, their analysis does not distinguish between the supply-side and the demand-side channels linking ethnic diversity to public good provision.[3] Although evidence from experimental games has

[2]Miguel and Gugerty also instrument for local ethnic diversity with regional diversity, and include a rich set of statistical controls in their estimating equation. However, the drawbacks of these strategies are well known and well understood.

[3]To provide indirect support for the supply-side channel, based on group-specific cooperation, Miguel and Gugerty estimate the relationship between the threat of social sanctions and ethnic diversity. However, this analysis has theoretical and empirical shortcomings. Social sanctions are an off-equilibrium phenomenon (as the authors acknowledge) and, on the few occasions that they are observed, they should not be predictable. Moreover, these sanctions are applied within the ethnic group, whereas Miguel and Gugerty's data are at the school level.

provided prior support for the supply-side channel (Habyarimana et al. 2007), our analysis provides the first credible evidence of group-specific cooperation based on actual election data, and the resulting negative relationship between ethnic diversity and the supply of public goods, in the context of local governments.

The root cause of the under-provision of public goods through the supply-side channel is that cooperation is group-specific. This effect is exacerbated in ethnically diverse populations because groups are small. To quantify the under-supply of public goods in Indian local governments due to the fact that cooperation does not extend beyond caste boundaries, we estimate the structural parameters of the model and conduct counter-factual simulations. In this analysis, the supply of public goods is measured by the fraction of the six major local public goods falling within the *panchayat's* jurisdiction for which there were expenditures on either new construction or maintenance in the ward in a given election term. In line with the observation that rural Indian households are under-served, the average fraction of the ward population receiving a given public good, across the six public goods we consider, is less than one-third in 40% of wards. The first counter-factual result is that if the ward representative internalized the benefit derived from the public goods by all residents, then the entire population would receive all public goods in all wards.

Ethnic diversity lowers public good provision through the supply-side channel by reducing the size of the representative's group. Ethnic diversity in the constituency is not subject to policy manipulation. For a given level of diversity, particular policies could still affect the supply of public goods by changing the size of the group in power. One example of such a policy is the increasingly common practice of making local representatives responsible for the administration of welfare programs in addition to their traditional role of delivering public infrastructure. The location of the estimated threshold at which there is a discrete increase in the supply of public goods and the competence of the elected representative implies, from the model, that this policy makes representatives from larger groups unpopular and thus reduces the size of the representative's group, on average, with an accompanying decline in the supply of public goods. Our counter-factual simulations of the estimated structural model quantify this effect, which turns out to be modest.

We also assess the consequences for the supply of public goods of another important policy - political reservation for ethnic groups, which we exploited in the Indian context, to test our model. Reservation policies are in part based on the implicit assumption that the representatives of groups favor co-ethnics when they are elected. Although our results imply that reservation for disadvantaged minority castes in India will indeed channel targetable public resources in their direction, they also suggest that reservation can have adverse efficiency consequences. Reservation reduces on average the size of the winning candidate's group in equilibrium and, therefore, the supply of non-excludable public goods by restricting the set of ethnic groups that are eligible to stand for election. Given the actual distribution of castes across wards in our data, our counter-factual simulations indicate that the effect of this restriction on public goods supply could be substantial in rural India; a policy that combines the decoupling of welfare transfers and public goods provision with de-reservation would reduce, from 40% to 20%, the fraction of wards in which less than one-third of the population received each public good, averaged across the six public goods we consider. This result is still, however, some way from the first best, due to the inherent limitations of group-specific cooperation.

2 Institutional Setting

2.1 Local Governments

The 73rd Amendment of the Indian Constitution, passed in 1992, 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, are divided into 10-15 wards. *Panchayats* are 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. We focus on these two independent tasks in this paper, as does the analysis in Besley et al. (2004). Our analysis diverges from previous research on Indian local governments, however, by examining the supply of public goods at the most local – ward – level. This allows us to focus on the "last mile," which is believed to be critical to the delivery of public services in developing countries (World Bank 2016). Although 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, the ward's own representative clearly plays a critical role in determining the resources that it receives.[4] Our analysis thus focuses on the representative's competence and effort, but the research design will take account of the role played by other ward representatives and the *panchayat* president in determining the supply of public goods in the ward.

The data that we use to examine the supply of public resources at the most local level 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, which covers 242 of the original 259 villages in 17 major states of India. We make use of three components of the survey data in this paper - 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.[5] 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 whether the household received a Below the Poverty Line (BPL) card in each of the last three election terms.

[4]Key informants in the 2006 Rural Economic and Development Survey (REDS), which we use for much of the analysis in this paper, were asked who in the *panchayat* decided the allocation of expenditures. Although 81% of informants reported that the president had a say, 93% said that is was, nevertheless, a joint decision of all *panchayat* members. In contrast, just 5% of respondents said the allocation decisions were determined by an influential caste group in the village.

[5]The states are Andhra Pradesh, Bihar, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Rajastan, 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.

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 and employment schemes account for 15% of total expenditures.[6] As in Besley et al. (2004), we measure welfare transfers by access to BPL cards. BPL card holders are eligible for subsidized food through a public distribution scheme. In addition, most central and state welfare programs administered by the *panchayat* restrict eligibility to BPL households. These programs provide funds for housing construction and repair and private electricity and water supply.

The village inventory obtained information on whether new construction or maintenance of specific public goods *actually* took place on each street in the village in each election 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 which fall under the purview of the *panchayat* and 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.[7] As reported in Appendix Table C1, these six goods account for 67.5 percent of the *panchayat's* discretionary spending. Nevertheless, and despite decades of rapid economic growth in India, 43% of households do not have electric connections, 74% lack running water, 59% live on streets without functional lights, 56% do not own a toilet or live on a street with a public toilet, and 47% live on a street that is unpaved, highlighting the magnitude of the problem that we are studying.

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, in principle, 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 et al. 2004). In practice, local constituencies with a higher fraction of disadvantaged minorities appear to be reserved earlier (Dunning and Nilekani 2013). The research design allows for this departure from randomization, as discussed below.

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 re-elected.[8] The representatives with the greatest chance

[6] Although *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.

[7] Public 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. *Panchayats* play a marginal role in the delivery of health and education services, which are largely administered by the state government (Bardhan and Mookherjee 2006).

[8] Chattopadhyay and Duflo (2004) note that not a single woman in their sample of reserved constituencies in the state

for re-election are men elected in unreserved seats. However, the probability that a ward election is unreserved in a given election term is just 0.4.[9] Given that the type of reservation is independent over time within a constituency, and assuming that leaders in reserved seats are never re-elected in the subsequent election, the maximum fraction of incumbent representatives that can 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. Exogenous turnover, generated by rotating reservation or by term limits, is observed in many local governments. We will incorporate this feature of local politics in the theoretical model and exploit the exogenous variation in the pool of eligible candidates that it generates in the empirical analysis.

We incorporate two additional features of local politics, which are also relevant in India, in the model. First, local representatives are poorly compensated for their efforts throughout the world. Ferraz and Finan (2011) note that 98% of municipal legislators in Brazil hold a second job. In the Indian local governments, *panchayat* presidents are paid 50-60 dollars per month (less than the minimum wage) while ward representatives earn even less. Second, political parties tend to be less active in local elections in India. Many states prohibit candidates in local elections from contesting on party lines. Symbols of recognized parties are not allocated to candidates in village *panchayats* and there is no separate nomination form – as in higher-level *Panchayat Samiti* and *Zilla Parishad* elections – for recognized political parties.[10]

2.2 Caste in Indian Local Politics

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. This has been the motivation for one of the world's most ambitious affirmative action programs, which reserves positions in the central government, institutions of higher education, and local governments (as described above) for historical disadvantaged caste groups; the Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Castes (OBC). There is, however, another side to this system, associated with solidarity and social connectedness *within* sub-castes or *jatis*, that is relevant for the current analysis. For expositional convenience, we will refer to the *jatis* as *castes* in this paper and the broader caste groups as *caste groupings*. There are approximately 4,000 castes in India and thus hundreds of castes within a caste grouping.

The basic marriage rule in Hindu society is that individuals must match within their caste. Muslims also follow this rule, matching within *biradaris*, which are equivalent to *jatis*, while converts to Christianity continue to marry within their original *jatis*. Evidence from nationally representative surveys such as the 1999 Rural Economic Development Survey (REDS) and the 2005 India Human Development Survey (IHDS) indicates that 95% of Indians marry within their caste or its non-Hindu

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.

[9]In our sample of ward-terms, 60 percent were open to all castes (see Table 6 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.

[10]We are grateful to Clement Imbert for bringing these features of Indian local politics to our attention.

equivalent kinship community and recent genetic evidence indicates that these marriage patterns have been in place for over 2,000 years (Moorjani et al. 2013). Marriage ties built over many generations result in a high degree of social connectedness within castes, which has been used, and continues to be used, to support economic cooperation. For example, mutual insurance arrangements have long been organized around the caste in rural India (Caldwell, Reddy, and Caldwell 1986, Mazzocco and Saini 2012, Munshi and Rosenzweig 2016). When urban jobs became available in the nineteenth century, with colonization and industrialization, 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, including business, in the contemporary economy (Damodaran 2008, Munshi 2011).

Previous research has documented economic cooperation within but not between castes in the village; e.g. Mazzocco and Saini (2012), Munshi and Rosenzweig (2016). We would expect political cooperation to be similarly organized within but not between castes. The 2006 REDS includes a module in which 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, an overwhelming majority of elected ward representatives received support from their caste (inside and outside the village).

Caste-based campaign support extends to voting behavior. If citizens vote on caste lines, then there should be a discrete increase in the probability that the largest caste's representative is elected when it just accounts for a majority of the ward's population (and can be assured of electoral success). The 2006 REDS village census obtained information on all households in each of the sampled villages, by ward. This enables us to measure the caste size-distribution in each ward.[11] Rotating reservation for historically disadvantaged groups changes the set of castes that are eligible to stand for election in each ward over time. Figure 1 thus plots the relationship between the population share of the largest *eligible* caste and the probability that the elected representative is drawn from that caste.[12] There is a discrete increase in this probability, from 0.4 to 0.7, just around the point where the caste attains an absolute majority, which is indicative of block voting.

Table 2 reports estimated coefficients in an equation corresponding to Figure 1. The dependent variable indicates whether the elected ward representative is drawn from the largest eligible caste in a given election term. The benchmark specification in Column 1 includes a binary variable indicating whether the largest eligible caste in the election term has an absolute majority in the ward as the only regressor. The augmented specification in Column 2 includes the size of the caste in the ward and a vector of election characteristics. The analysis covers three election terms and so it is necessary

[11] A caste is any set of households within a village reporting the same caste or *jati* 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. 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.

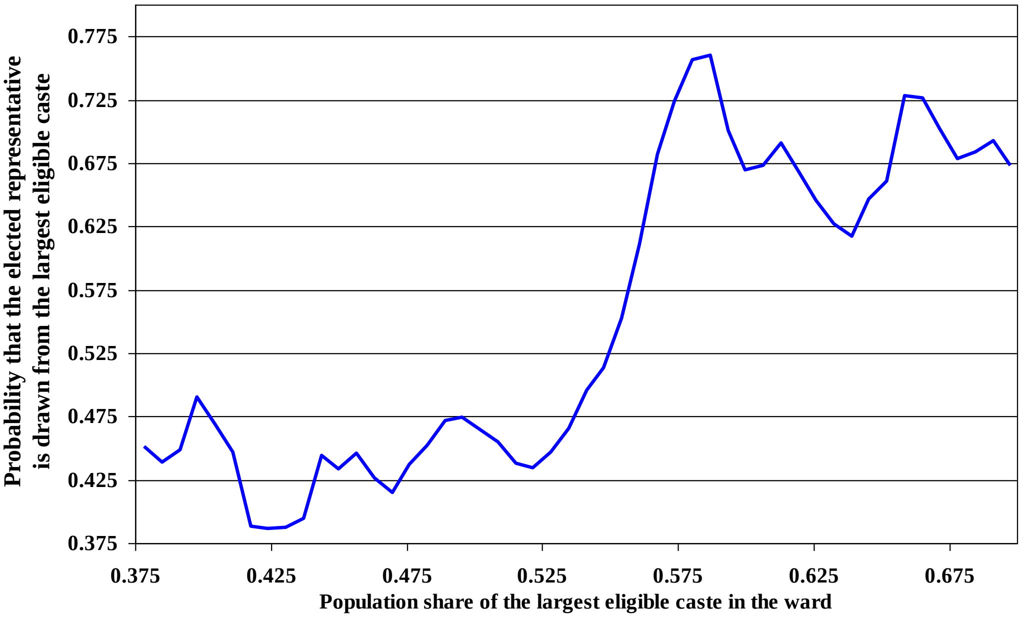
[12] The sample is restricted to wards with more than one caste and more than one street in all the analyses in this paper. Outlying ward-terms in the top 1% of the *panchayat*-level public goods expenditure distribution are also excluded. In addition, all analyses that require information on the representative's caste affiliation are based on 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). Analyses that do not require this information use the full sample.

Table 1: **Sources of Campaign Support (Percent) for Ward Representatives**

Source / domain	Within village (1)	Outside village (2)
From caste	82	29
From religion	28	13
From wealthy individuals	38	–
From a political party	–	41

Source: 2006 Rural Economic Development Survey (REDS) Village Inventory. The statistics are computed over the last three election-terms in each ward. Each statistic reflects the percentage of representatives who received financial and organizational support from a given source.

Figure 1: **Evidence on Caste-based Voting**



to account for changes in the total resources made available to Indian local governments as well as changes in reservation status at the ward-level over time. *Panchayat* elections are not synchronized across states and so the augmented specification includes election-term fixed effects, the election year, and reservation (SC, ST, OBC) fixed effects. The probability that the representative is drawn from the largest eligible caste increases by 0.4 when it has an absolute majority, matching the discrete increase documented in Figure 1. Notice that the size of the largest eligible caste has no effect on its probability of coming to power, despite the fact that this variable will later be seen to have a positive and significant effect on the supply of public goods. Our model provides an explanation for this finding.

Table 2: **Electoral Outcome for the Largest Eligible Caste in the Ward**

Dependent variable:	elected representative is drawn from the largest eligible caste (1)	(2)
Whether that caste has a majority in the ward	0.429** (0.0427)	0.402** (0.0458)
Log size of that caste in the ward	–	0.000195 (0.000465)
Reservation fixed effects	No	Yes
Election year	No	Yes
Election-term fixed effects	No	Yes
Sample mean of dependent variable	0.708	0.708
N	747	747

Standard errors clustered at the ward level in parentheses. ** p<0.05

3 A Model of Ethnic Groups in a Representative Democracy

3.1 Ethnic Groups and Public Resources

The residents of a political constituency receive two types of public resources: a public good and welfare transfers. In the Indian context, the welfare transfers would be BPL cards, which entitle beneficiary households to publicly funded private goods. Public goods include drinking water, sanitation, improved roads, electricity, etc. Public resources are delivered by a representative who is a resident of the constituency and is elected by the constituency. Given the emphasis on the supply-side channel linking ethnic diversity to public good provision, we assume that individuals are heterogeneous in their ability as representatives but are homogeneous in their preferences for public resources. The latter assumption will be relaxed in the empirical analysis. In addition, we make the following assumption about the level and distribution of the public resources, which is validated empirically:

A1. (a) The total amount of the welfare transfers is exogenously determined, whereas the supply of the public good depends on the effort exerted by the representative. (b) The welfare transfers are targetable, whereas the public good is non-excludable.

For this local government to function effectively, the representative who is elected must have the ability and the incentive to exert the optimal level of effort. However, standard political incentive mechanisms are typically inadequate in developing countries. Local political representatives are poorly compensated and term limits restrict their ability to credibly commit to exert effort once elected. Political parties, which have a long-term reputation to maintain, could potentially reduce these individual incentive problems, but they are less active at the local level in many countries, including India. In practice, representatives could receive private benefits from public office through corruption or political career advancement. The possibility of re-election could be relevant even with rotating reservation if the individual is sufficiently patient or the system can be manipulated. And political parties could find ways to circumvent legal restrictions and involve themselves in local elections. For analytical convenience, we capture the idea that the standard incentive mechanisms will nevertheless be incomplete by making the following stronger assumption:

A2. Parties are not involved in local politics. Representatives receive no monetary compensation and are elected for a single term.

When the standard political incentive mechanisms are unavailable, cooperation within ethnic groups can be used to increase the supply of public goods. 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. As in Miguel and Gugerty (2005), we assume that cooperation is possible within, but not across, ethnic groups. Given assumption A2, a long-term reciprocal arrangement between the representative and the entire constituency is infeasible. However, the representative's ethnic group could credibly commit to compensate him *ex post*, even if he is elected for a single term, because they are connected to each other in many ways outside the political system. The representative will target welfare transfers to co-ethnics in his constituency. In parallel, side-transfers will flow from the representative's ethnic group to him, resulting in the selection of a candidate whose ability, and the effort he subsequently exerts if elected to represent the constituency, are optimal from the group's perspective.

3.2 Representative Effort and Candidate Ability

N individuals drawn from K ethnic groups reside in the 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; $N_{k-1} < N_k, \forall k$. We first derive the effort, a, that is optimal from the ethnic group's perspective, taking as given the candidate's ability, ω. Because the level of welfare transfers is exogenously determined, from assumption A1, the representative's effort only affects the level of the public good. The effort is thus chosen to maximize

$$N_k a^\beta - \frac{a}{\omega}.$$

The first term in the expression above measures the utility derived from the non-excludable public good by the N_k members of ethnic group k if their candidate is elected. a^β is the level of the public good received in the constituency. While more effort will obviously increase the supply of the good, $\beta > 0$, 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 $\beta < 1/2$ below (which we later verify empirically). We assume, in addition, that the level of the public good maps linearly into the utility derived from its consumption by each resident of the constituency. We normalize so that this mapping is one-for-one. The second term in the expression above measures the effort cost of the ethnic group's chosen candidate, conditional on being elected. We make the usual assumption that the unit cost is decreasing in ability.

Based on the solution to the maximization problem, the optimal level of effort from the ethnic group's perspective is thus an increasing function of its candidate's ability and the size of the group,

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

The level of effort with group-specific cooperation 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) because outsiders are ignored.

The next step is to determine which individual will be selected as its candidate by the ethnic group. When making this decision, the group will take account of the relationship between the candidate's ability and the effort he will exert, conditional on being elected, derived above. It will also take account of the opportunity cost to the candidate of holding public office. Although the representative could extract personal rents for himself and advance his political career, we make the following assumption:

A3. The payoff in the private sector exceeds the payoff in public office, with the payoff gap increasing in ability.

Based on the preceding discussion, ethnic group k will put forward as its candidate the individual with ability ω that maximizes

$$N_k \left[a(\omega, N_k)\right]^{\beta} - \frac{a(\omega, N_k)}{\omega} - \alpha\omega. \tag{2}$$

The first term in the expression above measures the utility derived from the public good by the N_k members of the ethnic group. The second term measures the candidate's effort cost, conditional on being elected, and the third term the corresponding opportunity cost. The α parameter measures the difference in the returns to ability in the private sector and public office, which from assumption A3 is positive. Notice that the informal compensation received by the representative from his ethnic group does not appear in expression (2) because it is an internal transfer within the group. Substituting the expression for $a(\omega, N_k)$ from equation (1) and then maximizing with respect to ω, it is evident that larger groups will choose candidates with higher ability from among their members if $\beta < 1/2$;[13]

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

Substituting this expression back in equation (1), the candidate's effort (conditional on being elected) is also increasing in the size of his group if $\beta < 1/2$;

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

It follows from equations (3) and (4) that,

Proposition 1. *If there is group-specific cooperation, and $\beta < 1/2$, then candidates representing larger ethnic groups in the constituency will have higher ability and will supply a higher level of the public good (conditional on being elected).*

Although the expressions in equations (3) and (4) are specific to our model, the result is more general. Because more individuals benefit from non-excludable public goods in a large group, it will

[13]This result is independent of the ability distribution among potential representatives; i.e. in the population, 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.

be in its collective interest to select a more competent candidate who will, in turn, exert greater effort. If N_k is sufficiently large, candidates could be positively selected on ability even if there is a substantial opportunity cost to standing for public office (α is large). Notice that what matters for the supply of public goods is the size of the group *in the constituency*, rather than the overall size of the group. This distinction will help us rule out alternative explanations for the group-size effect below.

We made two assumptions when deriving Proposition 1. The first assumption, A2, says that the standard political incentive mechanisms are inadequate. If this was not true, then intra-ethnic cooperation would be unnecessary and the size of the representative's ethnic group would have no bearing on the supply of public goods. The empirical test implied by Proposition 1 is thus a joint test of assumption A2 and the assumption that cooperation is restricted to the representative's ethnic group.

The second assumption, A3, is that the representative's payoff in the private sector exceeds his private payoff from holding public office (ignoring the informal compensation that he receives from his ethnic group). If this assumption is false, then $\alpha \leq 0$ and expression (2) would be monotonically increasing in ω. The most competent individuals would stand for election in all ethnic groups, regardless of their size. Indeed, it can be shown that this would be true even if there is no within-group cooperation (by setting N_k to one in expression (2)). If groups consist of a small number of individuals in general, then the most competent individual could be more competent in larger groups (even if the ability distribution is independent of group size) just by chance. This would generate a positive correlation between group size and the representative's competence, even if within-group cooperation were absent.

A more stringent test of assumption A3 and within-group cooperation is that if the distribution of ability among potential representatives is uncorrelated with, or negatively correlated with, the size of the ethnic group in the constituency, as documented empirically for Indian local governments below, then representatives of larger groups will be systematically drawn from higher in their group's ability distribution. This is a statement about relative competence rather than absolute competence. If this test fails and assumption A3 is rejected, then this would rule out one channel through which within-group cooperation will increase the supply of public goods in larger groups; by having more competent candidates put forward for election. However, effort and, hence, the supply of public goods, would still be higher in larger groups from equation (1).

3.3 Representative Selection

In Miguel and Gugerty's model, each individual must decide whether or not to contribute to the public good, generating a relationship between the ethnic size-distribution in the population and public good provision. The advantage of our research setting is that a single political representative is responsible for the supply of public goods. A positive relationship between the size of the elected representative's ethnic group in the constituency and his ability and effort thus provides direct evidence of group-specific cooperation and the accompanying group-size effect from Proposition 1. Which group's candidate gets elected, however, will depend on the ethnic size-distribution in the constituency.

We thus proceed to endogenize representative selection. This will allow us to develop a robust test of group-specific cooperation and to explain, in part, why the largest group's candidate is often not elected, despite the associated reduction in the supply of public goods.

Welfare transfers were ignored when deriving Proposition 1. This is because the level of the transfers is exogenously determined from assumption A1 and we assume, in addition, that the distribution of these transfers does not require any effort. Once we take a step back and model which group's candidate gets elected, however, the representative's task of distributing welfare transfers becomes relevant. In particular, we will see that the welfare transfers affect public good provision by changing which ethnic group comes to power. This is one reason why the largest group's candidate is not always elected. For analytical convenience, we begin with the special case where the representative's sole task is to supply the public good. We then turn to the complete model, where the representative is responsible for both the public good and welfare transfers.

Elections are contestable. Each ethnic group in the constituency chooses whether or not to put its preferred candidate up for election. 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. This electoral process is the same as the citizen-candidate models of Osborne and Slivinsky (1996) and Besley and Coate (1997), except that ethnic groups put up their preferred candidates.

In our model, someone is always elected because the net benefit from public good provision,

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

is strictly positive for all groups.[14] If all ethnic groups fielded their preferred candidates, then the largest group's candidate would always be elected from Proposition 1. 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 the public good supplied by the other group's representative is sufficiently large. This is essentially Olson's (1965) free-rider problem, except that it is shifted up to the group level.

The largest group will, nevertheless, prefer to have its own candidate be the representative if the *incentive condition* is satisfied;

$$N_K\left[a(N_K)\right]^{\beta} - \frac{a(N_K)}{\omega(N_K)} - \alpha\omega(N_K) \geq N_K\left[a(N_{K-1})\right]^{\beta}. \tag{6}$$

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

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

[14]Equation (5) is derived by substituting the expressions for ability and effort from equations (3) and (4) in (2).

If this condition is satisfied, there is a unique equilibrium in which the largest group puts forward its preferred candidate for election and no other group fields a candidate.[15] If the incentive condition is not satisfied, there will be multiple equilibria. Replacing N_{K-1} with a smaller sized group, there will be a group $\underline{k}$ for which inequality (7) is just satisfied. Any strategy profile in which a group of size $N_k \in (N_{\underline{k}}, N_K]$ fields its preferred candidate, while all other groups stay out, will be an equilibrium.[16]

When will the incentive condition be satisfied? This will depend on the ethnic size-distribution of the constituency. In general, many measures of the ethnic size-distribution are available. We choose to measure the size-distribution by the population share of the largest ethnic group, as do Miguel and Gugerty (2005), because a relationship between this measure and the supply of public goods can be analytically derived from our model. 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 $\beta \in (0, 1/2)$ by assumption. If the largest group accounts for almost the entire population, $N_K \to N$ and $N_{K-1} \to 0$, the term in square brackets will reverse sign. Holding constant the population of the constituency, N, the average size of all other groups must decline as N_K increases. 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. 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. It follows that there is a discrete increase in the probability that the elected representative is drawn from the largest ethnic group when its population share reaches that threshold. The higher is the threshold, the greater is the under-supply of the public good due to the free-rider problem.

We next establish that the preceding result continues to be obtained with the complete model in which the elected representative is responsible for the supply of the public good and the distribution of welfare transfers. Assumption A1 states that welfare transfers are targetable and that their level is exogenously determined. We now make the stronger assumption that each constituency receives a fixed T units of the welfare transfers. Each beneficiary receives one unit of the transfer, which maps into its utility equivalent θ. Although welfare transfers are intended for economically or socially disadvantaged households in practice, we assume for analytical convenience that all households are eligible for the transfers in the model.[17] The representative will first ensure that each member of

[15](i) The strategy profile in which no one contests, and the public good is not provided, is not an equilibrium. Any group would be better off by deviating and fielding its preferred candidate, who would generate a positive net benefit for the group – from expression (5) – once elected. (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 candidate elected, 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 election, while having to bear the cost of entry.

[16](i) Because inequality (7) is *not* satisfied for group $k > \underline{k}$, 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.

[17]We could incorporate the eligibility requirement by assuming that a fixed fraction of the households in each ethnic group are eligible for the welfare transfers. The representative first targets the transfers to the eligible members of his own group, randomly assigning the transfers that remain (if any) to eligible outsiders. Alternatively, we could assume that the representative first channels the transfers to eligible households and a fixed fraction of ineligible households in

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 completely 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) candidate and that candidate will continue to exert the same level of effort if elected. While the representative of a larger group thus continues to supply a higher level of the public good, outsiders are worse off with respect to the welfare transfers when a larger group is in power. This changes which group's candidate gets elected, but does not change the supply of public goods conditional on who is elected.[18]

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 in Appendix A, 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 preferred candidate elected, an additional *feasibility condition*, derived formally in the Appendix, must also be satisfied to ensure that its candidate 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 the non-excludable public good than any other group's representative (from Proposition 1). 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. When deriving the electoral outcome for the complete model, which includes both the public good and welfare transfers, there are two regimes to consider: (i) $S^{**} > 0.5$ and (ii) $S^* < 0.5$.[19]

In the first regime, the feasibility condition is irrelevant because the largest ethnic group can win the election at S^{**}, which is greater than 0.5, without outside support. The probability that the representative is drawn from the largest group thus increases discontinuously at S^{**}, which is less than S^*, when the representative is responsible for both the welfare transfers and the public good. In the second regime, the threshold will also be at S^{**}, which is now less than 0.5, if the feasibility condition is satisfied at that population share. If it is not, the threshold will be located at the lowest population share at which the feasibility condition is satisfied or 0.5, whichever is smaller. Based on the structure of the model and placing the more stringent restriction that $\beta < 1/4$, we derive a stronger result in the Appendix, which is that if the feasibility condition is not satisfied at S^{**}, then it

his group, distributing what remains to eligible outsiders. Either way, this would add an additional parameter to the model, without changing any of the results that follow.

[18] An 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.

[19] We ignore the special (and unlikely) case, $S^{**} < 0.5 < S^*$.

will not be satisfied at any population share greater than S^{**}, and the threshold will thus necessarily be located at 0.5, which is greater than S^{*}.

Proposition 2. *(a) In a sample of constituencies ordered by the population share of the largest ethnic group, group-specific cooperation implies that there will be a discrete increase in both the elected representative's ability and the supply of public goods when the population share reaches a threshold. (b) If the location of the threshold is equal (not equal) to 0.5, then this indicates that adding welfare transfers to the representative's list of responsibilities decreases (increases) the supply of public goods.*

The model predicts that there will be a discrete increase in the probability that the largest group's representative is elected when its population share crosses a threshold level. Given the accompanying increase in the size of the elected representative's group, it follows from Proposition 1 that there will be a discrete increase in the representative's competence and the supply of public goods at the same threshold. This prediction is summarized in Proposition 2(a), providing us with a robust test of group-specific cooperation and the group size effect.

Both our model and Miguel and Gugerty's model predict that the largest group will stop free-riding on smaller groups when it accounts for a sufficient share of the population, resulting in a discontinuous increase in the supply of the public good at a threshold. Our model has a further motivation for a threshold, at the point where it just has an absolute majority, if the crowding out of outsiders with respect to the welfare transfers dominates the benefit from the extra public goods it supplies. The location of the threshold in our model is thus informative about preferred policy from Proposition 2(b). If the threshold is located precisely at 0.5, it follows that the current policy of making local representatives responsible for the administration of welfare programs will have reduced the supply of public goods by reducing the size of the group in power on average. If the threshold is located anywhere else, then this implies that the responsibility of distributing welfare transfers has reduced the free-rider problem and increased the supply of the public good.

The negative relationship between ethnic diversity and both the competence of the elected representative and public good provision also allows us to distinguish our model from alternative models in which the supply of public goods at the local level is determined from above by political parties, which predict a positive relationship. If political parties decide the allocation of resources across constituencies, then Lindbeck and Weibull's (1987) model of redistributive politics implies that swing constituencies will be favored. To the extent that ethnic groups align with particular parties, greater resources will be allocated to more competitive, ethnically diverse, constituencies (Casey 2015). Parties will also assign more competent candidates to those constituencies (Banerjee and Pande 2009). This is exactly the opposite of what we observe, indicating that the supply of public goods in Indian local governments (at least at the ward level) is driven from below rather than from above.

4 Evidence on Group-Specific Cooperation and Targeting

4.1 Targeting of Public Resources

Assumption A1 states that welfare transfers are targetable, whereas public goods are non-excludable. In this section, we use our data to validate this assumption at the ward level. To identify targeting of welfare transfers on caste lines, we examine the receipt of Below the Poverty Line (BPL) cards by households in our sample villages. BPL cards are meant to be received by economically disadvantaged households, but it is well known and well documented; for example, by Besley, Pande, and Rao (2011) that ineligible households who are politically connected can also benefit from them. To identify targeting, we estimate the following equation using the REDS household survey data:

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

where the dependent variable indicates whether household i receives a BPL card in ward j in election 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 fixed effects, election-term fixed effects, and the election year. ξ_{ijt} is a mean-zero disturbance term. This specification closely matches the specification used to estimate the probability that the elected representative is drawn from the largest eligible caste in Table 2.

The conditional (fixed effects) logit model is used to estimate equation (8) because the mean of the dependent variable is far from 0.5. The coefficient on the representative's caste in Table 3, 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 card 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 on caste lines. Assumption A1 also states that the total amount of welfare transfers is exogenously determined. This implies that the probability of receiving a BPL card in the population as a whole will not depend on the size of the representative's caste, regardless of targeting, as observed in Column 1.

Table 3: **Targeting of Public Resources**

Dependent variable	Household receives BPL card			Public goods placed on household's street		
	(1)	(2)	(3)	(4)	(5)	(6)
Household belongs to the representative's caste	1.377**	1.533**	0.213	–	–	–
	(0.577)	(0.639)	(0.752)			
Fraction of households belonging to representative's caste on street	–	–	–	-0.00915	-0.0167	-0.0451
				(0.0573)	(0.0596)	(0.106)
Size of representative's caste	-0.0192	-0.0217	-0.0308*	0.000463	0.000404	-0.000268
	(0.0190)	(0.0208)	(0.0158)	(0.000957)	(0.000900)	(0.00152)
Interaction[a)]	–	–	0.053***	–	–	0.00151
			(0.0201)			(0.00264)
Household belongs to the representative's caste grouping	–	-0.531	–	–	–	–
		(0.795)				
Fraction of households belonging to representative's caste grouping on the street	–	–	–	–	-0.0488	–
					(0.0510)	
Sample mean of dependent variable	0.258	0.258	0.258	0.343	0.343	0.343
N	1387	1387	1387	1387	1387	1387

All specifications include household fixed effects, reservation fixed effects, election-term fixed effects, and election year.
Public goods are measured by the fraction of the six major public goods received on the household's street in each election-term.
a) In column 3, the interaction variable is the representative's caste size × household belongs to the representative's caste. In column 6, the interaction variable is representative's caste size × the fraction of households in the representative's caste on the household's street. Caste grouping is SC, ST, OBC, all other.
Standard errors clustered at the ward level in parentheses. *p<0.1, ** p<0.05, *** p<0.01. Columns 1-3 are estimated using conditional logit.

The assumption in our analysis is that cooperation can be supported within, but not between, castes or *jatis*. This is in line with the political science literature, which has long assumed that clientelist arrangements in India are organized on *jati* lines. Clientelism is characterized by the transfer of targeted public goods, jobs, or services to groups of voters in return for their political support (Stokes 2015). The same social ties that allow transfers to flow from the *jati* to its representative in our model, would allow the *jati* to credibly commit to honoring its electoral obligations in a long-term clientelist arrangement with a political party. Although political parties did form patron-client relationships with many castes, they were traditionally not associated with a particular caste identity (Yadav 1999). There has, however, been a change in recent decades with the emergence of state-level parties that are explicitly identified with caste-groupings, such as upper castes, OBC, or SC (Yadav 1999). Although it is believed that the *jati* continues to be the social unit around which political activity is organized in the village (Brass 1990, Yadav 1999), we allow for the possibility that castes form coalitions in the ward, as they sometimes do at the state level, by adding a variable in Table 3, Column 2, which indicates whether the representative belongs to the household's caste grouping (SC, ST, OBC, all other). The coefficient on the caste-grouping indicator variable is small with the wrong (negative) sign, and statistically insignificant, while the coefficient on the caste indicator variable retains its magnitude and significance. Previous research on caste-based targeting in Indian local governments has documented that SC/ST households are more likely to receive publicly provided private goods when the *panchayat* president's position is reserved for SC/ST's; e.g. Besley et al. (2004), Bardhan, Mookherjee, and Torrado (2010). While it may make sense from a policy perspective to assess whether political reservation for particular caste groups benefits the members of those caste groups on average, our results indicate that targeting (at least at the ward level) is occurring at a finer *jati* level.[20]

We assume in the model that the total amount of welfare transfers is fixed and that the representative favors his own group when distributing the transfers. A larger group will capture more transfers simply because it has more numbers. This implies 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 relationship, which has not been previously examined in the literature, we allow the size of the representative's caste to affect insiders and outsiders separately by adding an interaction term to equation (8):

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

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. Providing evidence that outsiders are increasingly crowded out when a larger group is in power, this coefficient in Table 3, Column 3 is negative and significant (at the 10% 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 a BPL card by 6 percentage points (25%). At the mean size of the representative's caste (37 households), being an insider increases the probability of receiving the card by 4 percentage points.

[20]This may also explain why other studies; e.g. Dunning and Nilekani (2013), which have examined targeting at the coarser caste group level, have failed to uncover evidence of targeting.

Assumption A1 states that public goods are non-excludable at the ward level, in contrast with the welfare transfers which are assumed to be targetable. Although previous research at the *panchayat* level documents the targeting of public goods to the president's village (Besley, Pande, and Rao 2011), targeting has not been examined at the most local – ward – level. We thus replace access to a BPL card with public good provision as the dependent variable in equations (8) and (9). 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 public good provision at the street level is the fraction of the six major public goods – drinking water, sanitation, improved roads, electricity, street lights, and public telephones – for which there were expenditures on new construction or maintenance in a given election term. This measure is closely related to the public goods index constructed by Besley et al. (2004). Because the value of the 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.[21] Columns 4-6 report the estimates of the public goods equation. In contrast with what we observe for the welfare transfers in Columns 1-3, there is no evidence that public goods are being targeted to the representative's caste (or caste grouping) within the ward.

4.2 Group Size and the Supply of Public Resources

Proposition 1 indicates that group-specific cooperation results in larger groups putting forward more competent representatives. 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 competence of the elected ward representative by his years of schooling. Apart from the skills that it provides, education is associated with many individual characteristics that determine political competence in Indian local governments.[22] The village census in the 2006 REDS provides the years of schooling of each household head. The coefficient estimates in Table 4 indicate that among household heads in the 2006 REDS census, there is a positive association between schooling, having managerial experience, and the size of landholdings, within the caste and the ward. Individuals who manage large enterprizes, 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.

[21] If 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 4-6 at the household level, with household fixed effects, to be consistent with the BPL regressions in Columns 1-3. Street-level regression results, with street fixed effects, provide qualitatively similar estimates (available from the authors).

[22] Education has also been associated with the competence of political representatives in other settings; see, for example, Dal Bo et al. (forthcoming).

Table 4: **Schooling and Measures of Competence**

Dependent variable	Managerial occupation	Land wealth
	(1)	(2)
Schooling	0.0261***	8658***
	(0.00758)	(3447)
Age	0.109***	6268
	(0.0167)	(9257)
Age squared	-0.00083***	14.6
	(0.000175)	(70.1)
N	76,428	76,428

The sample consists of male household heads aged 25-60 in the 2006 REDS Village Census. Managerial occupation indicates wether the head is a landowning farmer or a businessman. Land wealth is measured in Rupees. All specifications include ward-caste fixed effects.
Standard errors clustered at the village-level in parentheses. *** p<0.01

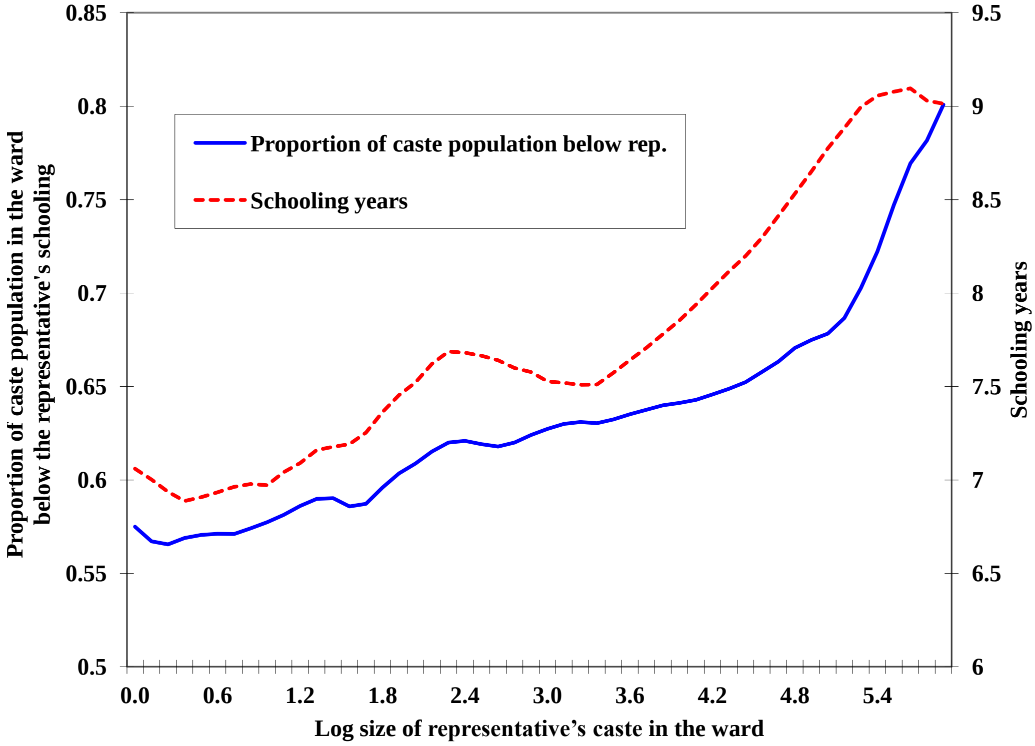
The village inventory in the 2006 REDS includes a special module that collected information on the education of the elected representative from each ward in each of the last three election terms.[23] Figure 2 provides support for Proposition 1, showing that the education of the elected representative, our measure of competence, is increasing continuously in the size of his caste in the ward.[24] If the education (competence) distribution is uncorrelated with or decreasing in group size, as shown below, then Proposition 1 implies, in addition, that representatives of larger castes will be drawn from a higher point in the education distribution of their caste in the ward. We see in Figure 2 that there is indeed a positive and continuous relationship between the size of the representative's caste in the ward in a given election term and his position in his caste's education distribution; starting from just above the median level of education and reaching as high as the 80^{th} percentile. This result provides support for assumption A3. Recall that if this assumption was false; i.e. if the returns to ability were increasing more steeply in public office than in the private sector, then the representative would always be drawn from the top of the ability distribution in his caste, regardless of its size.

The model also indicates, from equation (4), that there should be a positive relationship between the supply of public goods in the ward and the size of the elected representative's ethnic group. As noted, we measure public good provision at the street level by the fraction of the six major public goods for which there were expenditures on new construction or maintenance in a given election term. The corresponding ward-level measure is the population weighted average of the street-level statistic. The implicit assumption when constructing a single aggregate measure is that caste size, the source of

[23]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. Years 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.

[24]We do not distinguish between male and female representatives in the empirical analysis. The gender of the representative is uncorrelated with the size of the representative's caste in the ward, conditional on the usual vector of election characteristics (these results are available from the authors). Female reservation is also independent of the reservation by caste, which is the source of forcing variation in the robust tests of group-specific cooperation that follow.

Figure 2: **Group Size and Representative's Schooling**

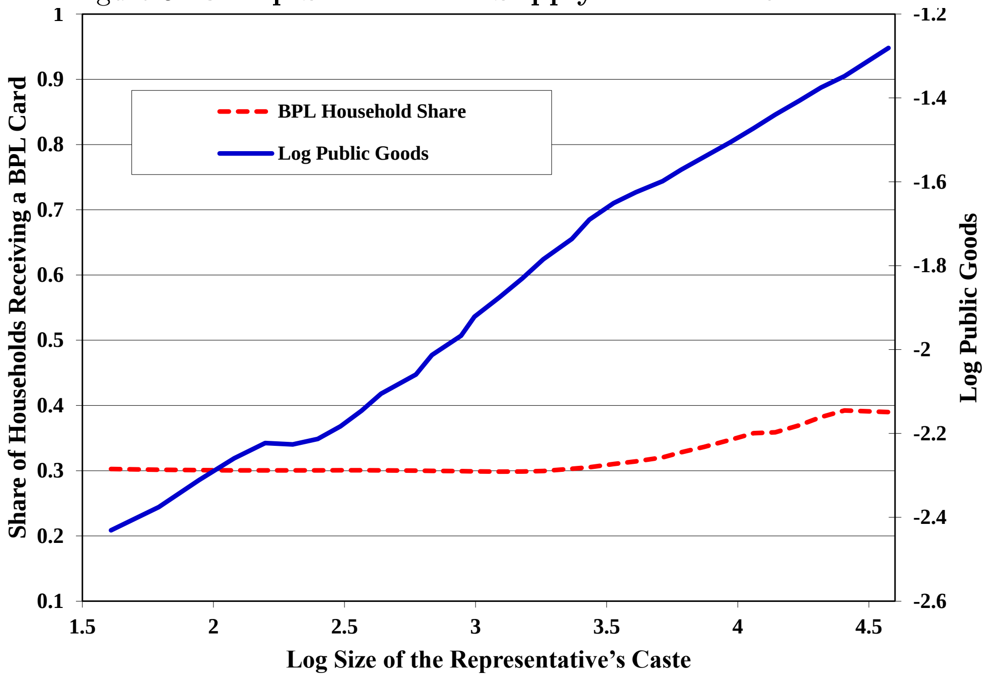


forcing variation in our model, (weakly) positively affects the supply of each public good. Empirical support for this assumption will be provided below; indeed, we will be unable to reject that the caste size effect is the same for each public good.[25] This implies that our aggregate measure can be used to estimate the relationship between caste size and the supply of public goods with a single equation. We estimate this equation in logs to be consistent with the specification of the structural equation that is estimated below. As implied by the model, the supply of public goods in a given election term is increasing continuously in the size of the representative's caste in the ward in Figure 3. In contrast, the probability that a household in the ward receives a BPL card is uncorrelated with the size of the representative's group (on average) in Figure 3, which is consistent with assumption A1 and the results reported in Table 3.

Proposition 1 states that the competence of the elected representative and the supply of public goods will be increasing in the number of co-ethnics who benefit from the public good; i.e. on the size of his ethnic group *in the constituency.* A strong test of our theory is that what matters is the size of the caste in the ward rather than the size of the caste in the village or the size of broader caste groupings such as SC, ST, OBC, in the ward. An additional assumption of our theory is that candidate selection does not depend on the probability that the group will come to power. In particular, we do not expect to observe a change in the selection process when a caste has an absolute majority in the

[25]Note that this does not imply that the average level of each public good is the same. Appendix Table A2 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.

Figure 3: **Group Size and the Supply of Public Resources**



ward, which we know from Figure 1 is accompanied by a discrete increase in the probability that its representative will be selected.

Table 5 reports the results of these tests, with the representative's education in years, his position in his caste's education distribution in the ward, and our measure of public good provision as the dependent variables. The benchmark specification, matching Figures 2-3, has the size of the representative's caste in the ward as the only regressor. The augmented specification includes the size of the representative's caste in the village, the size of the broader caste grouping in the ward, and a discrete variable indicating whether his caste has an absolute majority in the ward. The usual election characteristics – election-term fixed effects, the election year, and reservation (SC, ST, OBC) fixed effects – which could independently determine the representative's competence and the supply of public goods are included as regressors. The results with the benchmark specification indicate that the positive relationship between the size of the representative's caste and each outcome in Figures 2-3 is statistically significant. Results with the augmented specification indicate that conditional on the size of the representative's caste in the ward, its size in the village, the size of the broader caste grouping in the ward, or the fact that the caste has an absolute majority in the ward have no bearing (individually or jointly) on any of the outcomes.[26] Contrast these results with what we obtained in Table 2, where the outcome was the probability that the representative was drawn from the largest eligible caste. In that table, what mattered was whether the largest eligible caste had an absolute majority, with its size having no bearing on the outcome.

[26]Appendix Table C3 reports the estimated coefficients with the same augmented specification, except that the dependent variable is different points in the representative's caste's schooling distribution in the ward; the 25^{th} percentile, the median, and the 75^{th} percentile. These statistics can be computed from the 2006 REDS village census, which collected the schooling of all household heads. In contrast with the representative's schooling, each point in the schooling distribution is uncorrelated with or decreasing with the size of the representative's group.

Table 5: **Log of the Ward Representative's Caste Size in the Ward and His Schooling, His Location in his Caste's Schooling Distribution in the Ward and the Log Public Goods Fraction in the Ward**

Dependent variable	Representative's schooling (years)		Proportion of caste in the ward below representative's schooling		Log public goods	
	(1)	(2)	(3)	(4)	(5)	(6)
Log size of elected representative's caste in the ward	0 .614**	0.950**	0.0367**	0.0390**	0.336**	0.345**
	(0.187)	(0.418)	(0.00918)	(0.0211)	(0.130)	(0.178)
Log size of the elected representative's caste grouping in the ward (cgw)	–	-0.578	–	0.00250	–	-0.178
		(0.430)		(0.0208)		(0.231)
Log size of the elected representative's caste in the village (sv)	–	0.0816	–	-0.00383	–	0.179
		(0.265)		(0.0122)		(0.159)
Whether the representative's caste has a majority in the ward (mw)	–	-0.165	–	-0.0236	–	-0.320
		(0.449)		(0.0210)		(0.212)
Reservation type fixed effects	No	Yes	No	Yes	No	Yes
Election year	No	Yes	No	Yes	No	Yes
Election-term fixed effects	No	Yes	No	Yes	No	Yes
$H_0 : mw, cgw, sv = 0,\ F(3, x),\ [p]$	–	0.65 [0.586]	–	0.60 [0.613]	–	0.85 [0.469]
Sample mean of dependent variable	8.02	8.02	0.791	0.791	-2.04	-2.04
N	813	813	800	800	681	681

Samples include only wards where the caste of the elected representative can be identified and there is information on the dependent variable. Standard errors clustered at the ward level in parentheses. ** p<0.05. One tail tests of joint significance.

5 Ethnic Size-Distribution and the Supply of Public Goods

The preceding analysis provides empirical support for group-specific cooperation as implied by Proposition 1; the elected representative's competence, measured by his education, and the supply of public goods are increasing in the size of his caste in the ward. Proposition 2(a) provides a robust test of this relationship, which accounts for the fact that the caste that the elected representative is drawn from is determined by the electoral process. Across a sample of wards sorted by the population share of the largest caste, there will be a discrete increase in the size of the elected representative's caste, which maps into an accompanying increase in the representative's competence and the supply of public goods, when the share crosses a threshold. The location of the threshold is of independent interest from Proposition 2(b). Given the evidence on the targeting of welfare transfers provided above, the location of the threshold tells us how the additional responsibility of administering the welfare transfers changes the supply of public goods (by changing which caste's representative gets elected).

5.1 Identification Strategy

Proposition 2 relates the ethnic size-distribution of the constituency, measured by the population share of the largest caste in the ward, to the supply of public goods. However, the size-distribution could also determine public good provision through the demand-side channel. To disentangle the supply-side channel from the demand-side channel, we take advantage of the panel data on elections, and 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 size-distribution over time allows us to subsume all permanent characteristics of the ward, including its population-size, the demand for public goods by the electorate, and the effect of any dominant (elite) caste in a fixed effect.[27]

The ward fixed effects will also capture the effect of any dominant caste in the ward on the supply of public goods. Anderson et al.'s (2015) model of local politics describes an environment in which an elite group drawn from a particular caste captures political power, compensating the rest of the electorate in return. The primary determinant of public good provision in their model is the share of land owned by a specific caste (the Marathas). Although the population share of that caste in the village does interact with the land ownership share to determine the supply of public goods, there is no direct mapping to standard measures of ethnic diversity (which are not based on the population share of a specific caste). In addition, when political power is captured by a single caste, public resources will not vary over time within a ward depending on the set of eligible castes, as implied by our model. Anderson et al. note that elite capture of this sort will not be universally salient. Indeed, other studies; e.g. Bardhan and Mookherjee (2006) find little evidence of elite capture in Indian local governments. Where elite capture is salient in our data, the effect will be subsumed in the ward fixed

[27]The implicit assumption when we include the ward fixed effect is that the population of the ward, or the electorate, remains unchanged over time. This is a reasonable assumption given the low spatial mobility that is characteristic of rural India. Munshi and Rosenzweig (2016), for example, estimate that the annual rate of permanent household migration is less than one percent. Consistent with this statistic, the 2006 REDS village census indicates that since 1991 when the local governments were first established, less than 3 percent of new households had migrated into the sample villages.

effect.

Proposition 2(a) states that there will be a discontinuous increase in the competence of elected representatives and the supply of public goods when the population share of the largest ethnic group (caste) crosses a threshold. 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. 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 their share of 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 6 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.

Table 6: **Population Share of the Largest Eligible Caste, by Election Type**

Type of election	Open (1)	SC (2)	ST (3)	OBC (4)
Panel A - Distribution of shares				
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
Panel B - Fraction of ward-terms where largest share exceeds				
0.25	0.97	0.48	0.64	0.69
0.50	0.66	0.29	0.50	0.44
0.75	0.36	0.18	0.41	0.21
Fraction of ward terms	0.66	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.
Source: REDS Village Census and Village Inventory.

Panel B of Table 6 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 6, 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 (lower caste) elections. All our augmented specifications have included a full set of reservation fixed effects, and the specifications that follow will do the same (in addition to ward fixed effects).

Ward representatives operate independently in our model. 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 determine the supply of public goods, the equilibrium in other wards, which determines who is elected to represent them, will also be relevant to the extent that there is a fixed component to the total resources available in the *panchayat*. Nevertheless, if reservation is independent in elections across wards and for the president's position, then this implies that the political equilibrium in a ward-term, measured by the population share of the largest eligible caste, will be orthogonal to the equilibrium in that term in other wards. The political equilibrium in other wards can then be ignored when estimating the relationship between the population share of the largest eligible caste and the supply of public goods in the ward. To test the orthogonality assumption, we constructed the following two variables for each ward in each election term: (i) the population share of the largest eligible caste averaged across all *other* wards in the *panchayat*, and (ii) the corresponding average with the population share replaced by a binary variable indicating whether the largest eligible caste in each ward had an absolute majority. We then verified that both the population share of the largest eligible caste in the reference ward-term and the binary variable indicating whether it had an absolute majority in the ward were uncorrelated with the corresponding *panchayat*-level statistics (results available from the authors).

5.2 Locating a Threshold

To test Proposition 2 we first semi-parametrically examine the relationship between the population share of the largest eligible caste and each of the outcomes by estimating the following equation:

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

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 competence, 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 fixed effects, election-term fixed effects, and the election year; and ζ_{jt} is a mean-zero disturbance term. The additional regressors, X_{jt}, are partialled out using a procedure described in Appendix B that is robust to alternative specifications, leaving us with the conditional outcome, $\tilde{y}_{jt}$.

The conditional outcome $\tilde{y}_{jt}$ is then regressed nonparametrically on S_{jt}:

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

Figure 4 reports nonparametric estimates of equation (11) with the elected representative's education and the supply of public goods in each ward-term as outcomes.[28] Both outcomes increase (discontinuously) at a population share of 0.4, and continue to increase till 0.6, in Figure 4. Although the increase at the threshold is very steep, it is not vertical. We will see momentarily that this is a consequence of the estimation procedure, which uses a local polynomial smoother. Notice that the two outcomes track closely over the entire range of population shares, highlighting the close relationship between the representative's competence and public good provision that is an important feature of our model.

Figure 4: **Public Goods and the Elected Representative's Education by the population Shgare of the Largest Eligible Caste**

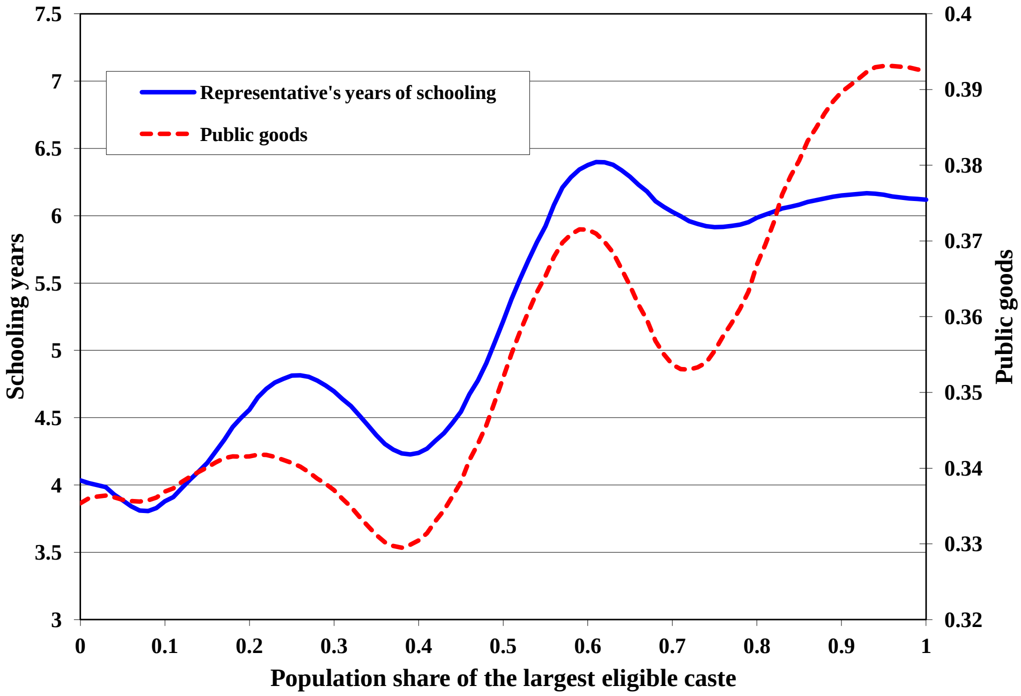


Figure 4 suggests that there is a discontinuous increase in the representative's competence and the supply of the public good when the population share of the most numerous caste crosses a threshold. We now proceed to formally test for the presence of a threshold, place statistical bounds on its location, and estimate the gain in the representative's competence 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:

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

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.

[28]The model implicitly assumes that there is a large number of individuals in each ethnic group, which allows it to select its optimal representative. In practice, the elected representative's competence, and the associated supply of public goods, will be closer on average to the optimal level in larger wards. Because the population share threshold is consequently more accurately located in large wards, we weight both the first-step and the second-step regressions by the ward population.

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 model, 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,

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

which is simply a normalization of the Residual Sum of Squares statistic, $RSS(S)$, based on its minimum value, $RSS(\hat{\underline{S}})$, 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, $\hat{\underline{S}}$, 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 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 5 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, exhibits an increase from around 0.4 that continues until 0.6. This is precisely what we observed for the conditional outcomes in Figure 4, 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 increases continuously with the population share. The test that we describe below will pin down the threshold for the V_S variable at precisely 0.5.

Figure 6 plots the $LR(S)$ statistic for hypothesized thresholds, $S \in [0.25, 0.75]$ in increments of 0.001.[29] This statistic reaches its minimum value, which we noted is zero by construction, when the

[29] 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.

Figure 5: **Simulated Variables, by the Population Share of the Largest Caste**

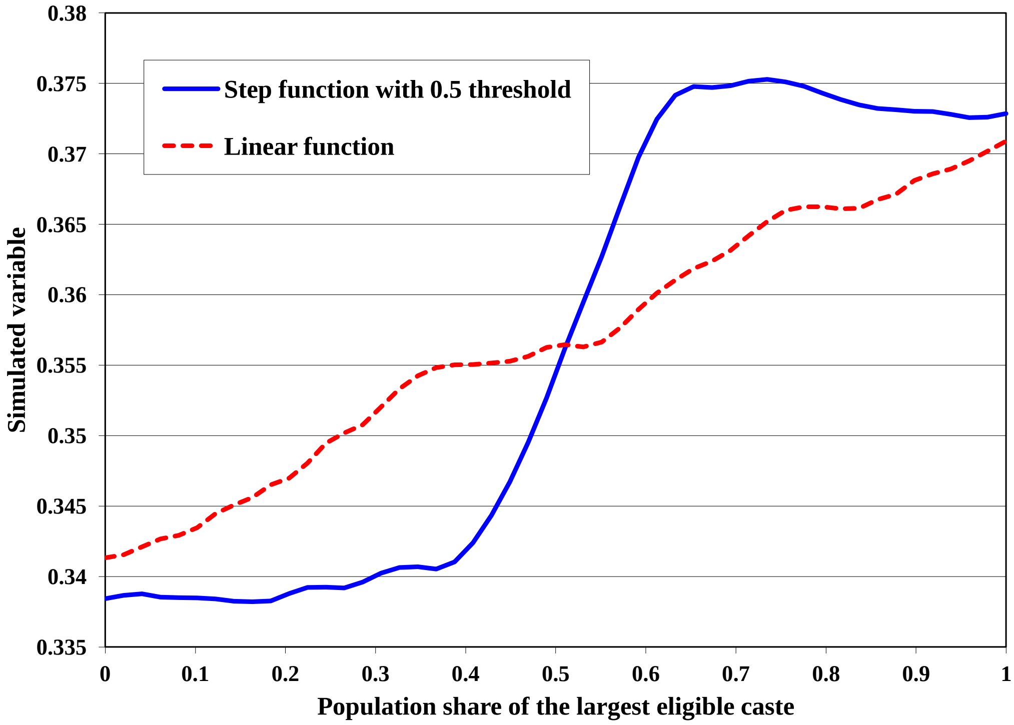
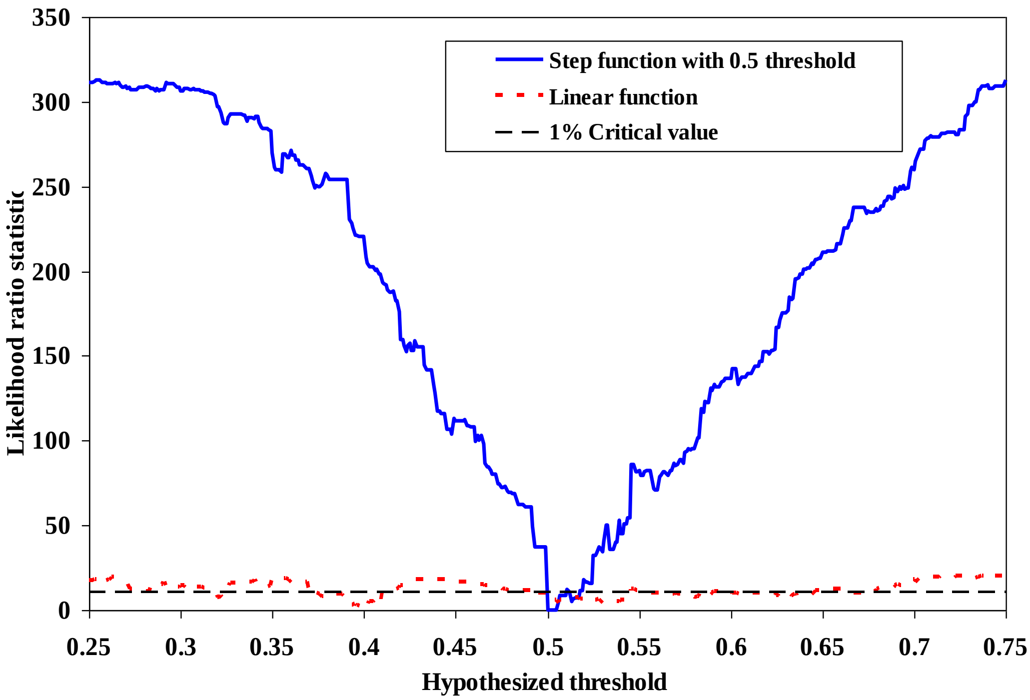


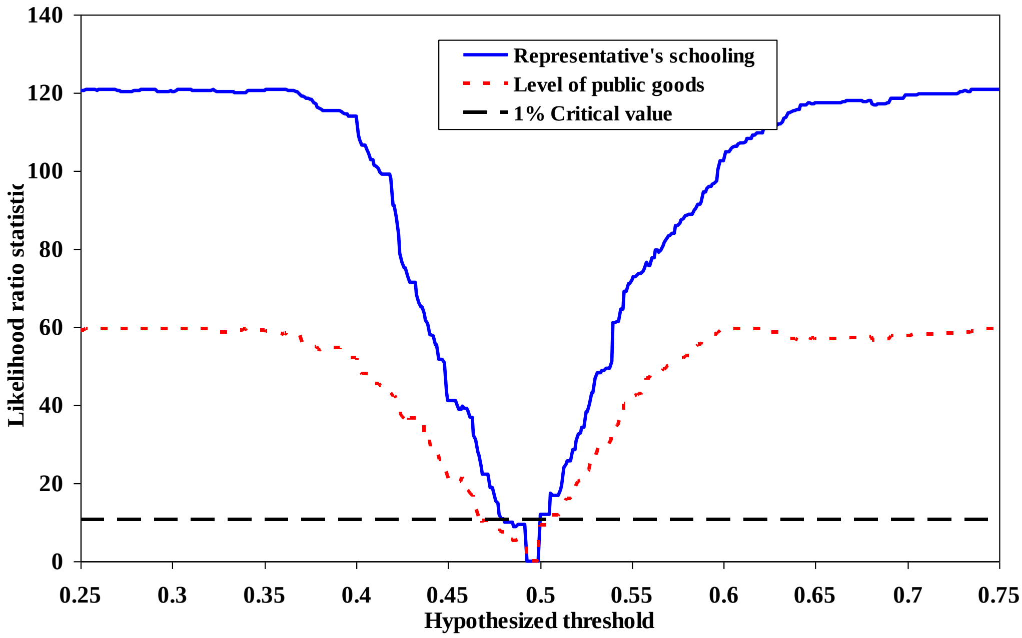
Figure 6: **Threshold Tests on Simulated Data: Step versus Linear Function**



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 6, 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, but not for the V_L variable.

We now proceed to apply the same procedure to the REDS data used to estimate Figure 4. 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 7, separately for each outcome.[30] Variation in the LR statistic across the range of hypothesized thresholds matches closely with the pattern for the V_S variable in Figure 6, which was constructed to increase vertically at a population share of 0.5. The Critical Value (CV) line bounds the location of the threshold. Given the tight bounds, we can infer with a high degree of confidence that a threshold exists for both the representative's education and public good provision. The LR statistic reaches its minimum value, which is zero by construction, at a hypothesized threshold of 0.5 for both outcomes. We can thus infer, in addition, that the threshold is the same, and that it is located at 0.5, for both outcomes.

Figure 7: **Threshold Tests on Real Data: Public Goods and the Elected Representative's Schooling**



As observed in Figure 1, there is a discrete increase in the probability that the largest caste comes to power when it has an absolute majority in the ward. Given the associated increase in the size

[30]As discussed in Appendix B, Appendix Figures C1 and C2 assess the robustness of the results to alternative specifications of the $\phi(S_{jt})$ function in equation(10). Appendix Figure C3 assesses the robustness of the threshold test to a flexible specification of equation (12) that allows for a linear relationship between the outcome $\tilde{y}_{jt}$ and the share S_{jt} below the threshold, a mean-shift at the threshold, and a linear relationship (with a possibly different slope and sign) above the threshold. This robustness test is useful because while the model predicts that there will be a discrete increase in the outcomes of interest at a population share threshold, it is less informative about the relationship between the population share of the largest eligible caste and these outcomes above and below the threshold.

of the group in power, the discrete increase in the representative's competence and the supply of public goods at the 0.5 threshold provides us with robust evidence of group-specific cooperation from Proposition 1. The same research design, exploiting a discontinuous change in the size of the group in power at a threshold, can also be used to test the assumption in the model that the total amount of welfare transfers, T, is exogenously determined. Recall from Table 3 and Figure 3 that the size of the representative's caste in the ward has no bearing on the probability that its residents receive a BPL card (on average). A more stringent test is that at the 0.5 population share threshold, where there is a discrete increase in the size of the caste in power, there should be no change in the average probability of receiving a BPL card in the ward. Appendix Figure C4 uses the procedure described above to show the relationship between the population share of the largest eligible caste and both the probability of receiving a BPL card and public good provision (as a basis for comparison) at the ward-term level. We see in the figure that there is no change in the probability of receiving a BPL card at a population share of 0.5. Formal tests for a threshold in Appendix Figure C5, moreover, find no evidence of a threshold for welfare transfers, in contrast to what we observe for public goods.

5.3 Gain at the Threshold

Having precisely located the population share 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:

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

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 2 in Table 7 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. The size of the estimated coefficients matches the discontinuous increase in public good provision and the representative's schooling in Figure 4. Our estimates indicate that there is a 13 percent increase in the supply of public goods and a two-year increase in the schooling of the elected representative at the 0.5 threshold. Recall that the effect of caste size on the supply of each public good must be (weakly) positive to interpret the increase in our aggregate measure of public goods, the fraction of goods received on each street and then averaged across streets in the ward, as an unambiguous increase in the total supply. Appendix Table C4 tests this assumption by separately estimating the mean-shift at the threshold for each of the six public goods. All of the individual point estimates are positive and, moreover, we cannot reject the hypothesis that they are equal, while the joint effect is highly significant, matching the result in Table 7.

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

Table 7: **Change in Public Goods Levels and Representative's Schooling at the Population Share Threshold of 0.5**

Dependent variable	Public goods (1)	Representative's schooling (years) (2)
Mean-shift at threshold	0.042** (0.021)	2.264*** (0.802)
Threshold location	0.50	0.50
Mean of dependent variable below the threshold	0.335	4.157
N	1666	1591

All specifications include ward fixed effects, reservation fixed effects, election-term fixed effects, and election year. Standard errors clustered at the panchayat-term level are in parentheses. ** p<0.05, *** p<0.01.

6 Alternative Explanations

The empirical results provide robust evidence that the competence of the elected representative and the supply of public goods is increasing in the size of his caste in the ward. Our interpretation of these findings is based on within-group cooperation, with the gains from public good provision increasing in group size (the number of beneficiaries). In this section we examine other interpretations of the group-size effect.

1. The demand for public goods varies with group size. Although a caste will be spread over a wide area covering many villages, there will typically be variation in the number of caste members across villages. Suppose that particular types of individuals, due to selection or historical accident, end up in villages with a large number of co-ethnics. Suppose, in addition, that these individuals have a particular preference for public goods. Then this would independently generate a positive correlation between group size and public good provision.

Note, however, that the preceding argument is based on the size of the caste in the village. The size of the caste in the ward, which is a recent political demarcation, should have no bearing on the preference for public goods. In contrast, what matters in our model is the number of beneficiaries of the public good; i.e. the number of co-ethnics in the ward rather than the village. Consistent with our model, and contrary to the implications of the alternative explanation, recall from Table 5 that it is the size of the representative's caste in the ward that determines his competence and the supply of public goods. Conditional on the size of the caste in the ward, its size in the village has no bearing on either outcome.

2. Ability in the population varies with group size. Each caste in a ward consists of a relatively small number of households. If the objective of all castes was to put forward their most competent candidate, then representatives belonging to larger castes could be more competent, even if the competence distribution is the same in all castes, just by chance (because of the small numbers). Alternatively, if the competence distribution is superior in larger castes, then representatives from those castes would be more competent even if they were randomly selected.

We verify empirically in Appendix Table C3 that the competence distribution is weakly decreasing in group size, ruling out the latter argument. Moreover, in our model the objective of each caste is *not* to put forward its most competent member as the candidate. Given that the competence distribution is weakly decreasing in group size, the candidates selected by larger groups will be systematically drawn from higher in their group's competence distribution. This is precisely what we observe in Figure 2; the position of the elected representative in his caste's education distribution increases monotonically from the 55^{th} percentile to the 80^{th} percentile over the range of caste sizes. This result cannot be explained by the small numbers argument.

3. Clientelism. In our model, larger ethnic groups put forward more competent candidates, who supply a higher level of public goods if elected, because more co-ethnics benefit. An alternative explanation for the positive relationship between group size and public good provision is based on clientelism; a larger group is a larger vote bank and, thus, it may be in the interest of the *panchayat* president or someone else at a higher level of government to allocate additional public resources to the ward when a larger group is in power in return for its votes (in the future). This top-down explanation cannot, however, explain the accompanying positive relationship between group size and the competence of the representative. This is a key feature of our model and is one channel through which group size affects the supply of public goods. We would not expect to observe this relationship if the residents of the ward were passive beneficiaries of patronage from above.

7 Structural Estimation and Counter-Factual Simulations

In this section we estimate the structural parameters of the model to (i) validate the assumption that $0 < \beta < 1/4$, (ii) quantify the under-supply of public goods due to the fact that the scope of cooperation is restricted to the caste, and (iii) quantify the impact of specific policies that based on the theory and the location of the threshold at 0.5 are expected to reduce the supply of public goods.[32]

From equation (4), the supply of public goods, $G \equiv a^{\beta}$, is:

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

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 model and the data diverge is the assumption that all households in the ward receive the same supply of public goods. In the data,

[32] Although we require that $\beta < 1/2$ to derive Proposition 1, recall that we need $\beta < 1/4$ to derive the stronger condition, in Appendix A, that if the feasibility condition is not satisfied at $S^{**} < 0.5$, then it will not be satisfied for any $S > S^{**}$.

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 because of permanent differences in the cost of construction across streets.[33] Public good provision on street, s, in ward, j, and term, t, is thus denoted by $G_{sjt} = F_{sj} \cdot G_{jt}$, where F_{sj} maps ward-level effort into street-level public goods.

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

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

where $f_{sj} \equiv log(F_{sj})$ and ν_{sjt} is a mean-zero measurement error term. Notice from equation (15) that the relationship between public goods and the size of the representative's caste in the ward is linear in logs, which is consistent with Figure 3.

The structural parameters, α and β, are evidently identified in the preceding equation and can be estimated using Nonlinear Least Squares.[34] The point estimates are reported in Table 8. 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 Proposition 1, which implies 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. 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 in magnitude for lower castes (although the difference is not statistically significant) possibly because members of these disadvantaged groups have less opportunities to use their schooling. However, β 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) there was no cooperation, and (ii) all residents of the constituency were able to cooperate to achieve the first-best allocation of resources. 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 8. The representative's caste in the ward comprises 29 households on average. Thus, while the counter-factual simulations indicate that caste-specific cooperation does increase the supply of public goods, it is evident from the figure that there is a substantial under-supply relative to the first-best on average.

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

[34]We weight by the number of households on the street when estimating equation (15) to allow for heteroscedasticity in the residual term, ν_{sjt}. Standard errors are clustered at the ward level.

Table 8: **Structural Parameter Estimates**

Parameter	(1)	(2)
α	0.0148	–
	(0.0128)	
α Upper castes	–	0.945
		(0.746)
α Lower castes	–	0.0151
		(.0144)
β	0.122***	0.122***
	(0.0450)	(0.0450)
$H_0 : \beta < 0$, p	0.004	0.004
$H_0 : \beta > 0.25$, p	0.000	0.000
$H_0 : \frac{2\beta}{1-2\beta} = 0$, p	0.041	0.041
$H_0 : \alpha$ Upper, α Lower< 0, p	–	0.043
$H_0 : \alpha$ Upper $= \alpha$ Lower, p	–	0.217
N	1957	1957

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. All specifications include street fixed effects. Standard errors clustered at the ward level in parentheses. *** p<0.01.

The fact that the threshold is located at 0.5 implies, from Proposition 2(b), that the supply of public goods would increase if the representative was no longer made responsible for distributing welfare transfers. The estimated structural parameters can be used to quantify the magnitude of this increase.

In our data, the representative of the largest eligible caste is elected 80% of the time when it has an absolute majority and 40% 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 assume, as in the model, that the representative of the largest eligible caste is surely elected 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, where the representative is no longer responsible for distributing welfare transfers, the representative of the largest caste will surely 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, S^*, 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 under-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 6, 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 small groups who would otherwise be excluded from the welfare transfers have access to them in at least a fraction of ward-terms. Because small groups capture a smaller fraction of the total transfers, their representatives 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, with the benchmark scenario and the decoupled scenario.

The cummulative density function of our measure of public goods supply across wards is reported in Figure 9 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, the fraction of the ward population receiving each public good, averaged across the six public goods we consider, is less than one-third in 40% of wards in the benchmark simulation.[35] The fraction of under-served wards declines to 35% when welfare transfers are removed from the representative's list of responsibilities. This fraction 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

[35] We specify one-third as the cutoff below which a ward is under-served because this is the average level of public good provision below the 0.5 population share threshold (see Table 7).

larger. However, there are equity consequences if the decoupling policy is not applied simultaneously.

Figure 8: **Caste Size and the supply of Public Goods, Relative to Two Benchmarks**

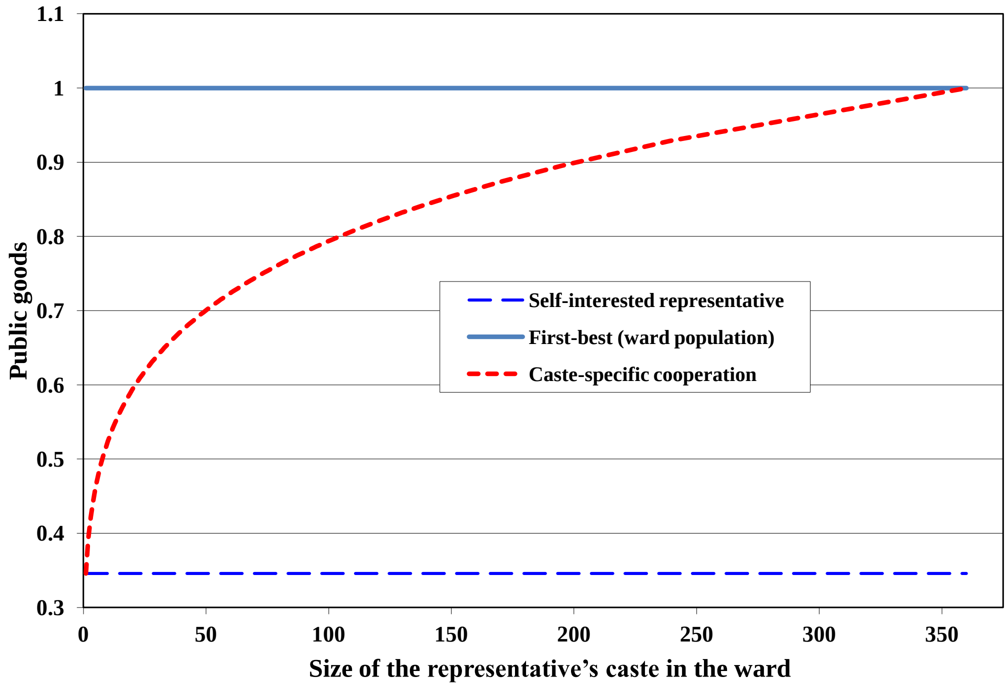
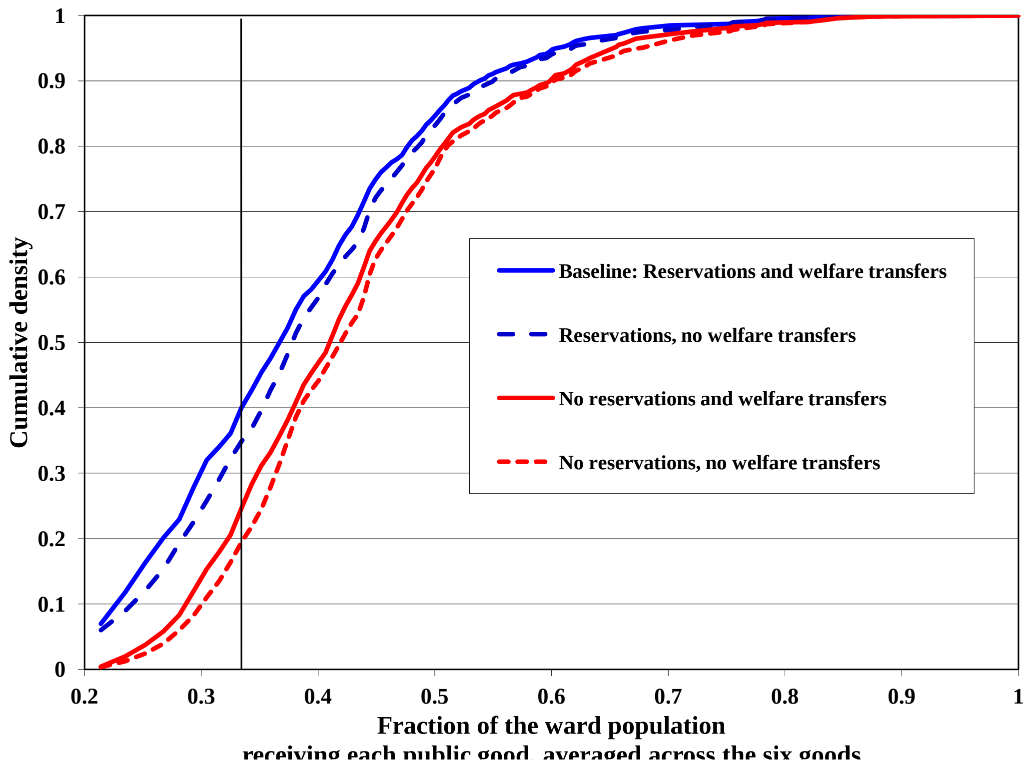


Figure 9: **Counterfactual-Simulation Cumulative Distributions: Fraction of Public Goods Received by a Ward**



The implicit assumption in our policy evaluation is that increased effort by the elected representative independently translates into an increase in the supply of public goods in each ward; i.e. there is no supply constraint at the *panchayat* level. If the *panchayat*-level supply is fixed, then the increase in the representative's caste-size due to decoupling or de-reservation will not increase the ward-level supply on average and will, in addition, have negative distributional consequences (persistently favoring wards with larger castes). However, if more competent representatives are able to bring back more resources from the center or use existing resources more effectively, then these policies could increase the supply of public goods in all wards. To the extent that the *panchayat*-level supply constraint binds, Figure 9 provides an upper-bound on the magnitude of the increase. Note that this caveat does not affect our tests of caste-specific cooperation and the model more generally because more competent representatives who exert more effort will increase the supply of public goods in their ward even if the *panchayat*-level supply is constant.

8 Conclusion

In this paper, we extended the citizen-candidate model to allow for cooperation within, but not across, ethnic groups. Our model delivers predictions for the relationship between ethnic diversity, the distribution of welfare transfers, and the supply of public goods by democratically elected local governments, which are increasing in influence throughout the world. In common with previous models that incorporate group-based cooperation as a solution to the collective action problem, our model predicts that representatives belonging to larger ethnic groups will supply more public goods and that there will be a discrete increase in public good provision when the population share of the largest group crosses a threshold level. An additional element of our group-based model, which captures a common feature of local governments in developing countries, is that the representative is responsible for the delivery of two types of public resources: a non-excludable public good and targetable welfare transfers. If the total supply of the welfare transfers is exogenously determined, and the representative first targets his own group, then outsiders will be increasingly crowded out with respect to the welfare transfers when a larger group is in power (as we find empirically). This makes large groups relatively unpopular with the electorate and is another reason, apart from free-riding on smaller groups, for why the elected representative may not always be drawn from the largest group.

Based on the model, the precise location of the threshold is informative about preferred policy. If the threshold is located at a population share of 0.5; i.e. that there is a discrete increase in the probability that the largest group comes to power when it has an absolute majority (with an associated increase in the supply of public goods) then this implies that the crowding out of the outsiders with respect to the welfare transfers dominates the increased supply of public goods that the largest group's representative provides. The supply of public goods would then be increased by removing the distribution of welfare transfers from the representative's list of responsibilities. If the threshold is located anywhere else, above or below 0.5, then making representatives responsible for the welfare transfers increases the supply of public goods, by reducing the incentive of the largest group to free-ride on smaller groups.

We test the hypothesis that cooperation is restricted to the representative's ethnic group (which is the caste or *jati* in India) and, hence, that the supply of public goods is increasing in its size, and identify the location of the threshold, using newly available data over multiple election terms at the most local (ward) level. The panel data allow us to subsume the demand for public goods in a ward fixed effect, while rotating set asides, which change the pool of castes that are eligible to stand for election from one term to the next, generate exogenous variation in the size of the representative's ethnic group over time. We find support for both the hypothesis of within-group cooperation as well as for competition between groups for targetable welfare transfers. Moreover, we find that in Indian local elections, the threshold is precisely located at 0.5.

The location of the estimated threshold indicates that the average size of the representative's group and, hence, the supply of public goods would be increased if representatives were no longer made responsible for both the supply of public goods and the distribution of welfare transfers. The model also provides another channel through which the supply of public goods could be increased, by rolling back the caste-based reservation system at the ward level. This would increase the size of eligible castes and, hence, the average size of the elected representative's caste. Without decoupling public good provision and the administration of welfare programs, however, de-reservation will also have adverse equity consequences *within* the ward because outsiders will be disproportionately excluded from these programs when the representative belongs to a larger caste. Moreover, to the extent that the supply of public goods is fixed at the *panchayat* level, de-reservation will, in addition, have adverse distributional consequences *across* wards by persistently favoring wards with castes that have a dominant majority.

Our estimates of the theoretical model's structural parameters allow us to quantify the magnitude of the under-supply of public goods due to group-specific cooperation, indicating that ethnic diversity (by reducing the size of the representatives group on average) significantly reduces the supply of local public goods in India. Counter-factual policy simulations based on structural estimates of the model also indicate that the gains in the supply of public goods from decoupling public good provision and the administration of welfare programs are relatively modest, even under the assumption that there are no supply constraints at the *panchayat* level. However, the gains from removing the reservation system could be more substantial, especially if the increased competence and effort by ward representatives drawn from larger groups leads to an overall increase in the supply of public goods to the *panchayat*.

The policy interventions that we consider take as given the fundamental cause of the under-supply of public goods, which is that political institutions are weak in less developed economies. Ethnic-based networks have emerged in these economies because market institutions function imperfectly and we show in this paper that the same ethnic groups have expanded their domain from private economic activity to the public sphere because political institutions also function imperfectly. If political leaders are adequately compensated and political parties actively compete, then ethnic groups will play a reduced role in local politics and the supply of public goods will increase. Ethnic politics is not the root cause of the under-supply of public goods, but a symptom of (and only a second-best response to) a deeper institutional failure.

Appendix A: Characterization of the Threshold

We begin by establishing that the population share threshold 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

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

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.

The first term in square brackets is negative when all groups are of equal size, $N_K = N_j = \frac{N}{K}$, positive when $N_K \to N$ and $N_j \to 0$, and increasing in N_K (as N_j declines). By a continuity argument, 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 $S^{**} < S^*$.

We next examine how the location of the threshold is affected by the feasibility condition. 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,

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

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 of the public good. 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 $-\Sigma_j \frac{dN_j}{dN_K} = 1$. The properties of the PG and WT functions can then be derived as follows:

$$PG'(N_K(N_j)) = \left(\frac{\beta^2}{\alpha}\right)^{\frac{\beta}{1-2\beta}} \frac{2\beta}{1-2\beta}\left[\left(\frac{1}{N_K}\right)^{\frac{1-4\beta}{1-2\beta}} - \left(\frac{1}{N_j}\right)^{\frac{1-4\beta}{1-2\beta}} \frac{dN_j}{dN_K}\right] > 0$$

$$PG''(N_K(N_j)) = -\left(\frac{\beta^2}{\alpha}\right)^{\frac{\beta}{1-2\beta}} \frac{2\beta}{1-2\beta}\frac{1-4\beta}{1-2\beta}\left[\left(\frac{1}{N_K}\right)^{\frac{2(1-3\beta)}{1-2\beta}} - \left(\frac{1}{N_j}\right)^{\frac{2(1-3\beta)}{1-2\beta}} \left(\frac{dN_j}{dN_K}\right)^2\right]$$

$$PG''(N_K(N_j)) \leq 0 \quad for \quad \frac{N_K}{N_j} \leq \left(\frac{-1}{\frac{dN_j}{dN_K}}\right)^{\frac{1-2\beta}{1-3\beta}}$$

$$PG''(N_K(N_j)) > 0 \quad for \quad \frac{N_K}{N_j} > \left(\frac{-1}{\frac{dN_j}{dN_K}}\right)^{\frac{1-2\beta}{1-3\beta}}$$

For $N_K \leq T$:

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

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

For $N_K > T$:

$$WT'(N_K(N_j)) = -\theta(N-T)\frac{\frac{dN_j}{dN_K}}{(N-N_j)^2} > 0$$

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

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 inflexion $N^I(N_j)$, 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 $N_K > T$.

If the feasibility condition is satisfied for all values of N_K, the largest group will surely come to power at S^{**}. If the feasibility condition is not satisfied for any value of N_K, then the largest group will surely come to power when it has an absolute majority; i.e. the threshold will be located at 0.5. If the feasibility condition is satisfied for some but not all values of N_K, we show that there exists a unique population share, S^F, below which the condition is satisfied and above which which the condition is not satisfied.

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 (we are grateful to Soenje Reiche for suggesting this line of reasoning). 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 $S_K > S^F$.

Appendix B: Threshold Test

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

Appendix Figures C1 and C2 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, using each of these alternative specifications, Appendix Figure C1 reports the corresponding 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 4. Threshold tests, reported in Appendix Figure C2, 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 model 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). Appendix Figure C3 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 Appendix Figure C3 that greater flexibility increases the precision of the threshold test, while maintaining the estimated location of the threshold at 0.5.

Appendix C: Tables and Figures

Table C1: **Budget Shares by Discretionary Expenditure Item for the Most Recent Complete Panchayat 5-Year Term**

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 facilites	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

Statistics based on the most recent complete panchayat five-year term in each ward.
Source: 2006 REDS Village Inventory.

Table C2: **Fraction of Households Receiving Public Goods Each Term**

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

SC=scheduled caste, ST=scheduled tribe, OBC=other backward caste. Statistics are based on the last three election-terms in each ward. Means and standard deviations (in parentheses).
Source: REDS Village Census and Village Inventory.

Table C3: **Representative's Caste's Schooling Distribution in the Ward**

Dependent variable/location	25th-percentile	50th-percentile	75th-percentile
Log size of elected representative's caste in the ward	- 0.777***	- 0.450	0.178
	(0.232)	(0.336)	(0.394)
N	813	813	813

All specifications include the log size of the representative's caste in the village, the log size of the representative's caste grouping in the ward, whether the caste has a majority in the ward, election term fixed effects, election year, and reservation fixed effects.
Standard errors clustered at the ward level in parentheses. *** p<0.01.

Table C4: **Change in the Probability that Each Public Good is Received at the Population Share Threshold of 0.5**

Public Good (dep. variable)	Sanitation (1)	Street lights (2)	Clean water (3)	Roads (4)	Communication (5)	Electricity (6)
Mean-shift at threshold	.0873 (.0328)	.0773 (.0198)	.0534 (.0303)	.0242 (.0343)	.0273 (.0171)	.0248 (.0244)
Mean of dep. variable	0.483	0.366	0.748	0.609	0.222	0.369
N	1666	1666	1666	1666	1666	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]		

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

Figure C1: **Public Goods, by the Population Share of the Largest Eligible Caste: Alternative First-Stage Specifications**

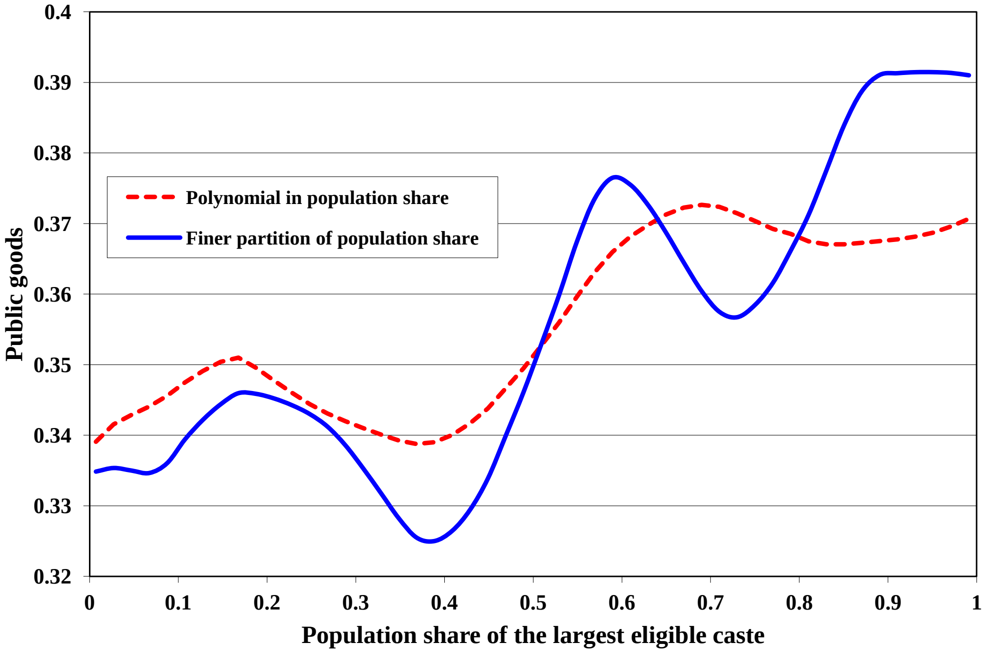


Figure C2: **Threshold Tests: Alternative First-Stage Specifications Likelihood Ratio by Hypothesized Threshold**

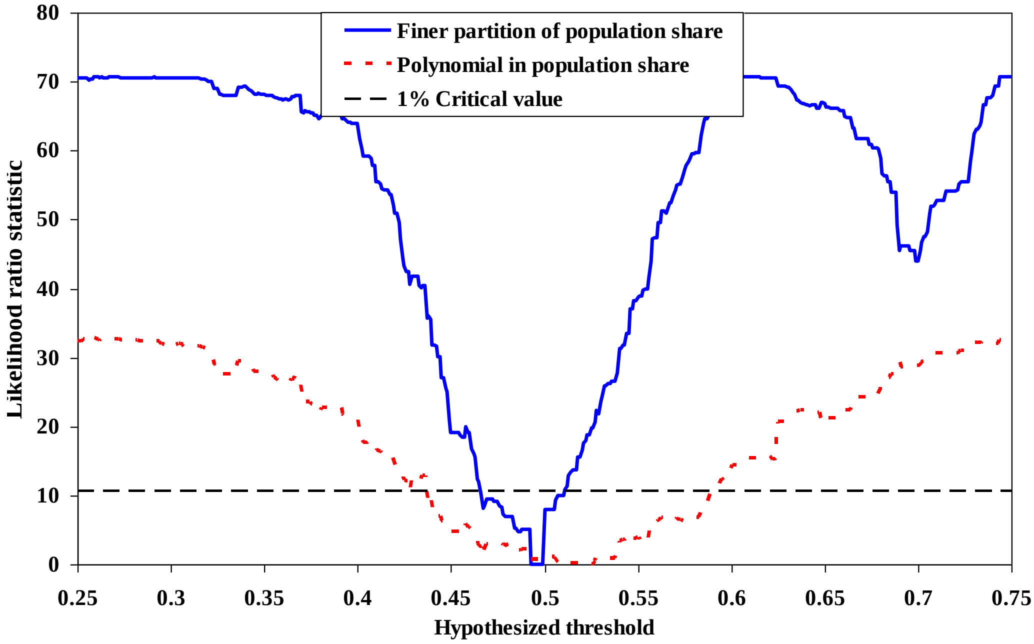


Figure C3: **Threshold Test: Alternative Threshold Functions Likelihood Ratio by Hypothesized Threshold**

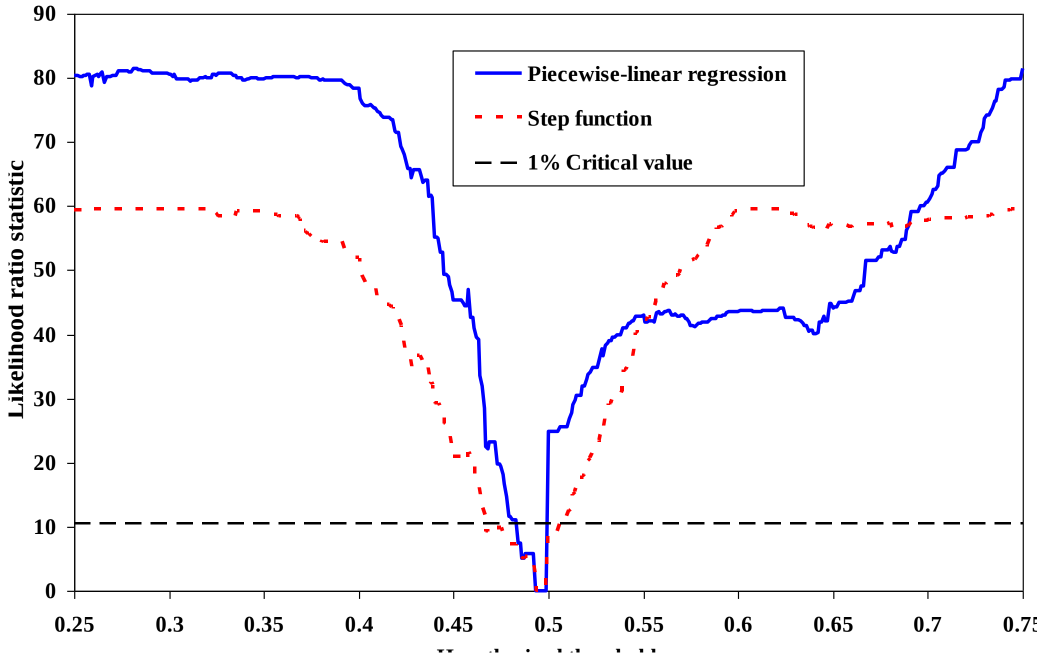


Figure C4: **Public Goods and Welfare Transfers, by the Population Share of the Largest Eligible Caste**

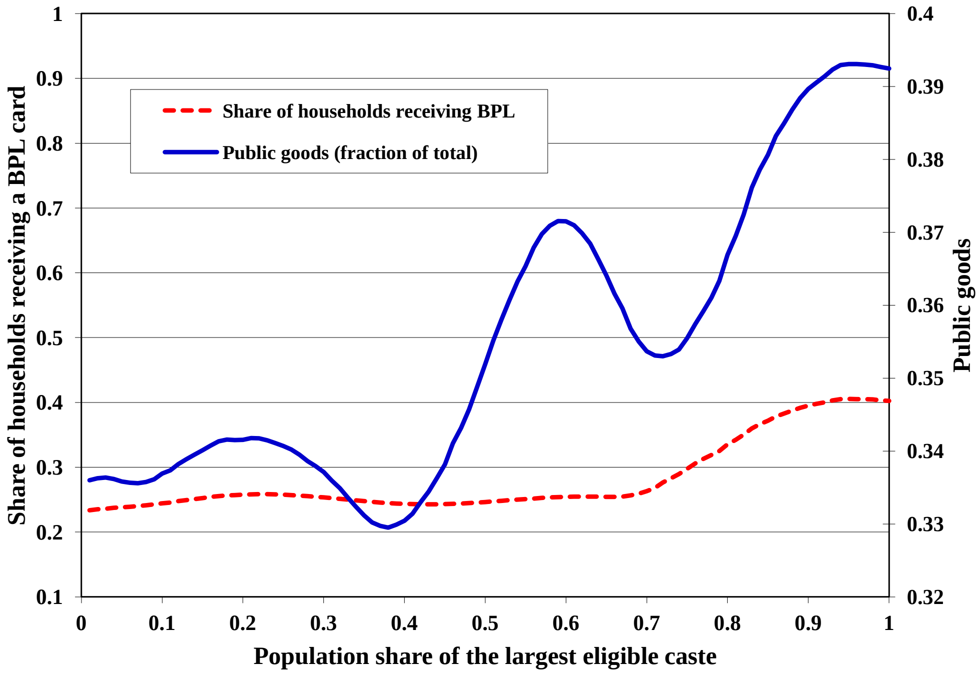
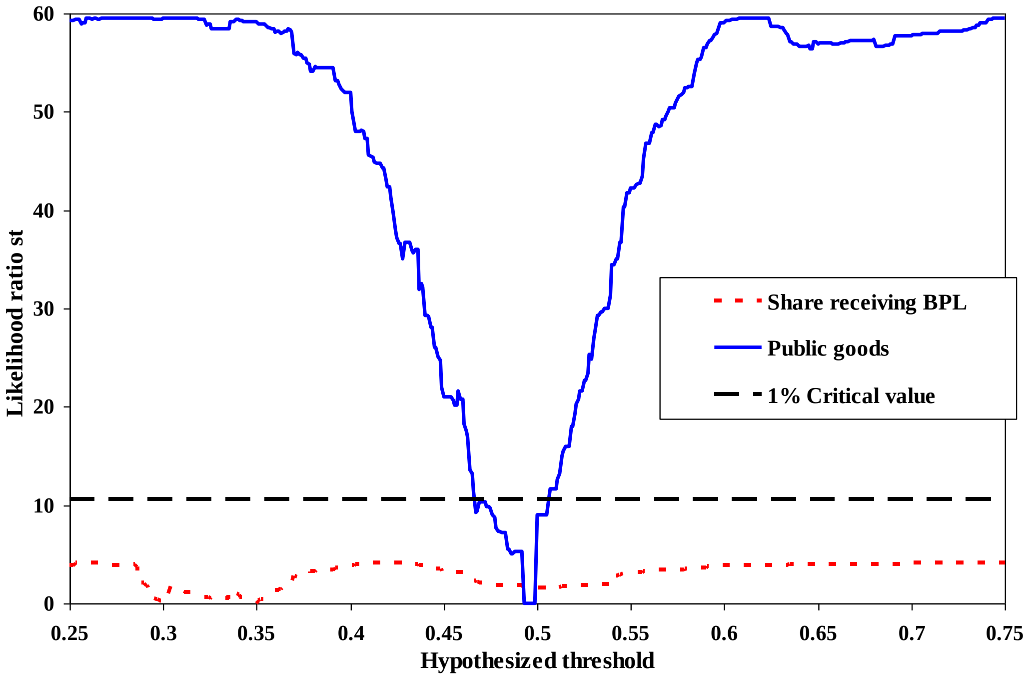


Figure C5: **Threshold Tests: Public Goods and Welfare Transfers**



References

[1] Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics* 114(4): 1243-1284.

[2] Anderson, Siwan, Patrick Francois, and Ashok Kotwal. 2015. "Clientilism in Indian Villages," *American Economic Review* 105(6): 1780-1816.

[3] Andrews, Donald W.K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* 61(4): 821-856.

[4] Banerjee, Abhijit and Rohini Pande. 2009. "Parochial Politics: Ethnic Preferences and Politician Corruption," Harvard University, Kennedy School of Government, typescript.

[5] Bardhan, Pranab and Dilip Mookherjee. 2006. "Pro-Poor Targeting and Accountability of Local Governments in West Bengal," *Journal of Development Economics* 79(2): 303-327.

[6] 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.

[7] 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.

[8] Besley, Timothy and Stephen Coate. 1997. "An Economic Model of Representative Democracy," *Quarterly Journal of Economics* 112(1): 85-114.

[9] Besley, Timothy, Rohini Pande, Lupin Rahman, and Vijayendra Rao. 2004. "The Politics of Public Good Provision: Evidence from Indian Local Governments" *Journal of the European Economic Association* 2(2-3): 416-426.

[10] Besley, Timothy, Rohini Pande, and Vijayendra Rao. 2011. "Just Rewards? Local Politics and Public Resource Allocation in South India," *The World Bank Economic Review* 26(2):191-216.

[11] Brass, Paul. *The Politics of India Since Independence.* Cambridge: Cambridge University Press.

[12] 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.

[13] Casey, Katherine. 2015. "Crossing Party Lines: The Effects of Information on Redistributive Politics," *American Economic Review* 105(8): 2410-2448.

[14] Chandavarkar, Rajnarayan. 1994. *The Origins of Industrial Capitalism in India: Business strategies and the working classes in Bombay, 1900-1940.* Cambridge: Cambridge University Press.

[15] Chattopadhyay, Raghabendra and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica* 72(5): 1409-1443.

[16] Dal Bo, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. Forthcoming. "Who Becomes a Politican?" *Quarterly Journal of Economics.*

[17] Damodaran, Harish. 2008. *India's New Capitalists: Caste, Business, and Industry in a Modern Nation*, Ranikhet: Permanent Black.

[18] Dunning, Thad and Janhavi Nilekani.2013. "Ethnic Quotas and Political Mobilization: Caste, Parties, and Redistribution in Indian Village Councils," *American Political Science Review* 107(1): 35-56.

[19] Ferraz, Claudio and Frederico Finan. 2010. "Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance." Working Paper, University of California at Berkeley.

[20] Ferraz, Claudio and Frederico Finan. 2011. "Electoral Accountability and Corruption in Local Governments: Evidence from Audit Reports." *American Economic Review* 101(4): 1274-1311.

[21] Hansen, Bruce E. 1999. "Threshold effects in non-dynamic panels: Estimation, testing, and inference," *Journal of Econometrics* 93: 345-368.

[22] Hayami, Yujiro and Masao Kikuchi. 1982. *Asian Village Economy at the Crossroads: An Economic Approach to Institutional Change.* Baltimore: Johns Hopkins University Press.

[23] Mazzocco, Maurizio and Shiv Saini. 2012. "Testing Efficient Risk-Sharing with Heterogeneous Risk Preferences," *American Economic Review* 102(1): 428-468.

[24] Miguel, Edward and Mary Kay Gugerty. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics* 89: 2325-2368.

[25] 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.

[26] 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.

[27] 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.

[28] Munshi, Kaivan. 2011. "Strength in Numbers: Networks as a Solution to Occupational Traps," *Review of Economic Studies* 78:1069-1101.

[29] Munshi, Kaivan and Mark Rosenzweig. 2016. "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review* 106(1): 46-98.

[30] Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*, Cambridge, Massachusetts: Harvard University Press.

[31] Osborne, Martin and Al Slivinski. 1996. "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics* 111(1): 65-96.

[32] Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

[33] Seabright, Paul. 1996. "Accountability and Decentralization in Government: An Incomplete Contracts Model," *European Economic Review* 40(1): 61-89.

[34] Stokes, Susan C. 2009. "Political Clientelism," in eds. Carles Boix and Susan C. Stokes, *The Oxford Handbook of Comparative Politics* Oxford: Oxford University Press.

[35] Tsai, Lily L. 2007. "Solidary Groups, Informal Accountability, and Local Public Goods Provision in Rural China," *American Political Science Review* 101(2): 355-372.

[36] Wade, Robert. 1988. *Village Republics: Economic Conditions for Collective Action in South India*. Cambridge: Cambridge University Press.

[37] World Bank. 2004. *World Development Report 2004: Making Services Work for Poor People*. Washington DC: World Bank.

[38] World Bank. 2016. *Policy Research Report: Making Politics Work for Development*. Washington DC: World bank.

[39] Yadav, Yogendra. 1999. Electoral Politics in the Time of Change: India's Third Electoral System, 1989-99, *Economic and Political Weekly* 34(34/35): 2393-2399.