Clientelism in Indian Villages*

Siwan Anderson, Patrick Francois, and Ashok Kotwal
Vancouver School of Economics

March 2014

Abstract

In many developing countries, institutional set-ups often feature a key group of players, the elite, seeking to manipulate extant institutions to their advantage. Their means of doing this vary across contexts and matter greatly for optimal institutional design and reform. We explore this process for village level governing institutions in India, using a survey that we designed for this end. The region we chose, rural Maharashtra, is known to exhibit functional local democracies, but also shows tremendous government inaction on poverty alleviation; perhaps due to elite control. We find a stunningly robust and participatory democratic process: elections are freely contested, fairly tallied, highly participatory, non-coerced and lead to appointment of representative politicians. However, beneath this veneer of ideal democracy we find evidence of deeply ingrained clientelist vote-trading structures maintained through extra-political means. Elite minorities seek power to undermine policies that would redistribute income towards the majority poor. We explore theoretically the means by which the elite are able to use their dominance of land-ownership and traditional positions of social superiority to achieve political control. Our theory suggests a large set of observables that should covary with the presence of a socially ascendant group (the Maratha caste). Estimated parameters inform us of the precise means by which this region's dominant elite maintain power in light of successful majoritarian institutional reforms.

*This project would not have been possible without the tireless work of Ashiwini Kulkarni and her team in collecting the data on which it is based. We thank seminar audiences at the World Bank, BREAD (NYU), University of Montreal, Queen's University, Bristol University, CEPR Public Economics (LSE), Monash University, and the Canadian Institute for Advanced Research, IOG group for their insightful comments. We have also benefited from detailed discussions with Ashwini Kulkarni, Milind Murugkar, Tim Besley, Gustavo Bobonis, Anil Deolalikar, Thomas Fujiwara , Jim Fearon, Biju Rao, Jim Robinson, Francesco Trebbi, John Hoddinott, Debraj Ray, Nava Ashraf, Robin Burgess, Garance Genicot, Dilip Mookherjee with special thanks to David Green and Paul Schrimpf for extensive comments, and Maja Krzic a Soil Scientist at UBC for guidance in our measurement of land quality. We are also very grateful for the helpful insights from Marianne Bertrand and four anonymous referees.

1 Introduction

There is a fair amount of consensus that successful development depends on the development of the right institutions.[1] By 'Institutions', we mean the rules of the game. Has a particular society been a democracy or a dictatorship? If it has been a democracy, what have been the voting rules? Is there universal franchise or only eligibility for property owners or educated elites?

India has long been characterized by functional and representative democratic political institutions – at federal, state and local levels. But at the same time, these democratically elected governments have been extremely rare champions of the interests of the poor; who are the vast majority of their constituents. Local governance in the Indian state of Maharashtra is a prime example. Though it is a state purported to have active political competition, and thought to feature free and fair elections, it is also a state where local governing bodies – Gram Panchayats – who are mandated and funded to deliver a raft of pro-poor policies, are known to perform this task poorly. The common view of Maharashtrian local politics is that beneath the veneer of representative democracy minority local elites are somehow able to capture majoritarian local institutions and run them in their own interests.

In a representative democracy, with high rates of voter participation, and a vast majority of the electorate exceedingly poor, it is somewhat puzzling that democratically elected governments should so rarely act in the interests of the poor. A possible explanation has been posited by the 'clientelist' hypothesis. Clientelism amounts to the buying of votes and hence power, by a cadre of political elite (patrons) in return for the delivery of direct benefits to the non-elite (clients) whose support is essential for maintenance of power. Elite patrons control government but promote benefits to their clients in a quid pro quo arrangement that may see direct transfers to clients, but that will feature governance largely in the interests of the elite.

Case studies describing this in numerous settings abound.[2] But case studies can only be suggestive of the pervasiveness of clientelist phenomena. Ultimately they leave us with little idea of how widespread it is, nor whether the factors that make it occur in the documented cases also contribute elsewhere.

We undertook an extensive data collection program in rural Maharashtra, India, in an attempt to understand the root causes of local (Gram Panchayat) level misgovernance.[3] Though a long-standing institution in village India, only since 1993 have Gram Panchayats in Maharashtra been responsible for program implementation, local public good provision, implementation of pro-poor policies and been subject to a regular electoral process. Since 1993 these rules have been uniformly applied and well respected throughout the state.

A problem with assessing the clientelism hypothesis is the difficulty of observing it. Poor governance may arise for a number of reasons, and omitted unobserved factors may lead both to local elites running the

[1] Consider, among many others, the seminal study of Acemoglu, Johnson, and Robinson (2001).

[2] Refer to Kitschelt and Wilkinson (2007) for an overview.

[3] Other work has focused on leader characterisitcs of Gram Panchayats (Besley, Pande, and Rao (2012) and Chattopadhyay and Duflo (2004)).

political show and poor governance outcomes, without a causal link. Ideally the researcher would like to identify some source of variation that would allow one to predict when a government is likely to be subject to elite capture, and for this variation to not directly affect governance outcomes except through the channel of capture.

For the case of clientelism, an opportunity presents itself in the Maharashtrian village context. Clientelism is a complicated social undertaking, it depends on an often dense network of interactions between patrons and clients. Patrons must be sure that the clients they deliver benefits to will vote as promised when the election presents itself. Clients must be sure that when they return a patron to power, the patron will deliver the promised benefits. Clientelist structures are generally necessary to make this vote-buying arrangement feasible. They are facilitated by the presence of a traditionally dominant group who naturally play the role of patrons, and who can exploit long-standing social networks to help maintain clientelist undertakings. In Maharashtrian villages such 'natural' patrons are members of the Maratha caste. They are the politically dominant caste within the region, and have been for centuries.[4] The economic elite within a village are the large land-owners and the electoral majority in it are the small holders and landless. We measure Maratha presence in a village on both dimensions: as economic elite through their land-holdings, and as electoral majorities through their population numbers. The literature on Maharashtrian villages identifies two prominent reasons for this caste's continued dominance of local politics; superior within group social-cohesion, and superior within caste trading networks. We embed these two reasons into a simple model of village governance that we use to predict when clientelism is likely to arise as a function of the variation in land holding and population numbers of this caste.

Our model considers the incentives of landlords (the elite) to obtain power, the incentives of workers (the majority) to cede power in return for benefits, and the instruments available to enact clientelist vote-trading transactions. The model suggests a large set of observables that should co-vary with the presence of the Maratha caste if clientelist vote-trading is at work. We map from this model to an estimating equation, the coefficients of which are interpretable directly in terms of the model's parameters. The model thus places a set of sign restrictions on these estimated parameters, and moreover allows us to explore the relative importance of long-standing postulates for the continued hegemony of the region's dominant (Maratha) caste.

In a broad sense, our paper is close to the work of Acemoglu and Robinson (2008). They have as their central motivation, analyzing how political institutions influence economic outcomes and distribution. At a simple level, democracies should favor 'citizens', and dictatorships the 'elite'. But there needs to be a clear distinction between 'de jure' and 'de facto' political power. In the course of history, a society may

[4]Maharashtra is almost unique in the Indian context by the degree to which its politics has been dominated by a single caste (Palshikar and Deshpande 1999, Vora 1996, Palshikar (2007)); the Marathas. They are an intermediate ranking group (sub-caste (or *jati*), traditionally from a Warrior Caste), that are both the most populous, and the largest land owning caste in the state. It is impossible to understand Indian village politics without taking some account of caste (Munshi and Rosenzweig (2009), Banerjee and Pande (2007)).

move from being a dictatorship to a democracy (a change in de jure political power) but the elite may take actions to neutralize this change by building their de facto power, using it to their advantage, and in turn perhaps retarding development in the process. Similar forces highlighted in their general treatment of this phenomenon are at play in our context. A subtle perversion of democracy, leveraging existing social and economic hierarchies, can explain persistence of elite (minority) control despite the implementation of democratic (majoritarian) structures.

Bardhan and Mookherjee (2011) present a model of 'political clientelism' that characterizes democracies in developing societies and how it differs from the phenomenon of 'elite capture'.[5] The clientelism they analyze has a few things in common with what we observe in Maharashtra, but also some marked differences which we will argue are explicable with our model given the institutional context.[6] Other papers that have explored evidence of clientelistic politics are Wantchekon (2003), Wantchekon and Fujiwara (2014), and Vicente and Wantchekon (2009) in West Africa. Finan and Shechter (2012) demonstrate how vote-buying can be sustained by an internalized norm of reciprocity using data from Paraguay.[7] In Maharashtrian villages, cases of clientelism have also been documented by Vora (1996).

Our identification strategy here relies on land holding leading to political leverage and so resembles Baland and Robinson (2008). In their analysis of Chile, landowners "bought" the support of their workers, and owning lands meant owning votes. The quid pro quo arrangement was higher wages in return for votes. Our model predicts the opposite for Maharashtra where income security seems paramount for workers. In our context, the clientelist undertaking takes the form of insurance for workers in return for their support of landlord candidates. In fact, a primary reason we identify for landlords to control governance is to thwart implementation of centrally mandated initiatives that would raise wages at the village level. We predict clientelism leading to fewer programs, more insurance, but lower wages when vote trading occurs - this is precisely what we find.

An interesting finding is that the gratitude the clients feel toward their patrons is so internalized that answers to social capital questions in our survey give a very positive picture of the social relations in villages where the historical elite dominate. A recent paper by Acemoglu, Reed, and Robinson (2013) examines the role of chieftaincy in Sierra Leone. They find that villages with fewer ruling families and more powerful chiefs have lower economic development but higher social capital. They conjecture this reflects the capture of civil society organizations by chiefs whose authority is highly respected because people rely on them for patronage. Our evidence is consistent with a similar story occurring within rural Maharashtrian villages.

[5]Robinson and Verdier (2013) also provide a theory of clientilsm, wherein the two sided problems of enforcement explain why the redistribution often takes the form of public sector employment rather than income transfers.

[6]Their model is based on the political process in the state of West Bengal where the ruling party used this sort of clientelism to win elections; the caste of the patrons plays no role there and the vote buying is also accomplished by using government resources. Other work of theirs, Bardhan and Mookherjee (2000, 2005, 2006), analyses whether a move toward decentralized governance in India has been effective in delivering government services and poverty alleviation schemes to the poor.

[7]Acemoglu et al. (2013) find some muted effects of democratic structures on economic outcomes that are consistent with a type of elite capture.

Our paper proceeds as follows. We start with a description of our context and the main hypotheses guiding both our modeling and data analysis. In Section 3 we develop the theoretical model that we use to determine our main estimating equations. Section 4 provides the empirical results and their interpretation. Section 5 considers alternative explanations and Section 6 concludes.

2 The Context

From November 2006 to May 2007, we surveyed 9132 households from a sample of 320 villages in the state of Maharashtra, which is located on the west coast of central India. Our data are from three main regions: Western Maharashtra, Marathwada, and Vidarbha (we excluded only the Konkan coastal region whose economic hub is Mumbai). To focus on villages which are primarily agricultural (as opposed to factory based or small market towns), which are large enough to generally have their own Gram Panchayat, and where society is caste based, rather than tribal, our criteria for village selection was a total population of 1500-2500 with a tribal population representing less than 10%.[8] From the universe of such villages within our geographic area (a total of 22 565) 320 were randomly chosen and visited by our enumeration teams. Within the villages, neighbourhoods were identified and their approximate population shares computed. Surveying intensity within a neighbourhood was proportional to its population share and households within neighbourhoods were randomly selected. Our sample ends up extremely poor; 42% are below the state poverty line (less than $1.25 ppp/day/capita).

We administered questionnaires at the household level, village level, and to the Gram Panchayats (GPs from now on) directly. Some information, particularly the balance sheets of the GPs, were accessed from higher level state government offices using the "Right to Information Act". In Maharashtra, a given GP typically covers a population of approximately 2000. As a result, in our data the GPs are generally village specific.

GPs implement centrally funded poverty alleviation programs, provide some public goods, represent village interests to higher level administrative units, and obtain resources from centralized funds for village projects. There is substantial variation in all of these performance indicators across our sample of villages.[9]

An important GP activity is pro-poor policy delivery; supposed to be available in the full universe of our sample. There are programs directly targeted to individuals below the poverty line (BPL). There are also non-targeted programs that are still primarily intensively utilized by the poor but nominally available to all residents. The mean number of programs available in a village is 5.33 out of a possible 19 major programs that we asked about, and when restricted to those directly targeted to BPL individuals it is 1.71 out of a total of 8 (refer to Table A1 in Appendix A). Another important pro-poor policy is the state's Employment Guarantee Scheme (EGS), the precursor to the federal government's National Rural Employment Guarantee

[8]Indigenous Tribal society exists in a somewhat parallel relationship to the Caste system in India. It differs markedly in social organization, and will not be amenable to the identification procedures we use here.

[9]Refer to Table A1 in Appendix A for a summary of these outcomes.

Act (NREGA).[10] The EGS is a legal guarantee for 365 days of employment to adult members of rural households willing to do public work-related unskilled manual labour at the statutory minimum wage. To operate in a village, EGS projects must be activated by the GP from a set of possible projects, after petitioning for particular project approval from a higher level authority. The scheme is evident in only 20% of villages. This scheme, like all listed programs, is funded externally and administered by the GP upon request for implementation. The GP draws up lists of eligible recipients, and disburses entitlements to them.

The dominance of the Maratha caste in Maharashtra in terms of land control, political alliances, and rural networks of power has been well documented by political scientists (Deshpande 2004). From the Village questionnaires we obtained: (1) Maratha population numbers; and (2) Maratha landholdings, both at the village level. As will be clear subsequently, these variables will play a key role in our identification strategy.

As seen in Table A4 in Appendix A, Marathas are the main landowners and the economically dominant class. This pattern of relative economic advancement occurs within villages where land ownership is dominated by Marathas (what we call *Maratha Land Dominated* from hereon) and also over the sample as a whole. Maratha dominance is seen in village politics too. Taking into account reserved positions for the Pradhan, the leader of the GP, (that if applied, always exclude a Maratha Male from standing), Table A5 in Appendix A demonstrates that though Marathas comprise about 40% of the population, they are the Pradhan in over 60% of villages where a Maratha can stand. In *Maratha Land Dominated* villages, an unreserved Pradhan is more than 80% likely to be a Maratha. Even when non-Marathas are the village majority, a Maratha is village Pradhan in almost two thirds of cases. This over-representation of Marathas is even more pronounced where positions are reserved for women – reaching 78% even when the village is majority non-Maratha.

All of our villages contain a well recognized economic elite – large land-holders.[11] Large land-holders are relatively wealthy, they own the most important agricultural asset (land), control key employment opportunities, and trade in the most important village goods (agricultural output).

But the elite are a tiny electoral minority, as there are extraordinarily high rates of political participation. Over 89% of eligible individuals voted in the last GP elections in all the major caste groups (Refer to Table A3 in Appendix A). The main reason for the approximate 10% who did not vote is that they were in villages where candidates stood unopposed. Almost no one reported being forced to vote (less than 0.2 of one per cent for any caste), over 95% had met their Pradhan (GP head), and nearly everyone felt comfortable raising concerns directly with their Pradhan.[12] Approximately 83% of our sample rank the Pradhan medium to high (4-5 on a scale of 1-5) in terms of honesty and fairness, and roughly 85% of individuals feel that the representatives of the GP have the most support in the village.

[10] "The EGS in Maharashtra India is the most famous and most successful direct governmental effort at reducing absolute poverty in rural areas...." Ravallion, Datt, Chaudhuri (1991).

[11] There simply do not exist villages with equally distributed medium sized land-holdings in our sample, and this is typical of India as a whole. Exceptions are the Tribal villages which are numerically small, and which we have purposefully not sampled.

[12] These are documented more explicitly in Table A2 in Appendix A. Voters elect the council members of the GP, which then elects among its members a Pradhan (leader). The Pradhan is the only member of the GP with a full-time appointment.

There is little direct value of political control to the elite since most of the resources that come to a village through the GP are tied to recipients. But land-owning elites may still be concerned about schemes that target the poor. The single greatest expenditure category for large landowners is labor. Keeping labour's costs low and maintaining a compliant work force is of great importance. Laborers without access to government programs, or employment opportunities outside the village, are more likely to comprise such a work force.

In summary, we observe: 1. High electoral turnout, accountable political leaders, and a strong majority of poor voters. 2. Weak local provision of centrally funded pro-poor programs. 3. Land-owning elites who would prefer to not have pro-poor policies in place. This brings us to our clientelism hypothesis.

2.1 Clientelism Hypothesis

Workers always comprise a majority and GPs are accountable and participatory local democracies.[13] Why would this majority be willing to give up access to centrally provided benefits, and the employment generation scheme that they could secure with GP effort? We conjecture an explanation based on Scott's (1979) classic analysis: Large landowners seek political power because they can use it to undermine implementation of programs that directly benefit the poor. This provides no direct benefits to the landlords, but keeps labor compliant and cheap. Landlords are never a majority, and democracy is functional in these villages, so landlords attain power only with support of the poor. The poor, on the other hand, would like to have the programs. The cheapest way for landlords to buy their support and thus gain control of village politics is by providing the poor majority with insurance guarantees. The poor understand that a consequence of ceding political control is losing programs and the EGS. Moreover wages may be lower than they otherwise would have been, but the insurance they get from the landlords is the price they extract for this.

2.2 Empirical Strategy

We have information about programs implemented in the village, resources available to the GP, wages, yields, and profits. We also know the villagers' views of how the GP operates, program availability and the presence of the EGS. We further asked who villagers receive help from in times of need. We will thus be able to observe the correlations between these variables. But showing correlations amongst these variables consistent with our hypothesized clientelism will not be enough to prove it. These correlations may be a consequence of omitted factors. Moreover the variables that form the clientelism conjecture are all endogenous to each other.

To get around these problems we exploit village level information we have about the politically dominant upper caste, Marathas. We know whether the largest landholding group in the village is Maratha, we also know Maratha population numbers. The relevance of population and land ownership variables is in how

[13]By "workers" we mean agricultural laborers as well as small cultivators whose main source of income is derived from wage labor.

they affect feasible clientelist structures. Clientelist vote-trading is organizationally complex. Politicians, or their functionaries, who receive votes for promised benefits, are not obliged to deliver the benefits once in office. Voters may not want to vote as they have promised if they have already received benefits.[14] Central to our identification strategy is the advantage that Maratha landlords have in managing these clientelist vote trading opportunities over landlords from lower castes. After a brief discussion of the reasons for this Maratha political advantage in the subsequent section, we construct a model of clientelist vote-trading.[15] This model builds in two potential contributing factors to Maratha landlord advantage, and shows how with them at play we can use the observed village level population distribution of the Maratha caste, and the variation in village landholdings by Marathas, to map to policy, insurance transfers, and economic outcomes implied by the conjectured clientelism. If the model is correct, information on Maratha landholding and population frequencies allows us to predict village level outcomes. For this variation to identify clientelism we need a number of independence restrictions to be satisfied which are made clear in the model development. But prior to even considering these, we briefly discuss the history of Maharashtra's settlement patterns to understand where this variation comes from.

At least since the fourteenth century Marathas have been the dominant land owners in Maharashstra, owing to their hegemony as a military caste. Overall prevalence as landowners persists today but we also see village-level variation in its distribution. This is because of legislated large scale land reforms which were enacted after national independence in the 1950s. These acts, diligently implemented in the state of Maharashtra, effectively redistributed land from the large holders to their former permanent tenants ("Other Backward Castes" or OBCs under today's classification) leading to a dramatic change in ownership (but not cultivation) patterns. In villages where large landowning Marathas were mostly absentee landlords, the dominant land-owning caste today *can* be a low caste (OBCs, former tenants). In villages where Maratha landlords resided, although the lower castes typically also own some land, Marathas are highly likely to still constitute the dominant landowning caste.[16]

One implication of this history is a potentially direct correlation between agricultural productivity and

[14]Kitschelt and Wilkinson (2007) discuss at length the central impediment to clientelist vote trading which is the incentive compatibility of such agreements. Voter moral hazard is compounded with a secret ballot, mandatory in all our villages, so that monitoring individual votes is difficult. But we make little of this side of the problem in the present paper. As Bardhan and Mookherjee (2011) note, successful clientelist politicians build surveillance and enforcement structures. In Maharashtrian villages, a great advantage is provided by the ward based system of voting, with directly elected community representatives drawn from small clusters of households. With the right sort of collective organization, clientelist political transactions are feasible even under anonymous balloting. In our villages, there are about 5 to 6 wards in a GP and each one comprises on average 300 to 400 individuals which is about 70 to 80 households. Each ward elects two representatives. This implies that, at most, 50 households should be enough to deliver a seat on the GP, implying a not implausible level of monitoring by patrons of clients votes undert vote trading. Organized voting is common in the Indian context, and numerous schemes have been devised to circumvent the anonymity of secret balloting, see Chandra (2004) and Subramanian (1999) for in depth analysis and examples.

[15]We are not the first to attempt a measure of elite capture through an indirect strategy. Another way is to use underlying socio-economic inequality as an indicator for locales likely to be subject to elite capture, which is then related to the allocation of public services across socio-economic classes or corruption among elected officials (Bardhan and Mookherjee 2011).

[16]Refer to Appendix D for historical evidence that that our key measures Maratha dominance are historically pre-determined.

Maratha prevalence. Marathas – as the militarily dominant group – may have resided where the highest quality lands were found, choosing to maintain indirect control as absentee landlords over the rest. To control for this possibility, we measured an extensive set of village geographic variables using the FAO-UNESCO soil maps, using GPS data matched to the 2001 census of India, and using information obtained from our own village surveys. Table 1 reports the averages of these variables across Maratha Land Dominated (MLD) and non-MLD villages demonstrating no significant differences on any dimensions across the two. Population sizes and proportions of scheduled castes also do not differ. [****Insert Table 1****]

With land quality correctly measured, the Maratha variables (landholdings and population numbers) should be exerting no additional effect on productivity through this direct channel. We recognize, however, that despite these controls, omitted factors might still be at play, and we return to an extensive discussion of the main potential ones in Section 5 after our main results. What will be clear is it is extremely unlikely that alternative channels of omitted influence can explain the patterns we find in the data.

2.3 Sources of Maratha Advantage

The sociological and political science literature on rural Maharashtra suggests two prominent reasons why Maratha landlords may be more effective than other caste groups at establishing clientelist politics in the villages they dominate.[17]

2.3.1 Superior Social Cohesion

The fact of social cohesion being present between same caste members is not disputable. But it has additionally been argued that Marathas may be better at achieving such social cohesion today due to their greater experience of collective social organization. Carter (1974) describes a history of collective political deals between Marathas in Maharashtrian rural politics. Conceptually, superior social cohesion could sustain enhanced cooperation if a Maratha cheating another (in a political or other form of transaction) suffers a higher cost than a non-Maratha suffers cheating his own caste mate. Such costs are easier to impose in a socially cohesive group. For example, they could involve withdrawing/reducing or excluding violators from social exchanges (marriages, festivals, celebrations). The greater the costs, the easier for Marathas to sustain cooperative outcomes. This hypothesis is in the spirit of Munshi and Rosenzweig (2008), who postulate that the internal disciplining mechanisms within caste groups can act as an effective check on politician misbehaviour. It adds to that basic insight, which is generally true for all caste groups, an added weight for Marathas, for whom these links and internal disciplining mechanisms are posited to be stronger than in other castes. This is consistent with recent experimental evidence on within caste-group punishment obtained by Hoff, Kshetramade and Fehr (2012). In their experiment, conducted between members of high and low castes in Uttar Pradesh (North India), high caste members were found to be systematically more willing to impose costly sanctions on norm violators than were the low castes. Marathas are almost always

[17] Refer to Carter (1974), Lele (1981), and Sirsikar (1970).

the highest ranked caste present in our villages, suggesting a possible advantage vis-a-vis the others. From hereon, we refer to this as the "superior social cohesion" explanation for Maratha political dominance.

2.3.2 Maratha Trading Networks

Another potential underpinning of Maratha power advantages is from their unique system of caste-based trading networks in the state. Rosenthal (1977) describes how a small producer is typically at the mercy of Maratha agents with substantial commercial ties across rural areas. Maratha trading networks deal in seeds, fertilizers, credit and agricultural output marketing. Most agricultural and credit cooperative institutions are either owned or controlled by Marathas (Palshikar 2007). Almost all agricultural transactions in the rural parts of Maharashtra, and in our data, are conducted through either a Maratha trading network or using members of the traditional itinerant trading caste (the Marwaris). Since Marwaris are not resident in villages (there are almost none living in our sample) there are no caste based connections between farmers and Marwaris in our sample. A potential explanation for Maratha landlords' political power is that access to, and use of, these networks is a benefit that Maratha patrons grant to political clients in return for political support. This would make support relatively cheap for Marathas to 'buy'.

3 The Model

3.1 Formalities

Workers, denoted W, own negligible land and sell labor. Landlords, denoted L, hire labor and derive income from landholding. There are $2n$ workers in each village and $1 << n$ landlords. Workers are a majority, so landlord political control can only occur if they 'buy' at least n worker votes. Each individual has a caste (denoted c_i for person i). Either $c_i = M$, or $c_i = N$, denoting Maratha and Non-Maratha respectively. Each agent is identified by both class (W, L) and caste (M, N).

Programs

The GP is tasked with implementing pro-poor policies. If implemented correctly, they generate worker value P. If implemented badly, they generate $\tilde{P} < P$. Landlords do not value programs. Reducing program availability lowers wages and raises labor compliance.[18] Denote wages when programs are implemented by w_P, and those when not by $w_{\tilde{P}} < w_P$. The full benefit to a worker from programs in a village is denoted $\Delta_{wP} \equiv P - \tilde{P} + w_P - w_{\tilde{P}}$. Finally, denote land rents when programs are implemented by π_P and those when not by $\pi_{\tilde{P}} > \pi_P$.

Maratha advantages

Maratha trading network access generates benefit of amount T for both workers and landlords. Only Marathas have access directly, but any Maratha can grant it to a non-Maratha costlessly. Workers utilize these networks to procure inputs for their own small plots, loans for business activities, and for sale of

[18] This can be micro-founded (see Appendix C) but is an assumption here.

output. The majority of workers (roughly 70%) in our sample live in households with a small amount of land or running a small business.

Landlords divide vote buying responsibilities symmetrically. Since landlords are of measure 1 each is responsible for the votes of n workers. Landlords have incentive to free-ride on the vote-buying of their colleagues. To overcome this, landlords impose social punishments on individuals who cheat. Marathas potentially have advantages in the strength of these punishments. Let X_M denote the social punishments imposed on a cheating Maratha landlord, by other Marathas. Let X_N denote a Non-Maratha landlord group's analogous punishments with $X_M \geq X_N \geq 0$. Additionally, Maratha workers are unique in being able to impose social punishments on Maratha landlords who cheat them, denoted by $X \geq 0$.[19]

Insurance

Votes are bought by promising insurance transfers in a state of 'need'. There are two possible states; a normal state with consumption valued at 1, and a state of need where a worker's marginal valuation of consumption is $\phi > 1$; examples are medical expenses, loss or damage to a household asset such as livestock, employment/sickness shocks to an earner etc. The need state arises with probability μ. It is drawn each period independently from a distribution common to all workers. An insurance promise from a landlord, i, to a worker, j, is a commitment by the landlord to a transfer of S_i^j in the worker's need state. Such a state is observable to both landlords and workers but not enforceable by formal/legal mechanisms. For simplicity we assume from now on that $\mu \cdot \phi = 1$.[20]

Incentive compatibility of insurance promises

The timing is as follows: (1) Worker and Landlord can strike a vote trading deal specifying a transfer S_i^j from landlord to worker in state of need in return for the worker's vote. (2) The state is revealed to both parties. (3) The landlord chooses the transfer level if the need state arises. (4) Elections occur. If the need state arose and the transfer received is (at least) S_i^j, the worker votes for the landlord's candidate. If the need state arose and the transfer received is less than S_i^j, the worker votes for someone else, and social sanctions are imposed on the landlord. If the need state does not arise, the worker votes as promised.[21]

In addition to social costs common to all individuals we allow for each pair to share a common idiosyncratic history (not observed by the researcher) which may allow some level of social cost to be imposed on the landlord if he cheats the worker. This pair's history is summarized by a non-negative number, x_i^j, which

[19] Only Marathas are able to impose these as they are the only *Jati* (sub-caste) on both sides of a clientelist agreement – i.e., Marathas can be both landlords and workers in the same village. Occasional exceptions will arise in non-Maratha landlord dominated villages, but these are rare and are ignored here.

[20] Relaxing this will imply that the parameters ϕ and μ enter into the structural interpretation of the estimated coefficients. As will be seen, we restrict intepretation of coefficients in terms of model parameters to sign implications which are invariant to this normalization.

[21] We only focus on the incentive compatibility of promises made by landlords to workers in return for their votes. A more complicated version of the model would also analyze the incentive compatibility of worker promises to vote in favour of the landlord's candidate after having received insurance. Since any patron group must assure clients vote as promised, this is a problem that is common and for which no caste has a particular advantage. In the model, we thus focus on the landlord's (patron's) side of the problem, where Maratha landlord advantages are more likely to arise.

is randomly and independently drawn from the distribution $F(x_i^j)$ defined on a finite support. The following assumption plays a key role in identifying the model:

Assumption 1: *The pairwise idiosyncratic terms, x_i^j, are independent of village level landholdings by caste, and caste population numbers.*

Purely at the individual level, this is a reasonable assumption. The working relationship between a landlord and his employees is likely to reflect their personal history which is unlikely to be systematically affected by village level variables such as the caste frequencies or aggregate landholding patterns. But it is possible that more than the pair's personal history could be involved. Coordination amongst workers to punish landlords who transgress could possibly be affected by village caste frequencies. For instance, worker level social capital or social cohesion may vary with the share of land held by Marathas. This would make it more difficult for workers to punish landlords in MLD villages. Since our surveys asked respondents about social cohesion, we are able to verify in Table B7 of Appendix B that such an effect is not evident; our measures of worker social cohesion do not vary across MLD and non-MLD villages.[22] We also demonstrate, in Appendix C section 9.4, that all of our model results will continue to hold under this violation of assumption 1.

Additionally, the cheating landlord incurs social costs imposed by other landlords for not delivering his n votes; the X_i defined above. Both of these punishments bound the transfers that a landlord can credibly promise in return for the worker's support in a clientelist arrangement. Specifically, incentive compatible insurance transfers between landlord i and worker j must satisfy:

$$S_i^j \leq X_i + I_i^j X + x_i^j. \tag{1}$$

where $I_i^j = 1$ if $c_i = M$ and $c_j = M$, and $I_i^j = 0$ otherwise.

The cost of a vote

Individual rationality of vote trading for landlord i buying the votes of n workers j requires:

$$\pi_{\tilde{P}} - nS_i^j \geq \pi_P. \tag{2}$$

For the worker (j), similarly, individual rationality of vote trading with landlord (i) in village k requires: $U_{jk}(L_i) \geq U_{jk}(W)$, where U denotes the expected utility outcome corresponding to the group in parentheses controlling the GP. Without vote trading, workers (W) are the majority and control the GP. Consequently programs are implemented at value P, there is no insurance, and if the worker is not Maratha, there is no trading network access. In that case worker utility is:

$$U_{jk}(W) = w_P + P + d^j T \tag{3}$$

where $d^j = 1$ if $c_j = M$, and $d^j = 0$ if $c_j = N$. In contrast, with vote trading, landlords (L_i) control the GP. Programs are reduced to $\tilde{P}$ but there is insurance provided and network benefits arise for workers, so

[22] We show that the landless are not more likely to share information with other workers or trust people like themselves in MLD villages.

that only if both parties are non-Maratha does a worker *not* receive network benefits in this case; $d_i^j = 0$. An additional cost of vote trading to workers is that they must support the landlord's candidate instead of their own preferred one. We capture the cost to the workers of village k doing this via the term x_k. A higher quality landlord candidate increases x_k. The variable x_k is drawn from distribution $G(x_k)$.

$$U_{jk}(L_i) = w_{\tilde{P}} + \tilde{P} + S_i^j + d_i^j T + x_k \tag{4}$$

Once again, model identification depends on the draw of candidate specific quality being independent of village level caste characteristics. Similarly, we assume:

Assumption 2: *The realization of x_k is independent of village level landholdings by caste, and caste population numbers.*

This amounts to assuming that the random variation in the quality of landlord candidates across villages is not related to landlord or population caste numbers. An indirect indication can be obtained by checking for systematic differences in worker perceptions of leader quality in MLD and non-MLD villages. Table B7 of Appendix B shows that in terms of the perceptions that workers have of panchayat leaders in MLD villages versus non-MLD villages, there do not seem to be systematic differences. Though this is not an aspect directly addressed by the literature on Maharashtrian villages, numerous contributors to this literature have argued that the history of domination by the Marathas has made them resented by lower castes and the poor (Mandavdhare 1989). This would suggest that if Assumption 2 is violated at all it would likely be in the direction of landlord candidates in MLD villages being seen less favorably by workers than such candidates in non-MLD villages. If so, the x_k term in Maratha land dominated (MLD) villages would be drawn from a distribution that is stochastically dominated by that in non-Maratha land dominated villages; clearly violating Assumption 2. Since this cannot be ruled out, we explore the implications of this departure from Assumption 2 in Appendix C section 9.5. This also has no effect on model results.

We now establish incentives and feasibility of clientelist vote trading for a given worker/landlord pair:

Proposition 1 *Clientelist vote trading is both individually rational and incentive compatible for a worker (j)/landlord (i) pair if and only if:*

$$x_i^j + x_k \geq \begin{cases} \Delta_{wP} - X_M - X \equiv x_{MM}, \text{ for } c_i = M \text{ and } c_j = M. & (5) \\ \Delta_{wP} - X_M - T \equiv x_{NM}, \text{ for } c_i = M \text{ and } c_j = N. & (6) \\ \Delta_{wP} - X_N \equiv x_{NN} \equiv x_{MN}, \text{ for } c_i = N \text{ and } c_j = M \text{ or } N. & (7) \end{cases}$$

Proofs of all propositions are in Appendix C.

A high value of x_k makes it less costly to workers to vote for the landlord's preferred candidate, and individual rationality of vote trading easier to satisfy. A high value of x_i^j makes landlord reneging on promised worker transfers costly, and hence supports a greater range of incentive compatible transfers from them in return for workers' votes. These terms do not vary by caste but the right hand side of this expression

does. When both worker and landlord are Maratha, equation (5), superior caste cohesion sustains higher punishments, X, and hence makes higher transfers incentive compatible. This adds to the within landlord punishment effect X_M. When only landlords are Maratha, (6), the X term disappears to be replaced by trading network access, T, which can be granted to non-Maratha workers for their support; hence benefiting vote trading. In the final case – that of non-Maratha landlords in equation (7) – there is neither trading network access, nor superior caste cohesion between workers and landlord, but non-Maratha landlords can still punish landlord digressions, X_N.

3.2 Computing the likelihood of vote trading

To move from individual level vote trading incentives to village level predictions, we now consider the relationship between these individual conditions and the likelihood of vote trading occurring in the village as a whole. We already know that no single landowner acting as a patron can orchestrate control of the GP by vote trading with his own workers alone. The landowners as a group must be able to contract votes from enough workers to ensure a majority in the village. We make the following assumption about how many:

Assumption 3: *If and only if a majority of workers find it individually rational to accept incentive compatible transfers from landlords for vote trading, then vote trading occurs, and landlords exert political control.*

This assumption allows us to map from the incentive compatibility/individual rationality conditions for worker/landlord pairs, and the frequency of such pairs to vote trading at the village level. It amounts to assuming that landlords have the wherewithal to act in their collective interests; if there are gains to be made from vote trading, we assume vote trading occurs. If, however, the votes that can be feasibly bought by landlord patrons are not sufficient for them to wrest control of the GP, they do not bother.

3.2.1 Deriving the estimating equation

The model predicts outcome variables, denoted v_k, – wages, profits, programs, insurance – that will be affected by clientelism occurring in village k; these will be discussed further in Section 3.4. These variables are also potentially affected by the village level co-variates listed in Table 1. So we include such controls, as well as individual level controls for education, land ownership, and caste identity.[23] We denote these by the vector $\mathbf{Z_k}$. Thus for an outcome variable v_k, in village k we have:

$$v_k = \alpha_v I_{VTk} + \alpha \mathbf{Z_k} + \mu_{vk}, \tag{8}$$

where $I_{VTk} = 1$ if vote trading occurs in village k and 0 otherwise, and μ_{vk} is a mean zero village and variable specific error term. By now assuming an explicit process for the idiosyncratic shock term, x_k, we are able to map population frequencies by caste in to the proportion of likely vote traders, hence the likelihood of clientelism occurring, and in turn to the set of outcome variables in the village. Let σ_{jik} denote the frequency of worker caste j, landlord caste i pairs in village k. Then:

[23] In Appendix B we show that all results are robust to omiting these controls.

Proposition 2 *If x_k is drawn from a uniform distribution, the expected value of variable v_k in a village with population frequencies σ_{ijk} and characteristics $\mathbf{Z_k}$ is given by:*

$$E[v_k|\sigma_{ijk}, \mathbf{Z_k}] = \alpha_v[C + \sigma_{MMk}(X_M + X) + \sigma_{NMk}(X_M + T) + \sigma_{NNk}(X_N) + \sigma_{MNk}(X_N)] + \alpha\mathbf{Z_k}, \quad (9)$$

with C a constant.

The assumption of uniformity on the x_ks makes for an extremely simple relationship between the probability of clientelism occurring and the caste landlord/worker frequencies. In Appendix C (Section 9.3) we show that the interpretation of coefficients and the inferences drawn from the model are identical under any well-behaved alternative distributions satisfying Assumption 2.

Intuitively, the probability that a high enough proportion of the village are willing vote traders – and hence that clientelism occurs – depends on the frequency of worker/landlord pairs by caste type (the $\sigma_{ji}s$) because each pairing differs in its propensity to vote trade depending on its caste composition via Proposition 1. We compute the σ_{ji}s from land ownership and population frequencies by caste under the following assumption.

Assumption 4: *Political clients are randomly matched with political patrons.*

Assumption 4 implies that the probability of a particular clientelist pair occurring in a village will be directly related to the population frequencies of such pairs in the overall village population; caste members are neither more nor less likely to be matched in landlord/worker pairs. Random matching may seem a strong (and unlikely) restriction in India where segregation by caste can occur. We show, in Appendix C Section 9.6, that the perhaps more natural "positive assortative matching" by caste case does not alter the model interpretation of coefficients.

We can use our data to estimate the σ_{ji}s directly. We know the population distribution of each caste group in the village, and from this we construct the continuous variable, $MPROP \in [0, 1]$, which equals the proportion of the village population that is from the Maratha caste. We also have from the Talathi (village administrator) the dominant land owning caste group in the village. We code this as the binary variable, MLD, which equals 1 if Marathas are the Land Dominant group, and 0 otherwise.[24] Using these, we can obtain a relationship between observables as follows:

Proposition 3 *Under Assumption 4, expression (9) re-arranges to*

$$\begin{aligned} E[v_k|MLD_k, MPROP_k, \mathbf{Z_k}] &= \alpha_v C + MLD_k \cdot \alpha_v [T + X_M - X_N] + MPROP_k \cdot \alpha_v [0] \\ &\quad + MLD_k \cdot MPROP_k \cdot \alpha_v [X - T] + \alpha\mathbf{Z_k} + \mu_{vk}, \end{aligned} \quad (10)$$

where C is a constant.

[24] We have an alternative source of Maratha land holdings from our household surveys. From these 30 households per village we obtain an estimate of the overall proportion of village lands held by Marathas $\widetilde{MLD} \in [0, 1]$. We utilize the Talathi variable throughout the body of the paper as it is not subject to sampling error and hence more accurate, but show that all results are robust to instead using $\widetilde{MLD}$ in Appendix B.

3.3 Predicted Outcome Variables

Equations (3) and (4) generate the set of workers' dependent variables predicted to be affected by clientelist vote trading. For each one of the dependent variables (v) below we elicit its directional change by comparing its model predicted value when vote trading happens (equation (4)) with its value without vote trading (equation (3)). The comparison determines the sign of $\alpha_v$in equation (10).

For worker variables: **1. Programs ($\alpha_v < 0$)**: Without vote trading programs are P, they fall to $\tilde{P}$ under vote trading. **2. Insurance Receipts ($\alpha_v > 0$)**: Without vote trading, the value of insurance received by workers from landlords is 0, under vote trading this rises to S_i^j. **3. Wages ($\alpha_v < 0$)**: Without vote trading wages are w_P, they fall to $w_{\tilde{P}}$ with vote trading. **4. Trading Networks ($\alpha_v > 0$)**: Without vote trading, d^jT are Maratha trading benefits, where $d^j = 1$ if and only if worker $j's$ caste is Maratha. With vote trading, Maratha trading benefits are d_i^jT where $d_i^j = 1$ if either worker $j's$ caste is Maratha or landlord $i's$ caste is Maratha, and $d_i^j = 0$ otherwise.

For landlords the relevant variables are computed using equation (2), the landlords' individual rationality conditions. These are: **1. Profits ($\alpha_v > 0$)**: Without vote trading, the value of profits is π_P, which rises to $\pi_{\tilde{p}}$ under vote trading. **2. Insurance Liabilities ($\alpha_v > 0$)**: Without vote trading, the value of insurance liabilities from landlords is 0; under vote trading this rises to S_i^j.

3.4 Interpretation

We estimate equation (10) using a linear regression of the following form:

$$v_k = \beta_0 + \beta_1 MLD_k + \beta_2 MPROP_k + \beta_3 MLD_k \cdot MPROP_k + \gamma \mathbf{Z_k} + \epsilon_k. \tag{11}$$

Given Assumptions 1 and 2, and that MLD_k and $MPROP_k$ do not directly affect the outcome variables – the implications of which we discuss later – the linear regressions that we estimate will yield consistent estimates. The regression coefficients directly map to the model's parameters as follows:

$\boldsymbol{\beta_1 = [T + X_M - X_N]\alpha_v}$. Since $T > 0$ and $X_M \geq X_N$, the model implies this coefficient should be positive (negative) for $\alpha_v > 0$ (< 0). Maratha land dominance should have a direct positive effect on the probability of vote trading, thus raising (lowering) the value of variables the theory predicts are larger (smaller) under vote trading.

$\boldsymbol{\beta_2 = [0]\alpha_v}$ **.** There should be no direct effect of Maratha population numbers, and hence the prevalence of Maratha workers, on a clientelist vote trading variable. This is because having Maratha versus non-Maratha workers only effects Maratha landlords, and so is picked up via the interaction term. Non-Maratha landlords have the same capacity to enter into vote trading with workers of any caste as the costs imposed on cheating are X_N in both cases.

$\boldsymbol{\beta_3 = [X - T]\alpha_v}$. The sign of this variable tell us the relative importance of the two posited explanations for the political prominence of the Maratha caste; their greater social cohesion or their trading network access respectively. Intuitively, the interaction term applies where Marathas are both the landlords and the workers. In such cases jati benefits, X, affect the feasibility of vote trading but trading network benefits that are of use to Maratha landlords with non-Maratha workers cease to be relevant as Maratha workers already have access directly; hence the negative sign on T. So for $\alpha_v > 0$, if the interaction term enters positively (negatively) $\beta_3 > (<)0$, then the superior caste cohesion hypothesis is more (less) important than the network advantage hypothesis.

$\boldsymbol{\beta_1+\beta_3 = [X_M - X_N+X]\alpha_v}$. This is the magnitude of the total clientelism advantage of Marathas that derives exclusively from their superior social cohesion. The term $X_M - X_N$ reflects the superior punishment capacity on a transgressing landlord by other landlords when Maratha. The term X is the superior punishment of a transgressing landlord by a worker. The summation of these coefficients suggests the total extra help in sustaining clientelism deriving from Maratha social cohesion. If this summation exceeds zero (for $\alpha_v > 0$) then social cohesion generates benefits for Marathas in maintaining clientelist structures. If not, the total benefits of superior social cohesion do not aid Marathas in sustaining clientelist structures.

4 Estimation and Results

Based on (11), we run two main estimating equations depending on whether the dependent variable is village or household level. The household level regression is represented by the following:

$$Y_{ik} = \beta_0 + \beta_1 MLD_k + \beta_2 MPROP_k + \beta_3 MLD_k \cdot MPROP_k + \boldsymbol{\psi} Z_{ik} + \boldsymbol{\gamma} Z_k + \epsilon_{ik}. \tag{12}$$

Y_{ik} is an outcome of household i, residing in village k. Z_{ik} includes household controls (education, land ownership, and caste identity); Z_k includes village level geographic, demographic, and climate controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and reservation status of Gram Pradhan). MLD_k is equal to 1 if a village k is Maratha land dominated and equal to 0 if the village is land dominated by a lower caste. $MPROP_k$ is the proportion of Marathas in a village k.[25] ϵ_{ik} is a regression disturbance term clustered at the village level. Refer to Tables A1, A2, and A3 in Appendix A for summary statistics on all of the variables used in the estimations.

We also explore the impact of Maratha dominance on GP performance using village level outcome measures. Similarly we estimate the following:

$$Y_k = \beta_0 + \beta_1 MLD_k + \beta_2 MPROP_k + \beta_3 MLD_k \cdot MPROP_k + \boldsymbol{\gamma} Z_k + \varepsilon_k. \tag{13}$$

[25] We also include region fixed effects to control for variation in historical land revenue systems, but we do not report these.

Y_k is a village level GP outcome measure in village k and ε_k is a regression disturbance term.

The pairwise correlation between MLD_k and $MPROP_k$is only 0.54, so estimations should not suffer from multicollinearity. Also, the villages where Marathas dominate (by either of these measures) are geographically dispersed throughout our sample. Our results are robust to including regional fixed effects defined in several ways.[26]

MLD_k is equal to one for 59% of the villages in our sample. The remaining villages are dominated by a lower ranked caste group (primarily two cultivating castes, the Kunbis and Dhangars, both OBCs under today's classification). In the empirical estimation we leave these two sub-castes (*or jatis*) grouped together. All of our main results persist if we instead separate them in the analysis.

4.1 Evidence on Model Variables

4.1.1 Worker Cost of Vote Trading: Programs ($\alpha_v < 0$)

We obtain direct information about the availability of mandated and centrally funded programs from our household level surveys and from our survey of the village government. There are two broad types of programs that are relevant. Firstly, there are programs directly targeted to individuals below the poverty line (BPL). There are also non-targeted programs that are still primarily intensively utilized by the poor but nominally available to all village residents. For each one of these programs, household members were surveyed on the availability of the program within the resident village. The model predicts a negative effect of MLD on programs, represented by the coefficient $\beta_1 = \alpha_v[T + X_M - X_N]$; since $T > 0$ and $X_M \geq X_N$. The model also predicts that $MPROP$, represented by β_2 (of (12)), should have no effect on the level of programs once MLD and the interaction between MLD and $MPROP$ are included in the regression.[****Insert Table 2****]

As can be seen from Table 2, the results are consistent with these predictions of the model. Programs (1) and BPL Programs (1), from the household surveys are, on average, 1.01 and 0.42 lower respectively in MLD villages. Both are significantly different from zero at the 1% level. The means for these variables were 5.36 and 1.73 respectively, so the percentage decline in programs is 19% and 24% respectively. Since the interaction term captures the effects of MLD in villages with a positive $MPROP$, these numbers should be interpreted as the percentage decline in the availability of programs due to the effect of MLD in villages where $MPROP$ is negligible. We also aggregated individual level information and ran village level regressions (represented by (13)), Programs (2) and BPL Programs (2). These are almost identical.

Our model predicts that we should see the greatest sensitivity to MLD for those government programs which are expected to impact wages or income directly, as these are the ones landlords should have greatest interest in limiting. To this end, we separated out the set of programs that directly or indirectly affect household labour decisions and income (such as public good and housing improvement schemes which do

[26]That is, at the level relevant to administrative divisions today, at colonial administrative units, or at the district level. Results are also robust to excluding the few areas where no Marathas are present.

rely on villagers' labour in their construction and the targeted public distribution system), from those which would not (such as those targeted towards children (child development and mid-day meals) and the elderly (social security pensions and foodgrain transfers)). We refer to these as "Income Programs" and "Non-Income Programs" respectively. We also asked information from both household and village records on the availability of the state's Employment Guarantee Scheme (EGS), which is the precursor to the present NREGA, in the village. Under the clientelist hypothesis the EGS program and the Income Programs should be negatively related to MLD and unaffected by $MPROP$. Again this is borne out by the data. As we see in Table 2, the EGS is 0.10 times less likely to be reported as existing by households in MLD villages (with a neglible $MPROP$). The mean of this variable is 0.20 over all villages, hence this is a 50% decline in MLD villages. Income Programs are 0.98 lower in MLD villages, the mean of this variable is 4.85, so this represents a 20% decline. As also expected, we find no corresponding effects of MLD on Non-Income Programs.

We further use the GP survey and access to the accounts of the GP to check the consistency of individuals' reports with formal records. Programs are funded by the state government but must be activated by the GP. If a lower level of reported program availability is occurring because the GP is petitioning for fewer recipients, and expending less effort in activating them, then this should imply lower revenue inflows into the GP. Revenue (1) and Revenue (2) are per-capita village revenues from two independent sources. The first is from our own access to the village balance sheets – much of which was obtained from our petitioning of this information through the state's Right to Information Act.[27] As our sample size indicates, despite our petitioning through the act, the books were not available in almost a third of our villages. Revenue (2) is information we obtained from the government's Village Census (2001). Both coefficients are negative and significantly different from zero, suggesting that, indeed, MLD villages have lower revenues. Again $MPROP$ has no significant effect on this as predicted.

Turning to the coefficient on the interaction term, β_3 (in (13)). The model's interpretation of this coefficient is as: $\alpha_v[X-T]$. So its sign indicates the relative importance of the posited underlying sources of Maratha clientelist advantage, X or T. The evidence in Table 2 is supportive of trading networks being more important than superior social cohesion, since $\alpha_v < 0$ for programs and the coefficient on the interaction term is positive in all cases. Specifically Programs (1), BPL Programs (1), Income Programs (1) are positive and significant at the 1% level. EGS (1) is positive and significant at the 5% level. There is again no corresponding significant effect for Non-Income Programs. For Revenue (2) and Expenditure this coefficient is also significant and positive at the 5% level and hence also consistent with trading benefits being more important than superior caste cohesion of Marathas.

Column 4 reports the significance of $\beta_1 + \beta_3$. Recall from Section 3.4 that the model implies $\beta_1 + \beta_3 = \alpha_v[X_M - X_N + X]$ representing the total social cohesion benefit of Marathas. As can be seen from the column,

[27] An act that mandates delivery of the relevant documentation within 30 days of written request, or the imposition of penalties on the GP.

in very few cases is the summation of these coefficients statistically significant. In the few cases that it is, this significance is typically lost in the robustness checks conducted in Appendix B (refer to Section 4.3). Overall it is not possible to reject the hypothesis that, with respect to program availability, superior Maratha social cohesion provides no net benefit to Marathas in sustaining clientelism. This, together with the sign of the coefficient on β_3 points to trading access being more important in explaining Maratha hegemony than superior social cohesion; a finding we will see repeated for the variables that follow.

Finally, we explore the reasons programs and revenues may be affected by these variables using the GP Survey. Specifically, the survey asked about the number of meetings between the village Pradhan and higher levels of government. Since programs are activated by the village government through requests to higher level officials, a more active Pradhan should indicate a village where greater attempts are being made to implement programs; conversely for a village under clientelist vote trading where programs are diminished. The coefficients β_1 and β_2 (from (13)) are consistent with this interpretation. MLD leads to around 3 fewer meetings between the Pradhan and higher level functionaries – the Block Development Officer (BDO), the District Collector (DC), and the Member of Parliament (MP). We also estimated an average effect size (AES) of these three dependent variables following Kling, Liebman, Katz and Sanbonmatsu (2004). The AES estimates confirm the findings on the β_1 and β_2 coefficients when examining the outcomes independently. The interaction term β_3 is positive and significant for meetings with the MP and DC. The latter is the most important contact of the three, as the DC makes the final allocation decision and distributes the government transfers to the GPs. The sign on β_3 is thus also consistent with trading benefits being more important than superior caste cohesion of Marathas; $T > X$.

4.1.2 Landlord Cost and Worker Benefit of Vote Trading: Insurance (S_i^j)

The clientelism hypothesis suggests that the counterpart of reduced programs that provides incentives for workers to partake in vote trading is increased insurance. The term S_i^j should appear when vote trading occurs. Thus insurance related dependent variables should exhibit opposite signs to those predicted for programs: $\beta_1 > 0, \beta_2 = 0$, and $\beta_3 = \alpha_v[X - T]$ in (11).

Information about insurance was obtained from the household surveys. Seven questions were asked in a hypothetical form, describing transfers, in both cash and kind, by individuals of different social status. Specifically: Insured (1): "Would most people in your village help you with some money in times of crisis?"; Insured (2): "Would a higher caste member of your village help you with some money in times of crisis?"; Insured (3): "Would most people in your village help a lower caste villager with some money in times of crisis?"; Insured (4): "Would most people in your village help you with some grain in times of crisis?"; Insured (5): "Would a higher caste member of your village help you with some grain in times of crisis?"; Insured (6): "Would most people in your village help a lower caste villager with some grain in times of crisis?"; Insurer: "Suppose a lower caste man asks to borrow a good sum of money from you because someone in his family has fallen ill. He is from the village and has the ability to repay the amount. Would

you lend it to him?".[**Insert Table 3**]

The results in Table 3 are once more consistent with the sign restrictions of the model. The point estimate of β_1 run for the landless is positive at around 0.15 for each of the first six questions. Implying that landless households are 15% more likely to report provision of help in a state of need in MLD villages. The AES estimate of these six dependent variables confirms the findings on the β_1 coefficient when examining the outcomes independently.

The increased prevalence of Marathas (upper castes) in a village should make the reported provision of insurance from "upper castes" more likely, even if insurance is simply provided at random. This introduces a bias into the estimation of the coefficient β_2. Specifically, clientelist vote-trading predicts the effects of $MPROP$ through clientelist vote trading to be zero, but the regression that we run also picks up a direct effect of $MPROP$ through this mechanical channel. So the structural interpretation of this parameter is confounded in this case due to the direct effect.

The structural interpretation of the coefficient on the interaction term, β_3, however are uncontaminated by the direct effect of $MPROP$. These results are reported in the third column. Since $\beta_3 < 0$, and always significantly so, the structural interpretation is that $X < T$, again implying Maratha trading networks seem to be more important in explaining Maratha landlord advantages than their superior caste cohesion.

Column 4 reports estimates of $\beta_1 + \beta_3$. Here, once again, this linear combination of coefficients is rarely significant. The implication in terms of model parameters is again that $\alpha_v[X_M - X_N + X]$ is not significantly different from zero, and hence that superior social cohesion amongst Marathas does not contribute to their political power.

Considering now the other side of the insurance relationship we again see a similar pattern. Large landowners answer that they are more likely to lend to a hypothetical "lower caste man in need" (Insurer) in MLD villages. This is mitigated by an increase in $MPROP$. We also ran regressions for this insurance liability question (Insurer) over restricted samples: large land holders, Marathas with low and large land ownership, and low castes with large and low land ownership. Again, consistent with the theory, the estimated coefficients are only significant for Marathas with large land holdings. Lower castes with large landholdings and Marathas with small land holdings are *not* more likely to lend to lower castes in MLD villages, nor are they less likely to lend in villages with greater $MPROP$. This is again consistent with the Maratha land owners being the ones who act cohesively in providing insurance, and with that happening in return for vote-trading.

4.1.3 Landlord Benefit/Worker Cost of Vote Trading: Wages ($w(\tilde{P}) - w(P) < 0$)

According to the clientelism hypothesis, vote trading is a means to make labor cheaper, so wages should fall with the probability of clientelist vote trading; ($\alpha_v < 0$), a negative coefficient for β_1. As the first row of estimations in Table 4 documents, $\beta_1 = -2.2$ for the average worker in a village. This implies that the average daily wage of workers in MLD villages with negligible $MPROP$ compared with non-MLD

villages is 2.2 rupees lower, which is significant at the 1% level. This is 6-7% of the average daily wage. According to the model, the lowering of wages through clientelist vote trading should be indiscriminate since it works through a village general equilibrium channel by reducing external labor demand (EGS) as well as lowering labor's opportunity cost (Programs). This is in contrast with other vote trading schemes, for instance the one highlighted in Chile by Baland and Robinson (2008). There, the finding was that wages rose as a reward for political support. Here wages are predicted to fall as a general equilibrium consequence of landlord village control. Consequently, we should observe wage declines across the village as a whole, not just for the workers who report increased insurance benefits. To check for across the board effects, we measure whether lower wages are restricted to workers reporting insurance benefits, low caste workers, low caste and insured, or workers of either gender. The remaining rows in the first panel of Table 4 indicate an effect that is consistent with a general equilibrium channel. The magnitude of wages for all categories of workers falls by about the same amount irrespective of gender, insurance access and caste.[****Insert Table 4****]

The coefficient β_2 is, as expected, again generally not significantly different from zero. The sign of β_3 is in all cases positive, and usually highly significant; again consistent with $T > X$.[28] Finally, the implication of column 4, that $\beta_1 + \beta_3$ is in almost all cases not statistically different from zero, again suggests that the net benefits of Maratha superior social cohesion are likely to be small.

4.1.4 Landlord Benefit of Vote Trading: Yields and Profits ($\pi(\tilde{P}) - \pi(P) > 0$)

We obtain information about yields and profits from surveys of large agricultural producers. We focus on the *kharif* – 4 month wet season – growing period which is the main one in our villages. In our data, kharif crops are grown on 77% of cultivated land, they generate 73% of total annual yields and form 81% of total annual profits. Typical crops include grains and pulses (*tur*, *bajra*, *jowar*, *chana*, soybean, and wheat) as well as cotton. Labour comprises 31% of total kharif input expenses, fertilizer 30%, and seeds 24% (irrigation expenses are neglible as these are rain-fed).

The model predicts increased yields and profits to be the main source of benefit for farmers from vote trading, so $\alpha_v > 0$ and hence $\beta_1 > 0$. As the first row in the second panel of Table 4 indicates, the model's sign restriction on β_1 is again respected by the data. Log yields are higher in MLD villages, and this is significantly different from zero at the 1% level. The coefficient is very large. A MLD village has the log of kharif yields that are 0.32 higher (32% higher) than a non-MLD village (again for a village with negligible $MPROP$). But since the average MLD village has approximately 30% more Marathas than a non-MLD village, and the interaction term is -0.51, the overall rise in yields when moving from MLD to non-MLD villages is: $0.32 - 0.3\text{x}0.51$ or an increase of around 17%. Log profits are 48% higher in such villages, but

[28] Note these findings also strongly suggest that mismeasured land quality is unlikely to be an important contributor to our results. As conjectured earlier, mismeasured land quality would tend to manifest through land being higher quality in Maratha land dominated villages than we have controlled for. If this is the case, it makes the lower wages found in such villages even more difficult to understand.

similarly adjusting for the higher $MPROP$ in such villages yields 17% higher profits.

Since we do not have information on the hourly wages paid to workers by specific farmers, nor how many hours worked, we can only use the information on hired and total workers, for a given farmer, to estimate the channels through which yields and profits may be higher in MLD villages. Paid workers, and total workers have higher point estimates, but are not significant statistically. If worker compliance is the channel of effect, we should expect that clientelist vote trading will be associated with a lower proportion of total input expenditures going to labor. This implies a negative prediction of the model on β_1 for this variable. The ratio of labor costs to total costs, the last row in the middle panel of Table 4, is lower in MLD villages, suggesting some contribution to increased profits could be arising through lower wages. But the effect here is statistically weak; significant only at the 10% level.

The coefficient on β_2 is again not different from zero; as expected. The negative sign of β_3 again suggests a stronger contribution of trading networks than superior caste cohesion to Maratha dominance; this is statistically significant at the 1% level. Similarly, the sign of β_3 when Labor/Total Costs is the dependent variable is also consistent with trading networks being the more important cause of Maratha dominance. A similar conclusion about the unimportance of Maratha social cohesion is shown in column 4 where $\beta_1 + \beta_3$ is not statistically different from zero (the exception is the case of kharif profits).

4.1.5 Worker Benefit Trading Network Access ($T > 0$)

So far, coefficient signs are consistent with the predictions of the model and consistent with trading network access being a more important determinant of Maratha dominance than their superior social cohesion. We should therefore see this also directly reflected in the patterns of reported trading network access, which we now consider.

A number of survey questions allow us to obtain information about access to Maratha trading networks. We asked individuals about their trade in inputs and outputs, if they traded inputs, we asked them what the terms were for payment. Since agricultural inputs must be utilized before returns are realized (often with a lengthy delay), small scale producers often receive a form of implicit loan by being able to defer payment for inputs until returns are had. We also asked if individuals received loans, and if so what the interest rate was on the loan. For all of these variables we obtained the caste of the individual with whom the respondent was trading, and we asked whether the trader was a village resident, or someone residing outside the village. As mentioned earlier, there are effectively only two castes in our villages who perform these trading functions. The Marwaris, an itinerant trading caste, and Marathas. The variable 'Maratha Trader' equals one if the household has traded with a Maratha for any tradeable good (which includes agricultural inputs and outputs, farm enterprise and non-farm enterprise goods). 'Terms of payments' is an index variable equal to 0 if the trader requires advanced payments, 1 if full payment is required at the time of sale; and 2 if instead payment in installments is acceptable.

According to the model, if clientelist vote trading occurs, low caste individuals should report increased

access to various forms of Maratha trading network. They should be: more likely to transact with a Maratha, on better terms, and to borrow at a lower interest rate. Thus for each measure of trading network access we should see $\alpha_v > 0$ which implies $\beta_1 > 0$. There should be no such pattern for Marathas however, as they have access independent of vote trading. The final panel in Table 4 shows that this sign restriction of the model is again observed in the estimates. Low caste individuals in MLD villages are 0.09 more likely to trade with a Maratha (5% significance), 0.10 more likely to trade with a Maratha trader who is a non-resident of the village (1% significance), and 0.24 more likely to obtain a loan from a Maratha (1% significance). These all reflect increases of more than 50% since the averages of these probabilities are 0.18, 0.19, and 0.24 respectively for low castes. They are also more likely to obtain input trade on beneficial terms (non-advance or deferred repayment), and pay 8.4% lower annualized interest rates on their loans (5% significance). Note again that this does not simply reflect an increased village wide prevalence of Marathas as we are controlling for Maratha population in this regression. In contrast, again as predicted, the coefficient does not generally differ from zero for Marathas, except for terms of payments which also appear better for them; a finding for which the model provides no direct explanation.

The model's prediction on $\beta_2 = 0$ however is more problematic to test for this variable. This is because trading with a Maratha is mechanically more likely to occur in villages where there is a large Maratha population. Thus the coefficient β_2 should directly reflect not just the vote trading effect (0) but also the greater frequency of Marathas. Accordingly, we see positive and large effects of Maratha population numbers on a number of the Maratha trader variables due to this mechanical effect.

The coefficient on β_3 again indicates the relative strength of trading network benefits versus superior caste cohesion. The point estimates are always negative, and with $\alpha_v < 0$ would again imply $T > X$. Though these estimates are not all significantly different from zero at conventional levels. This evidence again points in the same direction: trading networks seem a more important underpinning of Maratha Dominance than superior Maratha social cohesion.

4.2 Indirect Evidence

There are a number of variables that we may intuitively expect to be affected under a system of clientelist vote trading even though we do not include them in the formal model. We consider those variables here.

Maratha Pradhan

One direct prediction of elite capture is that the caste of the elected village government leader (*Pradhan*) is a Maratha. As seen in Table A5 in Appendix A, though Marathas comprise about 40% of the population, they are significantly over-represented as the Pradhan. We now check whether this over-representation's relationship with our Maratha variables. Estimates reported in the village level results (Table 2) confirm that controlling for $MPROP$, MLD villages are 0.43 more likely to elect a Maratha Pradhan, this relationship is statistically significant at the 1% level. Again, this effect is over and above that arising through the mechanical effect of increased Maratha population numbers captured through the coefficient β_2. Moreover,

there is a negative and statistically significant coefficient on the interaction term $MLD*MPROP$. Marathas are thus *less* likely to become pradhan in MLD villages the more Marathas in the village population. This is a highly surprising result on a priori grounds, but would again be consistent with clientelist vote trading where the Maratha landlord advantage in sustaining such vote trading derives from trading network access.

Voter Motivation

We also asked respondents why they voted the way they did in the last election, suggesting a list of reasons ranging from candidate qualifications, personal characteristics, party affiliation and policies, to reasons related to a personal connection with the candidate. We construct the variable 'voted-personal' equal to one if the household voted for a candidate due to a personal connection rather than due to the other reasons, which are coded zero.

We break up the sample and report results for 'vote personal' for low caste individuals separately from Marathas. We would expect the coefficient β_1 on MLD to be positive for low caste individuals. Vote trading should be more likely in MLD villages, and if this is more likely, the probability that low caste individuals report that they 'voted-personal' should be higher. This should not, however, be the case for high castes (Marathas), as if trading network access is at the heart of the Maratha advantage, this is not a benefit that can be traded with a Maratha worker in return for vote support. As before, $MPROP$ should have no effect on 'vote-personal' once MLD and the interaction between MLD and $MPROP$ are included in the regression.**[**Insert Table 5**]**

The data is again consistent with these predictions; results from an estimation of (12) with 'voted-personal' as the dependent variable are reported in the top two rows of Table 5. The coefficient on 'voted-personal' for low caste individuals, β_1(of (12)), is equal to 0.11 (and significant at the 1% level), suggesting that in a MLD village, with negligible $MPROP$, lower caste individuals are more likely to report that they voted based on personal connections to the candidate than in villages where another caste is land dominant.[29] The average of this variable is 0.24 so this reflects roughly a 46% increase. The coefficient on $MPROP$, β_2 (of (12)), is not significantly different from zero. No such pattern of response was found for Marathas, the coefficient β_1 for Marathas is negative but not significantly different from zero. The model interpretation of this is that Maratha landlords do not have a comparative advantage in vote-buying from Maratha workers, again consistent with trading network access being a key benefit to be traded.

This is corroborated by the sign of the coefficient on β_3. The negative and significant sign indicating that MLD has weaker effects on vote trading from lower castes the larger is $MPROP$. The larger is $MPROP$, the less important for achieving GP control it is for landlords to 'buy' votes from non-Marathas.

Social Capital/Village Cohesion

A natural alternative to the clientelist explanation for the low level of programs in MLD villages could

[29]We also ran analogous estimations where we conditioned on being in an unreserved area (so that Marathas are free to stand for election). We then distinguished between whether there is a Maratha Pradhan or not. Interestingly, the sign of the coefficient β_1 is consistent with the model's predictions only when the outcome of voting is a Maratha pradhan. This coefficient is insignificant if the Pradhan is non-Maratha.

be that the traditionally dominant Maratha landlords and majority non-Maratha workers are unable to coordinate on effective governance that would yield a good level of programs. A large literature (an early prominent example of which is Putnam et. al. 1993) suggests that local governance depends critically on good grass roots level social cohesion – or social capital. This would suggest that less functionality in local government arises where Marathas are dominant land holders because of a legacy of mistrust between upper caste landowners and the poor that persists to the present day.[30]

We asked a number of questions about trust and social capital. The variable 'Trust' is the response to: "Would you say that the large landholders can be trusted? (1=Almost none, 2=Some; 3=Majority; 4=Almost All). 'No Cheat' refers to answering someone from a higher caste is less likely to cheat you (compared to individuals from other castes or wealth levels). 'Repair' is the answer to "If someone from your village noticed something wrong on your farm they would?" repair it themselves (compared to conditional answers, such as "alert you if he is from a lower caste"....etc.). Another measure of social capital, suggested by Robert Putnam, is indicated by voluntary donations on the part of individuals – both time and money. Accordingly, we include the variables, 'donated cash' and 'donated labour' , which are dummy variables equal to one if the household did donate (cash or labour respectively) in the past year to a development project within the village. 'Share Water' is equal to 1 if the household shares a water source with members of the Maratha caste. We report the answers to these questions for low castes.

If clientelist vote trading is a voluntary quid pro quo, then we may expect this to actually build social capital, or at least not lower it. Workers willingly reduce access to programs (benefits from the government) in return for insurance (benefits from patrons). If however this is a consequence of forced landlord hegemony, then the low castes should be more likely to report ill of their higher caste neighbours. As the sign of β_1 from the second panel of Table 5 makes clear, caste relations are clearly no worse, and indeed probably better, in villages where the dominant land owners are Marathas. Low caste individuals are more likely to report that large landowners in their village can be trusted, that they are less likely to be cheated by higher castes and that someone from their village will "repair themselves" damage they notice on the respondent's land. Low caste households are also more likely to have donated cash, or labor to village level development initiatives in the last year if they reside in MLD villages.

The interaction term, β_3, again informs the relative importance of the sources of Maratha dominance. As the trading network explanation suggests, the effects of MLD are mitigated by an increase in $MPROP$; $\beta_3 < 0$. This result is particularly striking as it suggests that the positive social capital enhancing effect of MLD is less likely to arise even as $MPROP$ increases. That is, there is a negative effect on social capital with increased caste coherence between the two economic classes; an a priori surprising finding. Once again, Maratha trading networks being a key factor in sustaining clientelism would predict this. Increased

[30]The recent and growing literature on the historical determinants of social capital (Nunn and Wantchekon 2011, Guiso et. al. 2009 Tabellini 2010, Algan and Cahuc 2010) emphasizes the importance of such historical factors in affecting contemporary social capital.

$MPROP$ in a MLD village makes trading network access less usable in vote trading. and thus lowers the probability of clientelist vote trading. If clientelism induces improved social relations between the classes, this explains the signs on both β_1 and β_3.

Additional variables of interest, include 'Agree' which refers to answering "that most people in the village would agree on the type of development project the village should have (compared to differences of opinions). 'Target Village' refers to GP funds should be targeted to village as a whole, compared to poor or low caste individuals. Likewise 'Shared Funds' refers to GP funds are shared across the village compared to going directly to poor or low caste individuals. 'Festivals' is equal to one if there are village funds for festivals. All of these variables are significantly higher in MLD villages. These results suggest that clients are not simply acceding to Maratha power in these villages grudgingly. Instead they view the village more as a collective undertaking, and the GP not as a possible source of redistribution, but as an expression of this collective will. This is consistent with good social capital in a village not being correlated with good governance but with its converse; the capture of democratic political structures by the landed elite. Consistent with Acemoglu and Robinson's (2008) argument, and the examples discussed in Acemoglu et. al. (2013), the advent of democratic structures may induce the elite to switch strategies in order to maintain control. Here this is consistent with their cultivating better social relations with the non-elite (workers) via insurance provision and would explain the higher social capital in MLD villages.

4.3 Robustness Checks

In Appendix B, we report the results from a number of robustness checks. Tables B1 and B2 exclude all village controls.[31] Tables B3 and B4 also interact all the key village-level variables with $MPROP$. Tables B5 and B6 utilize an alternative continuous measure of MLD constructed from household surveys. The main results discussed above all continue to hold.

5 Alternative Interpretations

In line with the theory we have developed, we have used MLD, $MPROP$, and their interaction to identify clientelism occurring in a village. We have estimated reduced form relationships between these variables and a set of outcome variables that our theory predicts will be affected when clientelism arises. The theory predicts that MLD should make clientelism more likely, that $MPROP$ has no independent effect on it, and that the sign of the interaction term indicates which one of two prominent explanations for Maratha political hegemony is most important. The two predictions almost always hold for all dependent variables, and the implication from the interaction term consistently favours Maratha trading networks as a key source of Maratha political dominance.

As discussed, interpreting the evidence this way depends on the outcome variables we have reported not

[31] Though not reported here, the main results are also robust to excluding the household level control variables.

themselves being directly affected by either MLD or $MPROP$. In Section 2.2, we pointed to some main threats to our exclusion restrictions, we now further assess these.

Mismeasured Land Productivity

One threat we have already discussed is through a potential correlation between village productivity and Maratha prevalence - quite conceivably Marathas, as the dominant caste, expropriated the highest quality lands for themselves. We do not find any significant differences on observable dimensions of soil quality and the geographic and climatic cultivatability of land across MLD and non-MLD villages. However, if these are inadequately controlled for and either of our Maratha variables are correlated with land productivity then the interpretation of the coefficients we have forwarded above is not correct. It it impossible to measure this directly, but we will now argue that the sign of our results suggest that even if there is some direct effect of our Maratha variables on land quality, and we have not picked it up with our controls, this cannot explain the findings we have presented.

If MLD is positively correlated with an unobserved land quality dimension then consider our specifications with log yields, log profits and wages as dependent variables. For any one of these, then the true relationship between the dependent variable and the co-variates includes, in addition to the effects of our clientelism channel (measured through the β variables as before) a direct effect (denoted α_1 and α_2 below):

$$Y_{ik} = \beta_0 + \beta_1 MLD_k + \beta_2 MPROP_k + \beta_3 MLD_k \cdot MPROP_k + \boldsymbol{\psi} Z_{ik} + \boldsymbol{\gamma} Z_k + \alpha_1 MLD_k + \alpha_2 MPROP_k + \epsilon_{ik}$$

Thus the estimated coefficient on MLD includes β_1 interpretable through the model and α_1 via direct effects of land quality (similarly for $MPROP$). Consider how such mismeasurement would tend to bias these dependent variables. For yields and profits the bias introduced by mismeasured land quality, correlated with either of our Maratha variables, makes α_1 and $\alpha_2 > 0$. This biases upwards our estimates of β_1 and β_2 and indeed could explain the positive coefficient that we reported on MLD in the middle panel of Table 4. But another implication of such mismeasurement of productivity would also be higher wages. Similarly, this would imply that MLD in the top panel of Table 4 where wages are the dependent variable are also upward biased. Here, however, we find a strong and significant negative coefficient on MLD. Mismeasured land productivity correlated with MLD would then suggest a true effect operating through our model larger than that reflected in the coefficient on β_1. Mismeasured land productivity cannot consistently account for the full pattern of results. To the extent that it explains some part of the coefficient on log profits and yields, it makes the negative coefficient on wages from MLD even harder to understand. Further, such mismeasurement provides no clear implication for the sign of the interaction term. It is positive for wages and strongly negative for log profits and yields. MLD villages having higher unmeasured productivity could not explain an interaction term that always enters in the opposite direction to the effect of MLD.

With respect to the other dependent variables – insurance, social capital and Maratha traders – there are simply no obvious implications of mismeasured land productivity in MLD villages. One may hypothesize that mismeasurement could perhaps rationalize insurance and social capital being higher in MLD villages

– a wealth effect could perhaps be at play – but it would not be possible to explain why such insurance or social capital would be mitigated by increasing $MPROP$, as implied by the sign of the interaction term. Moreover, this channel seems inconsistent with any systematic pattern of Maratha trader access.

Direct Maratha Power

Another threat to the exclusion restriction may arise through the power of Maratha landlords and population numbers making the village more effective in receiving state level support. In a direct form, this is entirely inconsistent with the data. Marathas' traditional positions of prominence and power in state organizations should lead them to be favoured in receipt of programs, the EGS, and have greater GP budgets. We find the opposite for all of these, and a zero coefficient on $MPROP$. Again, predictions that are expected with the clientelism channels of the model.

A more subtle variant of such a channel arises if Maratha landlords are able to use their power to get what they want from the state – i.e., as we have argued, reduce programs. This could then lead to knock on effects consistent with our model, and as observed in the data; higher yields, profits and lower wages in MLD villages. Extending this story further, where Maratha workers are numerous, the capacity of Maratha landlords to exert their power and block programs from the centre is mitigated. So the interaction term on $MPROP$ for programs is positive and the knock-on effect is similarly to mitigate higher profits/yields and the lower wages in MLD villages.

The difference between this explanation and one based on the quid pro quo of clientelism that we have explored comes from other elements of the posited vote trading. Since clients are not forced by their patrons, but must be coopted, landlords must "pay" workers for their political support. As predicted by the clientelist story then, workers in such villages should report greater insurance provision from landlords – something we would not see if landlords are able to force policy reduction on their own. We also see greater access to Maratha trading networks, again consistent with clientelism, and not something that would arise were it simply the case that Maratha landlords are able to get their way. Such an explanation would also be unable to account for the higher social capital we see in MLD villages – especially that reported from low castes to Marathas – if the Maratha landlords are merely forcing the policies that they want. In contrast, with the clientelism explanation, social capital is higher because the workers are in a voluntary agreement with the landlords. They obtain fewer programs due to their support, but receive valuable insurance instead. A further implication of this alternative explanation can be checked through our household surveys. In these we also asked individuals to report their approval of the GP's personnel and behavior. If Maratha landlords are forcing the policies they want, then workers should report lower approval of the GP and its personnel in such villages. In contrast, under clientelism, workers understand the price they pay for their personal benefits is landlord control of the GP, so they will not report dissatisfaction with the activities of the GP and more so than in villages without clientelism. The latter is what we find. Table B7 in Appendix B reports that the landless are not significantly more dissatisfied in MLD villages with regards to the quality and performance of the GP. In particular, they do not report significantly more dissatisfaction with regards

to honesty and fairness, qualifications, providing public goods, allocating funds fairly, not discriminating across villagers, obtaining upper level government funds, and solving village disputes. Even though, as we have shown, programs are lower, the EGS is less frequent, and village level expenditures are diminished.

Overall, it is not possible to definitively rule out that mismeasurement of land productivity or other factors of the village correlated with its Maratha variables – such as Maratha influence at higher government levels – could be biasing some of the estimates we report for some of the dependent variables. However, it is extremely unlikely that such channels of omitted influence could either explain the patterns that we have documented in the data, or not lead to inconsistencies on numerous other dimensions of the data. In contrast, the patterns in the data are consistent with the channels of the postulated clientelism model, and consistent with Maratha trading networks providing a key underpinning of Maratha landlord power. The model can also explain the reported reasons for voting, the provision of insurance, and the inferred levels of social capital.

However, we emphasize again that the test of the model we provide here is indirect. We do not observe vote trading directly, nor do we observe the provision of benefits directly in return for an individual's vote. It is thus impossible to rule out alternative explanations of the empirical regularities we had documented, perhaps due to omitted factors that we have not considered. Alternatively, a combination of the factors that we have argued are insufficient to explain the results on their own could also conceivably be behind these patterns.

6 Conclusions

An extensive data collection exercise in the Indian state of Maharasthra has allowed us to study the performance of local democracies there. We conclude that local governments – though elected through a process that is fully consistent with a vibrant democracy – do not act to benefit the poor majority that elects them. We have argued that government behavior, wages, profits, yields, the patterns of help in contingencies and trading network access observed across these villages all line up as predicted by clientelist vote-trading between large landowners who are buyers of votes, and workers who are selling them.

Our analysis makes an unambiguous conclusion about the welfare implications of vote trading. It is bad for the majority worker group taken as a whole. But we take this conclusion with some caution as it rests on modeling assumptions that, though defensible, cannot be definitively proved from our data. Two aspects are important. Firstly, we have assumed that all of the surplus generated by vote trading goes to the landlords as they are able to make take-it-or-leave-it offers to workers. The offers thus make workers just indifferent to the clientelist vote trading and hence no better off than if they were to control the GP themselves. That landlords should hold such control over surplus division is certainly consistent with a degree of monopsony power, with many other studies, both in Maharashtra and in other agrarian contexts (Padhi 2007) but not an assumption that we can directly prove. A second assumption makes workers as a whole worse off even

while the vote-trading subset are not. This derives from our modeling of the local democratic process where it was posited that landlord control of the GP necessitated a mere half of the village's workers to vote in support of their candidate. So the vote-trading half is no worse off – they receive insurance and trading network access in return for their political support which costs them lower wages and reduced programs. But the remaining half – whose support is superfluous – just suffer the lower wages and programs. Perhaps, in some villages, workers as a whole can stand together against such divided support. In that case, a simple majoritarian calculus – where 50% of votes buys control – understates the degree to which landlords must extend benefits to workers across the village. Though there is no organized labor in our villages, and almost no organized worker political movements, there may still be informal coalitions amongst workers able to hold out for both better conditions in the vote trade, and for the broader extension of benefits that we cannot observe. To the extent these are present, such welfare implications will be mitigated, and perhaps even overturned. This appears unlikely, but is impossible to rule out.[32]

Persisting with a negative welfare implication, the results here nod toward policy that removes pro-poor policy discretion to a higher level of government that may not be so easily captured by local landed interests. It also suggests, more generally, that progressive policies designed to improve welfare are less likely to spread if a single cohesive group dominates access to key economic opportunities such as insurance or labor markets. Again, administering benefits from higher levels of government should weaken the cohesiveness of such blocking agents – it is easier for landlords in a village rather than a coalition spanning a larger region to agree on a course of action. A national level program directly targeting beneficiaries is currently being trialed in India – the Unique Identification Project (commonly referred to as *UID*). The benefits of this cannot be bartered for votes locally, and should benefit all workers. Which, given the overwhelmingly poor nature of our worker sample, should also lead to declines in measured poverty too.

Taking local interests as given and immovable, the analysis suggests that policies uniformly beneficial in their distributional impacts should be preferred. It may be necessary to trade off the magnitude of benefits to the poor in order to coopt landlords in to accepting such policies. For instance, many of the policies here would raise worker wages without benefiting landlords. But policies that benefited landlords as well, even if these had lower effects on wages – such as private well construction projects – may be preferable as they would be more likely to pass through with elite support.

[32]Since clientelism reduces resources flowing in to the village, it is also possible to imagine a welfare improving Coasian bargain between workers and landlords acting as cohesive groups. In it, landlords would agree to allow programs in, which inject resources to the village but also raise wages. In return, workers would agree to share some of these benefits with landlords. In reality, such a scheme is unworkable due to standard commitment problems. It would not be possible for workers to agree to transfer benefits to landlords ex post, after the programs are implemented. Workers would be unable to resolve this by making ex ante transfers as they are likely to be liquidity constrained. Perhaps more importantly, such a bargain would require solving a large collective action problem.

References

[1] Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001) "The Colonial Origins of Comparative Development: An Empirical Investigation" *American Economic Review*, 91(5), 1369–1401.

[2] Acemoglu, Daron, Tristan Reed, and James A. Robinson (2013) "Chiefs: Elite control of civic society and economic development in Sierra Leone" *Journal of Political Economy*, forthcoming.

[3] Acemoglu, Daron and James A. Robinson (2008) "Persistence of Power, Elites, and Institutions" *American Economic Review,* 98(1), 267-293.

[4] Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A. Robinson (2013) "Democracy Redistribution and Inequality" NBER Working Paper 19746.

[5] Algan, Yann and Pierre Cahuc (2010) "Inherited Trust and Growth" *American Economic Review*, 100(5), 2060-92.

[6] Baland, Jean-Marie and James A. Robinson (2008) "Land and power: Theory and Evidence from Chile" *American Economic Review*, 98(5), 1737-1765.

[7] Banerjee, Abhijit and Rohini Pande (2007) "Parochial Politics: Ethnic Preferences and Politician Corruption". BREAD Working Paper No. 152.

[8] Bardhan Pranab K. and Dilip Mookherjee (2000). "Capture and Governance at Local and National Levels," *American Economic Review*, 90(2), 135-139. .

[9] Bardhan, Pranab and Dilip Mookherjee (2005). "Decentralizing antipoverty program delivery in developing countries," *Journal of Public Economics*, Elsevier, 89(4), 675-704.

[10] Bardhan, Pranab and Dilip Mookherjee, (2006) "Pro-poor targeting and accountability of local governments in West Bengal" *Journal of Development Economics*, 79, 303-327.

[11] Bardhan Pranab K. & Dilip Mookherjee (2011). "Political Clientelism-cum-Capture: Theory and evidence from West Bengal," Working Paper, Institute for Economic Development, Boston University.

[12] Besley, Timothy, Rohini Pande, and Vijayendra Rao (2012) "Just Rewards? Local Politics and Public Resource Allocation in South India" *World Bank Economic Review*, 26(2), 191-216.

[13] Cameron, C. A., J. Gelbach, and D. Miller (2011) "Robust inference with multi-way clustering". *Journal of Business and Economic Statistics*, 28(2), 238-249.

[14] Carter, Anthony T. (1974) *Elite Politics on Rural India: Political Stratification and Political Alliances in Western Maharashtra*. Cambridge University Press, Cambridge.

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

[16] Chandra, Kanchan, (2004) *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India.* Cambridge: Cambridge University Press.

[17] Deshpande, Rajeshwari (2004) "Kunbi Maratha as OBC: Backward Journey of a Caste" *Economic and Political Weekly*, 39(14/15), April 3-16, 1448-1449.

[18] Finan, Frederico and Laura Schechter (2012)"Vote-buying and reciprocity" *Econometrica*, 80(2), 863-882.

[19] Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2009). "Cultural Biases in Economic Exchange?" *Quarterly Journal of Economics*, 124(3), 1095-1131..

[20] Hoff, K., M. Kshetramade and E. Fehr (2012) "Caste and Punishment: The Legacy of Caste Culture in Norm Enforcement," World Bank Working Paper.

[21] Kitschelt, H. and S. Wilkinson (2007) *Patrons, clients and policies: Patterns of democratic accountability and political competition.* Cambridge: Cambridge University Press.

[22] Kling, J. R., J.B. Leibman, L. F. Katz, and L. Sanbonmatsu (2004) "Moving to opportunity and tranquility: Neighborhood effects on adult economic self-sufficiency and health from a randomized housing voucher experiment" Working Paper 481, Industrial Relations Section, Princeton University.

[23] Lele, Jayant (1981) *Elite Pluralism and Class Rule: Political Development in Maharashtra India.* New York.

[24] Mandavdhare, Surendra (1989) *Caste and Land Relations in India: A Study of Marathwada*, New Delhi: Uppal Publishing House.

[25] Munshi, Kaivan. and Mark Rosenzweig. 2008."The Efficacy of Parochial Politics: Caste, Commitment, and Competence in Indian Local Governments." BREAD Working Paper No. 182..

[26] Nunn, Nathan and Leonard Wantchekon (2011) "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 101(7), 3221-52.

[27] Padhi, Kulamani (2007) "Agricultural Labour in India - A Close Look" *Orissa Review*, February-March, 23-28.

[28] Palshikar, Suhas (2007) *Politics of Maharashtra: Local Context of the Political Process,* Pratima Prasashan, Pune.

[29] Palshikar, Suhas, and Rajeshwari Deshpande (1999) "Electoral Competition and Structures of Domination in Maharashtra" Economic and Political Weekly, 34 (34/35), 2409-2422.

[30] Putnam, Robert, Robert Leonardi and Raffaella Y. Nanetti (1993) *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press.

[31] Ravallion, Martin, Gaurav Datt, and Shubham Chaudhuri (1991) "Higher wages for relief work can make many of the poor worse off: Recent evidence from Maharashtra's Employment Guarantee Scheme", Policy Research Working Paper Series 587, The World Bank.

[32] Robinson, James. A. and Thierry Verdier (2013) "The political economy of clientelism" *Scandinavian Journal of Economics*, forthcoming.

[33] Rosenthal, Donald B. (1977) *The Expansive Elite: District Politics and State Policy-Making in India.* University of California Press, Berkeley.

[34] Scott, James (1979) *The Moral Economy of the Peasant: Rebellion and Subsistence in Southeast Asia*, Yale University Press.

[35] Sirsikar, V. M. (1970) *The Rural Elite in a Developing Society: A Study in Political Sociology*. New Delhi.

[36] Subramanian, Narendra (1999) Ethnicity and Populist Mobilization: Political Parties, Citizens and Democracy in South India. Delhi: Oxford University Press

[37] Tabellini, Guido (2010) "Culture and Institutions: Economic Development in the Regions of Europe" Journal of the European Economic Association, 8(4), 677-716.

[38] Vicente, Pedro and Leonard Wantchekon (2009) "Clientelism and vote buying: Lessons from field experiments in west Africa" *Oxford Review of Economic Policy*, 25(2), 292-205.

[39] Vora, Rajendra (1996) "Shift in Power from Rural to Urban Sector" *Economic and Political Weekly*, January, 13-20.

[40] Wantchekon, Leonard (2003) "Clientelism and voting behavior: Evidence from a field experiment in Benin" *World Politics*, 55(3), 399-422.

[41] Wantchekon, Leonard and Thomas Fujiwara (2014) "Can informed public deliberation overcome clientelism? Experimental evidence from Benin" *American Economic Journal: Applied Economics*, forthcoming.

Table 1 - Village Geographic, Climatic, and Demographic Measures by Caste Dominance

Variable	Maratha Land Dominated	Non-Maratha Land Dominated	Equivalence of Means
Population	2071.8 (639.7)	2032.3 (1012.0)	39.5 (95.1)
Households	369.9 (60.4)	374.8 (64.5)	-4.9 (7.8)
Proportion SC	0.14 (0.08)	0.15 (0.11)	-0.016 (0.012)
Distance to Town	21.9 (12.6)	20.3 (12.9)	1.6 (1.5)
Distance to Road	2.6 (2.7)	2.8 (2.9)	0.2 (0.3)
Distance to Water	2.9 (2.2)	3.0 (2.5)	-0.05 (0.3)
Uncultivatable Land	97.3 (116.6)	94.0 (122.7)	3.3 (14.1)
Area - Cultivated	2225.8 (964.1)	2170.4 (1655.5)	55.3 (182.8)
Area - Irrigated	504.7 (718.7)	489.9 (929.4)	14.8 (116.1)
Area - Rainfed	1586.2 (1074.1)	1501.0 (1107.6)	85.2 (156.7)
Area - Tree/groves/orchards	29.1 (64.5)	26.6 (59.6)	2.5 (9.2)
Area - Forest	99.1 (335.9)	54.8 (135.4)	44.3 (41.2)
Area - Pasture/grazing	58.6 (147.0)	42.7 (104.8)	15.9 (19.5)
Area - Fallow	168.5 (283.6)	167.9 (521.3)	0.6 (56.6)
Area - Inhabited	30.4 (84.6)	35.3 (94.5)	4.9 (13.1)
Area - Community/Panchayat	8.9 (17.8)	12.8 (32.7)	3.9 (3.6)
No Alkalinity	0.89 (0.31)	0.87 (0.34)	0.02 (0.04)
No Waterlogging	0.88 (0.33)	0.82 (0.39)	0.06 (0.05)
No Soil Erosion	0.92 (0.26)	0.87 (0.34)	0.06 (0.04)
Topsoil Dominant Nitrogen	2.82 (0.60)	2.82 (0.38)	0.003 (0.06)
Topsoil Dominant Organic Carbon	3.03 (0.43)	3.00 (0)	0.03 (0.04)
Topsoil Dominant Ph	3.98 (0.41)	3.93 (0.26)	0.05 (0.04)
Soil Depth	0.25 (0.15)	0.24 (0.16)	0.01 (0.02)
Soil Color	0.54 (0.21)	0.54 (0.21)	0.005 (0.03)
Salinity	1.89 (0.17)	1.89 (0.18)	0.0002 (0.02)
Percolation	1.81 (0.20)	1.83 (0.21)	0.01 (0.03)
Drainage	1.82 (0.22)	1.83 (0.19)	0.02 (0.03)
Rainfall	62.3 (14.7)	64.9 (19.7)	2.6 (2.0)
Observations	194	112	

Notes: The sample excludes areas where no Marathas reside (Eastern Vidarbha). Standard deviations are in parentheses in first two columns. Standard errors are in parentheses in the third column. The first seven varibles come from the Village Amenities and GPS Data from the Census of India 2001. SC refers to Scheduled Castes, the lowest ranking group in the caste hierachy. Our own data on village population numbers do not vary by caste dominance either. The next nine variables, pertaining to land-use patterns, also come from the Village Census of India 2001. The first three measures of soil quality come from our village survey. The next three variables pertaining to topsoil (30 cm) content come from FAO-UNESCO soil maps. Our soil quality measures we chose after consulting with an agricultural specialist in the Faculty of Land and Food Systems at the University of British Columbia. The last five variables on soil quality come from our household survey, aggregated up to the village level. Rainfall information, which is only available at the district level, comes from the India Meteorological Department.

Table 2 - Estimations of GP Measures

Dependent Variable	Coefficient (β_1) MLD	Coefficient (β_2) $MPROP$	Coefficient (β_3) $MLD \cdot MPROP$	Coefficient (β_1+β_3)	Obs
Maratha Pradhan	0.43 (0.15)***	1.21 (0.24)***	-0.51 (0.28)*	-0.08 (0.18)	290
All Programs (1)	-1.01 (0.36)***	-1.19 (0.57)**	1.84 (0.72)***	0.83 (0.49)*	8140
BPL Programs (1)	-0.42 (0.12)***	-0.37 (0.21)*	0.81 (0.27)***	0.39 (0.18)**	8140
EGS (1)	-0.10 (0.04)***	0.02 (0.08)	0.19 (0.10)**	0.09 (0.07)	8140
Income Programs (1)	-0.98 (0.33)***	-1.11 (0.54)**	1.84 (0.68)***	0.87 (0.46)*	8140
Non-Income Programs (1)	-0.03 (0.04)	-0.08 (0.07)	-0.004 (0.08)	-0.04 (0.05)	8140
All Programs (2)	-0.94 (0.37)***	-1.04 (0.57)*	1.44 (0.75)**	0.49 (0.50)	291
BPL Programs (2)	-0.40 (0.13)***	-0.35 (0.21)*	0.69 (0.28)***	0.29 (0.18)	291
EGS (2)	-0.08 (0.04)**	0.02 (0.09)	0.18 (0.11)*	0.10 (0.08)	291
Income Programs (2)	-0.92 (0.34)***	-1.00 (0.54)*	1.52 (0.70)**	0.61 (0.47)	291
Non-Income Programs (2)	-0.02 (0.04)	-0.04 (0.07)	-0.09 (0.09)	-0.11 (0.06)*	291
Revenue (1)	-157.4 (86.2)*	-173.5 (237.1)	122.8 (231.5)	-34.6 (200.5)	220
Revenue (2)	-10.4 (5.4)**	-16.0 (8.0)**	25.4 (10.7)**	14.9 (5.8)**	307
Expenditure	-12.6 (6.6)**	-17.6 (31.1)	29.0 (12.6)**	16.4 (6.8)**	307
BDO Meetings	-2.37 (1.04)**	1.46 (3.92)	0.15 (4.03)	-2.22 (3.45)	290
MP Meetings	-3.04 (1.54)**	-2.05 (2.49)	3.28 (1.92)*	0.25 (1.46)	290
DC Meetings	-2.04 (0.87)**	-3.96 (1.77)**	3.32 (1.67)**	1.29 (1.02)	290
Meetings (AES)	-2.48 (1.06)**	-1.51 (1.85)	2.25 (2.13)	-0.23 (1.53)	290

Notes: All estimations include village-level controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and whether the GP is reserved) and regional fixed effects. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Robust standard errors are in parentheses. Acronyms used are: Maratha land dominated (MLD); Maratha popluation proportion ($MPROP$); Below Poverty Line (BPL); Employment Guarantee Scheme (EGS); Block Development Officer (BDO); District Collector (DC); Member of Parliament (MP); and Average Effect Size (AES). Revenue (1) refers to data collected from the balance sheets (covers last 24 months) submitted by the GPs (obtained using RTI Act). Revenue (2) and Expenditure are annual per capita values from the 2001 Village Census. Information on programs (Programs (1); BPL Programs (1); EGS (1); Income Programs (1); Non-Income Programs) are reported from household level data, and regression disturbance terms are clustered at the village level for these estimations and household level controls are also included (education, land ownership, and caste identity). Programs (2); BPL Programs (2); EGS (2), Income Programs (2), and Non-Income Programs (2) are variables which aggregate this household level information up to the village level. Estimations are OLS except EGS (1), which is a probit estimation, where the reported coefficients are the partial derivates of the predicted probability. Meetings (AES) is the estimated average effect size of the three variables: BDO Meetings, DC Meetings, and MP Meetings.

Table 3 - Estimations of Insurance Relations

Dependent Variable	Coeff. (β_1) MLD	Coeff. (β_2) $MPROP$	Coeff (β_3) $MLD \cdot MPROP$	Coefficient $\beta_1 + \beta_3$	Obs
Insured (1) [Landless]	0.13 (0.05)***	0.10 (0.07)	-0.20 (0.09)**	-0.06 (0.07)	2564
Insured (2) [Landless]	0.14 (0.05)***	0.08 (0.07)	-0.22 (0.10)**	-0.08 (0.07)	2564
Insured (3) [Landless]	0.14 (0.05)***	0.10 (0.07)	-0.22 (0.10)**	-0.08 (0.07)	2564
Insured (4) [Landless]	0.17 (0.05)***	0.16 (0.07)**	-0.27 (0.10)***	-0.11 (0.07)	2564
Insured (5) [Landless]	0.16 (0.05)***	0.09 (0.07)	-0.26 (0.10)***	-0.10 (0.08)	2564
Insured (6) [Landless]	0.16 (0.05)***	0.11 (0.07)*	-0.28 (0.10)***	-0.12 (0.07)*	2564
Insured (AES) [Landless]	0.14 (0.03)***	0.05 (0.06)	-0.19 (0.08)***	-0.05 (0.06)	2564
Insured (AES) [Low Castes]	0.11 (0.03)***	0.13 (0.05)***	-0.25 (0.07)***	-0.14 (0.05)***	3246
Insurer [Large Land Owners]	0.08 (0.03)***	0.02 (0.06)	-0.15 (0.07)**	-0.07 (0.09)	1932
Insurer [Marathas/Large Land]	0.26 (0.06)***	0.14 (0.11)	-0.31 (0.12)***	-0.05 (0.07)	1066
Insurer [Low castes/ Large Land]	-0.02 (0.04)	-0.02 (0.09)	0.002 (0.11)	-0.02 (0.10)	807
Insurer [Marathas/ Low Land]	0.07 (0.06)	-0.13 (0.08)	-0.04 (0.09)	0.03 (0.05)	1110
Insurer [Low castes/ Low Land]	0.05 (0.04)	0.14 (0.08)*	-0.21 (0.10)**	-0.16 (0.08)**	2311

Notes: All estimations include village-level controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and whether the GP is reserved), household-level controls (education, land ownership, and caste identity), and regional fixed effects. Regression disturbance terms are clustered at the village level. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Acronyms used are: Maratha land dominated (MLD); Maratha popluation proportion ($MPROP$); and Average Effect Size (AES). Insured (1): "Would most people in your village help you with some money in times of crisis?". Insured (2): "Would a higher caste member of your village help you with some money in times of crisis?". Insured (3): "Would most people in your village help a lower caste villager with some money in times of crisis?". Insured (4)-(6) are the same questions with "money" replaced by "grain". Insured (AES) is the estimated average effect size of the six variables: Insured (1) to (6). Insurer: "Suppose a lower caste man asks to borrow a good sum of money from you because someone in his family has fallen ill. He is from the village and has the ability to repay the amount. Would you lend it to him?". Large land owners have more than 5 acres. Low land owners have between 0 and 2.5 acres. Estimations are probits (except AES) where the reported coefficients are the partial derivates of the predicted probability.

Table 4 - Estimations of Wages, Yields, and Trading Relations

Dependent Variable	Coeff.(β_1) MLD	Coeff. (β_2) $MPROP$	Coefficient (β_3) $MLD \cdot MPROP$	Coefficient $\beta_1 + \beta_3$	Obs
Daily Wage [All Labourers]	-2.23 (0.91)***	-1.70 (1.82)	5.41 (2.13)***	3.18 (1.58)**	13546
Daily Wage [Males]	-2.42 (1.19)**	-2.25 (2.22)	6.70 (2.71)***	4.28 (1.99)**	7480
Daily Wage [Females]	-1.98 (0.86)**	-1.05 (1.71)	3.76 (1.93)**	1.77 (1.42)	6066
Daily Wage [Landless]	-2.09 (1.02)**	-1.05 (2.29)	4.79 (2.50)**	2.70 (1.95)	5518
Daily Wage [Low Castes]	-2.34 (0.99)**	-1.92 (1.95)	4.75 (2.35)**	2.42 (1.78)	9167
Daily Wage (Insurance) [Low]	-2.79 (1.29)**	-1.91 (2.58)	5.93 (3.05)**	3.14 (2.27)	4342
Daily Wage (No Insurance) [Low]	-2.06 (0.98)**	-2.17 (1.87)	4.12 (2.35)*	2.05 (1.83)	4833
Log Kharif Yields	0.32 (0.12)***	0.13 (0.20)	-0.51 (0.23)**	-0.19 (0.16)	2320
Log Kharif Profit	0.48 (0.15)***	0.43 (0.28)	-1.04 (0.30)***	-0.56 (0.21)***	1838
Labour/Total Costs (Kharif)	-0.04 (0.02)*	-0.05 (0.04)	0.07 (0.04)*	0.03 (0.03)	1800
Maratha Trader [Marathas]	0.001(0.06)	0.08 (0.12)	0.10 (0.12)	0.10 (0.08)	2725
Outside Maratha Trader [Marathas]	0.04 (0.06)	0.12 (0.11)	0.03 (0.11)	0.07 (0.07)	2655
Maratha Lender [Marathas]	-0.26 (0.16)	0.12 (0.25)	0.04 (0.29)	-0.22 (0.21)	371
Terms of Payment (Inputs) [Maratha]	0.11 (0.06)**	0.08 (0.11)	-0.15 (0.11)	-0.03 (0.08)	9864
Interest Rate on Loan [Marathas]	-9.0 (5.9)	-32.3 (8.7)***	32.3 (9.9)***	23.3 (6.3)***	392
Maratha Trader [Low Castes]	0.09 (0.04)**	0.18 (0.09)**	-0.01 (0.11)	0.09 (0.08)	3012
Outside Maratha Trader [Low Castes]	0.10 (0.03)***	0.18 (0.06)***	-0.18 (0.08)**	-0.08 (0.06)	2793
Maratha Lender [Low Castes]	0.24 (0.08)***	0.45 (0.16)***	-0.21 (0.19)	0.04 (0.15)	452
Terms of Payment (Inputs) [Low]	0.08 (0.05)*	0.24 (0.10)**	-0.26 (0.11)**	-0.18 (0.09)*	10034
Interest Rate on Loan [Low Castes]	-8.41 (4.15)**	-2.03 (10.7)	1.73 (11.27)	-6.68 (9.51)	250

Notes: All estimations include village-level controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and whether the GP is reserved), household-level controls (education, land ownership, and caste identity), and regional fixed effects. Regression disturbance terms are clustered at the village level. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Acronyms used are: Maratha land dominated (MLD); and Maratha popluation proportion ($MPROP$). The sample of labourers are all those who work for a daily wage in agriculture. Additional individual controls (gender, age, education) are included in the wage estimations. Regression disturbance terms are clustered at the household and village level for these estimations using the approach of Cameron, Gelbach and Miller (2011). The sample for the yields, profits, proportion of labour costs regressions is large cultivators ($>$ 5 acres of land). All measures are per acre of land. Kharif yields are the total value of output per acre of land for a given crop, summed over all of the kharif crops for each household. Kharif profit is yields net of input costs (seeds, fertilizer, irrigation, electricity, pesticides, and labour). Workers include partime and fulltime, same results held if restricted ourselves to fulltime workers. Additional crop controls are included in the yields and profits estimations. Maratha Trader is equal to one if the household has traded with a Maratha for any tradeable good (which includes agricultural inputs and outputs, farm enterprise and non-farm enterprise goods) conditional on trading goods. Outside Maratha Trader refers to the trader residing outside of the village conditional on trading goods. Maratha Lender refers to lending money conditional on the household borrowing money. These estimations on Maratha trading relationships are probit estimations, where the coefficients reported are the partial derivates of the predicted probability. Terms of payments is an index variable equal to 0 if the trader requires advanced payments; 1 if full payment is required at the time of sale; and 2 if instead payment in installments is acceptable. These are ordered probit estimations. Terms of payments and interest are reported per individual loan, for these estimations regression

disturbance terms are clustered at the household and village level for these estimations using the approach of Cameron, Gelbach and Miller (2011).

Table 5 - Estimations of Inter-Caste Relations, Collective Village Life, and Voting

Dependent Variable	Coeff. (β_1) MLD	Coeff. (β_2) $MPROP$	Coeff (β_3) $MLD \cdot MPROP$	Coefficient $\beta_1 + \beta_3$	Obs
Voted-Personal [Marathas]	-0.07 (0.05)	-0.03 (0.07)	0.10 (0.07)	0.03 (0.05)	2780
Voted-Personal [Low Castes]	0.11 (0.04)***	0.11 (0.08)	-0.18 (0.09)**	-0.07 (0.07)	2108
Trust [Low Castes]	0.14 (0.06)**	0.14 (0.11)	-0.44 (0.13)***	-0.31 (0.10)***	4954
No Cheat [Low Castes]	0.04 (0.02)*	-0.02 (0.05)	-0.05 (0.06)	-0.08 (0.78)	4743
Repair [Low Castes]	0.08 (0.03)**	0.001 (0.06)	-0.11 (0.08)	-0.03 (0.05)	4927
Donated Cash [Low Castes]	0.08 (0.04)**	-0.001 (0.06)	-0.05 (0.09)	0.03 (0.06)	4965
Donated Labour [Low Castes]	0.08 (0.03)***	0.07 (0.06)	-0.06 (0.08)	0.02 (0.06)	4965
Agree [Low Castes]	0.07 (0.04)*	0.18 (0.07)***	-0.21 (0.09)***	-0.14 (0.06)**	4959
Social Capital (AES) [Low Castes]	0.07 (0.01)***	0.07 (0.03)***	-0.14 (0.03)**	-0.07 (0.02)***	4693
Share Water [Low Castes]	0.27 (0.06)***	0.54 (0.13)***	-0.26 (0.14)**	0.01 (0.10)	2929
Target Village [Marathas]	1.61 (0.47)***	0.46 (1.42)	-2.72 (0.80)***	-1.11 (0.56)**	3059
Target Village [Low Castes]	1.05 (0.32)***	0.44 (0.66)	-1.79 (0.77)***	-0.74 (0.53)	4865
Shared Funds [Marathas]	1.98 (0.52)***	0.78 (1.45)	-2.47 (0.78)***	-0.49 (0.48)	2795
Shared Funds [Low Castes]	0.85 (0.33)***	1.56 (0.70)**	-2.05 (0.80)***	-1.19 (0.63)*	4584
Festivals	0.08 (0.04)**	0.02 (0.05)	-0.07 (0.07)	0.01 (0.05)	8167

Notes: All estimations include village-level controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and whether the GP is reserved), household-level controls (education, land ownership, and caste identity), and regional fixed effects. Regression disturbance terms are clustered at the village level. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Acronyms used are: Maratha land dominated (MLD); Maratha popluation proportion ($MPROP$); and Average Effect Size (AES). Voted - Personal equals to one if the houshold voted for a candidate due to a personel connection rather than due to the characteristics of the candidate (honesty, good reputation, qualifications). Samples are conditional on voting. The sample of low castes in the voting regressions is SC/STs. Trust is response to: "Would you say that the large landholders can be trusted? 1=Almost none, 2=Some; 3=Majority; 4=Almost. Cheat refers to answering somone from a higher caste is most likely to cheat you (compared to other castes or wealth levels). Repair is the answer to "If someone from your village noticed something wrong on your farm they would?" repair it themselves (compared to conditional answers, such as "alert you if he is from a lower caste....etc). Donated cash or labour are dummy variables equal to one if the household did donate (cash or labour respectively) in the past year to a development project within the village. Share Water is equal to 1 if the household shares a water source with members of the Maratha caste. Samples are conditional on sharing a water source. Probit models are estimated for all of these variables except for the trust question. The reported coefficients for the probit estimations are the partial derivates of the predicted probability. Target Village refers to GP funds should be targeted to the village as a whole, compared to poor or low caste individuals. Shared funds refers to GP funds are shared across the village (e.g. for development projects; public goods) compared to going directly to the poor or low status; the rich and high status; or to GP members or other government officials directly. These four estimations are estimated as multinomial logit models. Agree refers to answering that most people in the village would agree on the type of development project the village should have (compared to differences of opinions within the village). Fesitvals is equal to one if there are village projects to finance festivals. Social Capital (AES) is the estimated average effect size of the six variables: Trust, No Cheat, Repair, Donated Cash, Donated Labour, and Agree.

FOR ONLINE PUBLICATION

7 Appendix A: Summary Statistics

Table A1 - Summary Statistics - GP and Village Measures

Variable	Mean	Standard Deviation	Observations
Maratha Pradhan	0.41	0.49	300
Reserved	0.58	0.49	319
Population	2271.2	659.9	319
Proportion Maratha	0.41	0.31	310
Proportion SC/ST	0.25	0.17	310
Maratha Land Dominated	0.61	0.49	320
Distance to Water	2.85	2.19	318
Distance to Road	2.60	2.75	318
Distance to Rail	22.6	18.9	318
River/Canal	0.26	0.44	320
Topsoil Nitrogen	2.02	0.89	318
Topsoil Organic Carbon	0.21	0.93	318
Topsoil Ph	0.53	1.21	318
Rainfall	70.98	20.06	318
Longitude	76.21	1.19	320
Latitude	19.46	1.04	320
Elevation	483.0	138.4	320
All Programs	5.36	2.53	304
BPL Programs	1.73	0.88	304
EGS	0.20	0.21	304
Income Programs	4.85	2.33	304
Non-Income Programs	0.51	0.28	304
Revenu (1)	149.8	360.8	229
Revenue (2)	9.7	27.6	318
Expenditure	8.9	25.4	318
BDO Meetings	3.22	6.31	319
MP Meetings	1.72	8.23	318
DC Meetings	1.26	4.59	319

Notes: Information on the caste of the Pradhan (the elected leader of the village government) and whether the position is reserved comes from our GP Questionnaire. Village population numbers, caste proportions, and caste land ownerhip patterns come from our Village Questionnaire. SC/ST refers to the Scheduled Castes and Schedule Tribes, the lowest ranking group in the caste hierarchy. Maratha land dominated is equal to one if a village is dominated by Marathas in terms of land ownership and equal to zero if instead the majority of landholdings in the village are in the hands of a lower caste. Distance to water, road, and rail come from GPS Census data. Whether there is a canal or river in the village comes from the 2001 Village Census. The three variables pertaining to topsoil (30 cm) content come from FAO-UNESCO soil maps. Rainfall information, which is only available at the district level, comes from the India Meteorological Department. Latitude, Longitude, and Elevation measures come from the GPS Census Data. Total Programs refers to the total number of the 15 Government Schemes implemented

in the village[33]. BPL refers to the number of the possible 8 programs targeted at individuals below the poverty line (these include: Housing Support Scheme; Sanitation Support Scheme; Indira Awas Yojana IAY, a housing construction program; Targeted Public Distribution System (TPDS)). EGS refers to the Employment Guarantee Scheme, which is the precursor to the present NREGA. Income programs is the set of programs that likely directly or indirectly affect household labour decisions and income (such as public good and housing improvement schemes which do rely on villagers' labour in their construction and the targeted public distribution system). Non-Income Programs refers to programs which do not directly affect household labour decisions such as those targeted towards children (child development (ICDS) and mid-day meals) and the elderly (social security pensions and foodgrains (Annapurna)). We obtained information on the availability of programs from our household survey and aggreated this information to the village level. Revenue (1) refers to data collected from the balance sheets of submitted by the GPs, these are computed per capita of the GP population. We obtained the majority of this information using the RTI Act. The information covers the last 24 months. Revenue (2) and Expenditure are annual per capita values from the 2001 Village Census. BDO (Block Development Officer) , MP (Member of Parliament), and DC (District Collector) meetings all refer to the number of times in the last year that the Gram Pradhan has met with officials from higher level governments to seek resources. This information is from the GP Questionnaire.

[33]The complete list of programs: Housing Support Scheme; Sanitation Support Scheme; EGS; SGYR (Sampoom Gram Rojgar Yojana); IAY (Indira Awas Yojana); SGSY (Swamjayanti Grameen Sawa Rogar Yojana); ICDS (Integrated Child Development Scheme); Social Security Pension, Mid-day Meal Program; ARWSP (Accelerated Rural Water Supply Program); PMGSY (Pradhan Mantri Gram Sadak Yojana); TPDS (Targeted Public Distribution System); Annapurna; Watershed Development Programs under DRAP and DDP; Total Sanitation Campaign; Swajaidhara; Business Support Program; FFW (Food for work program); PDS.

Table A2 - Summary Statistics - Household Measures

Variable	Mean	Standard Deviation	Observations
Maratha	0.38	0.49	8671
OBC	0.31	0.46	8671
SC/ST	0.29	0.45	8671
Landless	0.34	0.47	9136
Low Land (0-2.5 Acres)	0.21	0.41	9136
Large Land Owners (>5 Acres)	0.29	0.45	9136
Less than Primary School	0.28	0.45	8948
Insured (1)	0.63	0.48	9135
Insured (2)	0.61	0.49	9132
Insured (3)	0.62	0.49	9129
Insured (4)	0.64	0.48	9133
Insured (5)	0.63	0.48	9127
Insured (6)	0.66	0.47	9120
Insurer	0.79	0.41	9128
Voted-Personal	0.24	0.43	8266
Trust	2.57	0.95	9110
No Cheat	0.89	0.31	8772
Repair	0.36	0.48	9084
Donated Cash	0.52	0.50	9140
Donated Labour	0.29	0.45	9140
Agree	0.55	0.50	9133
Share Water	0.70	0.46	5093
Target Village	2.20	0.58	8964
Shared Funds	1.65	1.20	8329
Festivals	0.59	0.49	9132

Notes: All information comes from our Household Questionnaire. OBC refers to Other Backward Castes and SC/ST referes to Scheduled Castes and Scheduled Tribes. These caste groups are ranked below Marathas in the caste hierarchy, where the SC/ST category is the lowest ranked. Household education is measured by the highest level of education that any household male has reached. Less than primary refers to the highest category being less than primary school. Insured (1): "Would most people in your village help you with some money in times of crisis?". Insured (2): "Would a higher caste member of your village help you with some money in times of crisis?". Insured (3): "Would most people in your village help a lower caste villager with some money in times of crisis?". Insured (4)-(6) are the same questions with "money" replaced by "grain". Insurer: "Suppose a lower caste man asks to borrow a good sum of money from you because someone in his family has fallen ill. He is from the village and has the ability to repay the amount. Would you lend it to him?". Voted - Personal equals to one if the houshold voted for a candidate due to a personel connection rather than due to the characteristics of the candidate (honesty, good reputation, qualifications). Samples are conditional on voting. Trust is response to: "Would you say that the large landholders can be trusted? 1=Almost none, 2=Some; 3=Majority; 4=Almost. Cheat refers to answering somone from a higher caste is most likely to cheat you (compared to other castes or wealth levels). Repair is the answer to "If someone from your village noticed something wrong on your farm they would?" repair it themselves (compared to conditional answers, such as "alert you if he is from a lower caste....etc). Donated cash or labour are dummy variables equal to one if the household did donate (cash or labour respectively) in the past year to a development project within the village. Agree refers to answering that most people in the village would agree on the type of development project the village should have (compared to differences of opinions within the village). Share Water is equal to 1 if the household shares a water source with members of the Maratha caste. Samples

are conditional on sharing a water source. Target Village refers to GP funds should be targeted to the village as a whole, compared to poor or low caste individuals. Shared funds refers to GP funds are shared across the village (e.g. for development projects; public goods) compared to going directly to the poor or low status; the rich and high status; or to GP members or other government officials directly. Fesitvals is equal to one if there are village projects to finance festivals.

Table A3 - Summary Statistics - Household and Individual Measures

Variable	Mean	Standard Deviation	Observations
Daily Wage	41.49	15.96	15004
Male	0.55	0.50	15021
Illiterate	0.43	0.50	15014
Age	39.4	15.7	15007
Log Kharif Yields	8.91	1.17	5539
Log Kharif Profit	8.31	1.43	4269
Labour/Total Costs (Kharif)	0.31	0.23	5648
Household Members	5.51	2.63	9132
Jowar	0.45	0.50	5874
Rainfed	0.68	0.47	6105
Black Soil	0.53	0.50	6128
Clay Soil	0.64	0.48	6122
Salinity	0.18	0.39	6125
Percolation	0.29	0.45	6126
Drainage	0.29	0.46	6127
Maratha Trader	0.42	0.49	6341
Outside Maratha Trader	0.26	0.44	5945
Maratha Lender	0.43	0.49	901
Terms of Payment (Inputs)	1.02	0.47	21496
Interest Rate on Loan	20.0	22.1	920

Notes: The sample of labourers are all those who work for a daily wage in agriculture. The gender, literacy rate, and age of these workers are reported above. Yields, profits, and proportion of labour costs are all measured per acre of land. Kharif yields are the total value of output per acre of land for a given crop, summed over all of the kharif crops for each household. Kharif profit is yields net of input costs (seeds, fertilizer, irrigation, electricity, pesticides, and labour). Workers include partime and fulltime. Jowar is a dummy variable equal to one if the household grows this main staple crop. Rainfed refers to the percentage of land which is rainfed as opposed to irrigated. Black soil, clay soil, salinity, percolation, and drainage are all measures of the soil quality of the household land. Maratha Trader is equal to one if the household has traded with a Maratha for any tradeable good (which includes agricultural inputs and outputs, farm enterprise and non-farm enterprise goods) conditional on trading goods. Outside Maratha Trader refers to the trader residing outside of the village conditional on trading goods. Maratha Lender refers to lending money conditional on the household borrowing money. Terms of payments is an index variable equal to 0 if the trader requires advanced payments; 1 if full payment is required at the time of sale; and 2 if instead payment in installments is acceptable. Terms of payments and interest rate are reported per individual loan.

Table A4 - Characteristics by Caste

Variable	Marathas	OBCs	SCs
Cultivator (Overall)	0.83 (0.38)	0.65 (0.48)	0.33 (0.47)
Cultivator (Maratha Land Dominated)	0.84 (0.36)	0.66 (0.47)	0.35 (0.48)
Agricultural Labourer (Overall)	0.10 (0.30)	0.19 (0.40)	0.53 (0.50)
Agricultural Labourer (Maratha Land Dominated)	0.09 (0.29)	0.18 (0.38)	0.52 (0.50)
Landless (Overall)	0.13 (0.34)	0.31 (0.46)	0.62 (0.48)
Landless (Maratha Land Dominated)	0.12 (0.32)	0.29 (0.45)	0.60 (0.49)
Average Land Owned (Overall)	6.74 (7.38)	6.00 (6.22)	3.67 (3.28)
Average Land Owned (Maratha Land Dominated)	6.86 (7.53)	5.70 (5.96)	3.47 (2.99)
$>$ 5 Acres (Overall)	0.39 (0.49)	0.35 (0.48)	0.15 (0.36)
$>$ 5 Acres (Maratha Land Dominated)	0.40 (0.49)	0.33 (0.47)	0.14 (0.35)
$>$ 10 Acres (Overall)	0.16 (0.37)	0.13 (0.33)	0.03 (0.18)
$>$ 10 Acres (Maratha Land Dominated)	0.16 (0.37)	0.12 (0.33)	0.03 (0.18)
Voted	0.89 (0.31)	0.93 (0.26)	0.90 (0.30)
Supposed to Vote	0.08 (0.28)	0.09 (0.29)	0.10 (10.29)
Forced Vote	0.002 (0.04)	0.0004 (0.02)	0.001 (0.03)
Unopposed Election	0.09 (0.29)	0.04 (0.21)	0.07 (0.26)
Raise concerns to Pradhan	0.96 (0.19)	0.96 (0.18)	0.96 (0.18)
Met Pradhan	0.97 (0.16)	0.95 (0.22)	0.96 (0.20)
Observations	3259	2659	2019

Notes: Standard deviations are in parentheses. OBC refers to Other Backward Castes and SC/ST referes to Scheduled Castes and Scheduled Tribes. These caste groups are ranked below Marathas in the caste hierarchy, where the SC/ST category is the lowest ranked. Occupation categories (Cultivator and Agricultural Labourer) refer to main source of livelihood for household. Total land owned is in acres and are reported conditional on owning land. Voted refers to voted in the last GP election. Supposed to vote refers to "supposed to vote - does not mean anything". Forced vote refers to forced to vote for a given candidate by friends, family, or villagers. Unopposed election - refers to single candidate election (this was the main reason for not voting). The fourth variable is the response to "Do you feel you can raise concerns (bring oral requests) directly to the Gram Pradhan?"

Table A5- Control of Panchayat Measures

Variable	Overall	MLD	MLD Maratha Majority	MLD Non-Maratha Majority
Population Proportion of Marathas	0.41 (0.31)	0.54 (0.26)	0.71 (0.13)	0.28 (0.19)
Maratha Pradhan	0.41 (0.49)	0.57 (0.50)	0.64 (0.48)	0.45 (0.50)
Maratha Pradhan - Unreserved	0.63 (0.48)	0.82 (0.38)	0.94 (0.24)	0.64 (0.49)
Maratha Pradhan - Reserved for Women	0.62 (0.49)	0.89 (0.31)	0.95 (0.23)	0.78 (0.44)
Reserved Pradhan	0.58 (0.49)	0.57 (0.50)		
Reserved Pradhan - Women	0.27 (0.45)	0.26 (0.44)		
Reserved Pradhan - SC/ST	0.24 (0.43)	0.24 (0.43)		
Reserved Pradhan - OBC	0.49 (0.50)	0.50 (0.50)		
Proportion Reserved on GP	0.59 (0.19)	0.56 (0.16)		
Observations	315	193	120	73

Notes: MLD denotes Maratha Land Dominanted . Data on proportion of Marathas comes from the village survey. Data on characteristics of the Pradhan (elected leader of village government) come from the GP Survey. OBC refers to Other Backward Castes and SC/ST referes to Scheduled Castes and Scheduled Tribes. These caste groups are ranked below Marathas in the caste hierarchy, where the SC/ST category is the lowest ranked.

8 Appendix B: Alternative Estimations

8.1 Estimations without Village Controls

As a robustness check, we ran analogous estimations to those reported in Tables 2 to 5, which exclude all of the village-level controls. These estimation results are reported in Tables B1 and B2 below. We see that the main results on the significance of the coefficients on MLD and $MLD \cdot MPROP$, discussed in the paper, all go through. We have also run estimations which additionally exclude all of the household level controls in the household level regressions. Though not reported here, the main results are also robust to excluding these control variables as well.

Table B1 - Estimations of GP Measures without Village Controls

Dependent Variable	Coefficient (β_1) MLD	Coefficient (β_2) $MPROP$	Coefficient (β_3) $MLD \cdot MPROP$	Coefficient $\beta_1+\beta_3$	Observations
Maratha Pradhan	0.47 (0.13)***	0.93 (0.20)***	-0.50 (0.26)**	-0.03 (0.18)	292
All Programs (2)	-1.48 (0.53)***	-1.55 (0.86)*	1.82 (1.08)*	0.34 (0.76)	293
BPL Programs (2)	-0.58 (0.18)***	-0.51 (0.29)*	0.83 (0.37)**	0.25 (0.26)	293
EGS (2)	-0.10 (0.04)**	0.03 (0.10)	0.19 (0.11)*	0.09 (0.09)	293
Income Programs (2)	-1.42 (0.50)***	-1.49 (0.80)*	1.88 (1.01)*	0.45 (0.70)	293
Non-Income Programs (2)	-0.05 (0.05)	-0.06 (0.08)	-0.06 (0.10)	-0.12 (0.08)	293
Revenue (1)	-125.7 (61.8)**	-131.2 (198.2)	82.4 (208.0)	-43.3 (216.2)	222
Revenue (2)	-13.5 (4.9)***	-19.9 (6.5)***	27.9 (9.6)***	14.4 (5.1)	308
Expenditure	-12.4 (4.8)***	-17.8 (6.2)***	26.0 (9.4)***	13.6 (4.9)***	308
BDO Meetings	-2.35 (1.20)**	0.21 (3.86)	-0.09 (3.9)	-2.44 (3.32)	290
MP Meetings	-2.92 (1.46)**	-2.71 (1.76)	2.74 (1.75)	-0.18 (1.19)	290
DC Meetings	-2.19 (1.00)**	-3.91 (1.49)***	2.96 (1.42)**	0.77 (0.64)	290
Meetings (AES)	-2.5 (1.2)**	-2.15 (2.0)	1.9 (2.4)	-0.6 (1.7)	289

Notes: All estimations include regional fixed effects. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Robust standard errors are in parentheses. Acronyms used are: Maratha land dominated (MLD); Maratha popluation proportion (MPROP); Below Poverty Line (BPL); Employment Guarantee Scheme (EGS); Block Development Officer (BDO); District Collector (DC); Member of Parliament (MP); and Average Effect Size (AES). Programs (2); BPL Programs (2); EGS (2), Income Programs (2), and Non-Income Programs (2) are village level variables defined as in the notes to Table 2. Revenue (1) refers to data collected from the balance sheets (covers last 24 months) submitted by the GPs (obtained using RTI Act). Revenue (2) and Expenditure are annual per capita values from the 2001 Village Census. Meetings (AES) is the estimated average effect size of the three variables: BDO Meetings, DC Meetings, and MP Meetings.

Table B2 - Estimations of Household Measures without Village Controls

Dependent Variable	Coeff. (β_1) MLD	Coeff. (β_2) $MPROP$	Coeff (β_3) $MLD \cdot MPROP$	Coefficient $\beta_1 + \beta_3$	Obs
Insured (AES) [Landless]	0.11 (0.03)***	0.02 (0.06)	-0.18 (0.07)**	-0.06 (0.06)	2583
Insurer [Large Land Owners]	0.08 (0.02)***	0.06 (0.05)	-0.19 (0.06)***	-0.10 (0.05)**	2519
Daily Wage [All Labourers]	-2.06 (0.99)**	0.69 (2.0)	3.79 (2.36)*	1.73 (1.93)	13581
Daily Wage [Males]	-2.45 (1.18)**	-1.48 (2.31)	4.99 (2.87)*	2.54 (2.38)	7502
Log Kharif Yields	0.29 (0.12)***	0.23 (0.17)	-0.50 (0.23)**	-0.21 (0.16)	2334
Log Kharif Profit	0.31 (0.15)**	0.24 (0.28)	-0.82 (0.32)***	-0.50 (0.24)**	1849
Maratha Trader [Low Castes]	0.13 (0.04)***	0.22 (0.09)	-0.05 (0.11)	0.08 (0.09)	3025
Outside Maratha Trader [Low Castes]	0.11 (0.03)***	0.19 (0.06)***	-0.19 (0.08)***	-0.08 (0.06)	2800
Maratha Lender [Low Castes]	0.29 (0.07)***	0.47 (0.15)***	-0.23 (0.19)	0.06 (0.14)	454
Terms of Payment (Inputs) [Low]	0.09 (0.05)*	0.27 (0.10)***	-0.27 (0.11)**	-0.18 (0.09)*	10044
Interest Rate on Loan [Low Castes]	-8.60 (4.0)**	7.85 (8.70)	1.78 (11.54)	-6.81 (9.89)	252
Voted-Personal [Low Castes]	0.09 (0.03)***	0.15 (0.08)	-0.22 (0.10)**	-0.12 (0.08)	2121
Social Capital (AES) [Low Castes]	0.07 (0.01)***	0.03 (0.02)	-0.12 (0.03)***	-0.06 (0.02)**	4176
Share Water [Low Castes]	0.31 (0.06)***	0.55 (0.14)***	-0.34 (0.16)**	-0.02 (0.12)	2947
Target Village [Low Castes]	1.09 (0.31)***	0.70 (0.59)	-1.73 (0.72)***	-0.64 (0.52)	4888
Shared Funds [Low Castes]	0.95 (0.30)***	1.43 (0.65)**	-1.87 (0.75)***	-0.92 (0.62)	4608

Notes: All estimations include household-level controls (education, land ownership, and caste identity), and regional fixed effects. Regression disturbance terms are clustered at the village level. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Acronyms used are: Maratha land dominated (MLD); Maratha popluation proportion (MPROP); and Average Effect Size (AES). Insured (AES) is the estimated average effect size of the six variables, Insured (1) to (6), defined in the notes of Table 3. Insurer is a dummy variable equal to one if respondents answer yes to: "Suppose a lower caste man asks to borrow a good sum of money from you because someone in his family has fallen ill. He is from the village and has the ability to repay the amount. Would you lend it to him?". Large land owners have $\geq$ than 5 acres. The sample of labourers are all those who work for a daily wage in agriculture. Additional individual controls (gender, age, education) are included in the wage estimations. Regression disturbance terms are clustered at the household and village level for these estimations using the approach of Cameron, Gelbach and Miller (2011). The sample for the yields, profits, proportion of labour costs regressions is all large cultivators ($>$ 5 acres of land). All measures are per acre of land. Kharif yields are the total value of output per acre of land for a given crop, summed over all of the kharif crops for each household. Kharif profit is yields net of input costs (seeds, fertilizer, irrigation, electricity, pesticides, and labour). Workers include partime and fulltime, same results held if restricted ourselves to fulltime workers. Additional crop controls are included in the yields and profits estimations. Maratha Trader is equal to one if the household has traded with a Maratha for any tradeable good (which includes agricultural inputs and outputs, farm enterprise and non-farm enterprise goods) conditional on trading goods. Outside Maratha Trader refers to the trader residing outside of the village conditional on trading goods. Maratha Lender refers to lending money conditional on the household borrowing money. These estimations on Maratha trading relationships

are probit estimations, where the coefficients reported are the partial derivates of the predicted probability. Terms of payments is an index variable equal to 0 if the trader requires advanced payments; 1 if full payment is required at the time of sale; and 2 if instead payment in installments is acceptable. These are ordered probit estimations. Terms of payments and interest are reported per individual loan, for these estimations regression disturbance terms are clustered at the household and village level for these estimations using the approach of Cameron, Gelbach and Miller (2011). Voted - Personal equals to one if the houshold voted for a candidate due to a personel connection rather than due to the characteristics of the candidate (honesty, good reputation, qualifications). Samples are conditional on voting. The sample of low castes in the voting regressions is SC/STs. Social Capital (AES) is the estimated average effect size of the six variables: Trust, No Cheat, Repair, Donated Cash, Donated Labour, and Agree as defined in the notes of Table 5. Target Village refers to GP funds should be targeted to the village as a whole, compared to poor or low caste individuals. Shared funds refers to GP funds are shared across the village (e.g. for development projects; public goods) compared to going directly to the poor or low status; the rich and high status; or to GP members or other government officials directly. These two estimations are estimated as multinomial logit models.

8.2 Estimations with Village Controls Interacted with *MPROP*

As a robustness check, we ran analogous estimations to those reported in Tables 2 to 5 in the paper, which include additional interaction terms. In these estimations the key village-level variables, discussed in Table 1 (measuring presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, proportion of the population that is SC/ST as well as total village population) are interacted with $MPROP$. These estimation results are reported in Tables B3 and B4 below. We see that the main results on the significance of the coefficients on MLD and $MLD \cdot MPROP$, discussed in the paper, again all continue to hold.

Table B3 - Estimations of GP Measures with Village Controls Interacted with MPROP

Dependent Variable	Coefficient (β_1) MLD	Coefficient (β_2) $MPROP$	Coefficient (β_3) $MLD \cdot MPROP$	Coefficient $\beta_1 + \beta_3$	Observations
Maratha Pradhan	0.44 (0.14)***	0.83 (0.53)	-0.59 (0.28)**	-0.14 (0.19)	290
All Programs (2)	-1.27 (0.59)**	-2.22 (2.01)	2.02 (1.12)*	0.75 (0.75)	291
BPL Programs (2)	-0.43 (0.21)**	-0.35 (0.70)	0.77 (0.39)**	0.34 (0.26)	291
EGS (2)	-0.08 (0.04)*	0.22 (0.21)	0.21 (0.11)**	0.14 (0.08)	291
Income Programs (2)	-1.25 (0.55)**	-2.17 (1.91)	2.12 (1.05)**	0.86 (0.70)	291
Non-Income Programs (2)	-0.02 (0.05)	-0.05 (0.18)	-0.10 (0.11)	-0.11 (0.07)	291
Revenue (1)	-188.1 (112.7)*	165.0 (328.3)	146.8 (276.4)	-41.3 (233.7)	220
Revenue (2)	-13.1 (6.7)***	-20.6 (31.1)	30.2 (12.8)***	17.1 (7.0)***	307
Expenditure	-12.6 (6.6)**	-17.6 (31.1)	29.0 (12.6)**	16.4 (6.8)**	307
Meetings (AES)	-1.2 (1.1)	5.9 (4.2)	-0.4 (2.1)	-1.6 (1.5)	289

Notes: All estimations include village-level controls (presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, proportion of the population that is SC/ST, and total village population) and regional fixed effects. The estimations also include interaction terms between each of these village-level controls and $MPROP$. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Robust standard errors

are in parentheses. Acronyms used are: Maratha land dominated (MLD); Maratha popluation proportion (MPROP); Below Poverty Line (BPL); Employment Guarantee Scheme (EGS); and Average Effect Size (AES). Programs (2); BPL Programs (2); EGS (2), Income Programs (2), and Non-Income Programs (2) are village level variables defined as in the notes to Table 2. Revenue (1) refers to data collected from the balance sheets (covers last 24 months) submitted by the GPs (obtained using RTI Act). Revenue (2) and Expenditure are annual per capita values from the 2001 Village Census. Meetings (AES) is the estimated average effect size of the three variables: BDO Meetings, DC Meetings, and MP Meetings as defined in the notes to Table 2.

Table B4 - Estimations of Household Measures with Village Controls Interacted with MPROP

Dependent Variable	Coeff. (β_1) MLD	Coeff. (β_2) $MPROP$	Coeff (β_3) $MLD \cdot MPROP$	Coefficient $\beta_1 + \beta_3$	Obs
Insured (AES) [Landless]	0.09 (0.04)***	-0.04 (0.14)	-0.18 (0.08)**	-0.08 (0.06)	2565
Insurer [Large Land Owners]	0.05 (0.02)*	0.07 (0.12)	-0.14 (0.06)**	-0.09 (0.05)*	2501
Daily Wage [All Labourers]	-2.03 (1.0)**	-7.03 (4.3)	4.19 (2.14)**	2.16 (1.50)	13546
Daily Wage [Males]	-2.22 (1.26)*	-7.02 (5.58)	5.22 (2.72)**	3.0 (1/95)	7480
Log Kharif Yields	0.24 (0.14)*	-0.24 (0.51)	-0.43 (0.25)*	-0.18 (0.15)	2320
Log Kharif Profit	0.29 (0.18)*	-0.55 (0.61)	-0.77 (0.37)**	-0.47 (0.24)**	1838
Maratha Trader [Low Castes]	0.12 (0.04)***	0.09 (0.21)	-0.06 (0.11)	0.06 (0.08)	3012
Outside Maratha Trader [Low Castes]	0.11 (0.03)***	0.09 (0.14)	-0.19 (0.08)**	-0.08 (0.06)	2793
Maratha Lender [Low Castes]	0.20 (0.08)***	-0.50 (0.37)	-0.21 (0.20)	-0.02 (0.15)	452
Terms of Payment (Inputs) [Low]	0.10 (0.05)**	0.05 (0.15)	-0.25 (0.11)**	-0.15 (0.09)*	10034
Interest Rate on Loan [Low Castes]	-8.48 (4.0)**	2.77 (17.9)	-5.67 (12.95)	-14.15 (11.48)	250
Voted-Personal [Low Castes]	0.09 (0.04)**	0.12 (0.19)	-0.15 (0.11)	-0.06 (0.09)	2108
Social Capital (AES) [Low Castes]	0.05 (0.01)***	-0.002 (0.07)	-0.12 (0.03)***	-0.07 (0.03)***	4693
Share Water [Low Castes]	0.35 (0.06)***	0.84 (0.25)***	-0.46 (0.14)***	-0.11 (0.10)	2929
Target Village [Low Castes]	0.85 (0.37)**	-0.27 (1.41)	-1.48 (0.84)*	-0.63 (0.60)	4865
Shared Funds [Low Castes]	1.24 (0.35)***	2.11 (1.58)	-2.78 (0.83)***	-1.55 (0.66)	4584

Notes: All estimations include village-level controls (presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, proportion of the population that is SC/ST, and total village population), household-level controls (education, land ownership, and caste identity), and regional fixed effects. The estimations also include interaction terms between each of these village-level controls and $MPROP$. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Acronyms used are: Maratha land dominated (MLD); Maratha popluation proportion (MPROP); and Average Effect Size (AES). Insured (AES) is the estimated average effect size of the six variables, Insured (1) to (6), defined in the notes of Table 3. Insurer is a dummy variable equal to one if respondents answer yes to: "Suppose a lower caste man asks to borrow a good sum of money from you because someone in his family has fallen ill. He is from the village and has the ability to repay the amount. Would you lend it to him?". Large land owners have $\geq$ than 5 acres. The sample of labourers are all those who work for a daily wage in agriculture. Additional individual controls (gender, age, education) are included in the wage estimations. Regression disturbance terms are clustered at the household and village level

for these estimations using the approach of Cameron, Gelbach and Miller (2011). The sample for the yields, profits, proportion of labour costs regressions is all large cultivators ($>$ 5 acres of land). All measures are per acre of land. Kharif yields are the total value of output per acre of land for a given crop, summed over all of the kharif crops for each household. Kharif profit is yields net of input costs (seeds, fertilizer, irrigation, electricity, pesticides, and labour). Workers include partime and fulltime, same results held if restricted ourselves to fulltime workers. Additional crop controls are included in the yields and profits estimations. Maratha Trader is equal to one if the household has traded with a Maratha for any tradeable good (which includes agricultural inputs and outputs, farm enterprise and non-farm enterprise goods) conditional on trading goods. Outside Maratha Trader refers to the trader residing outside of the village conditional on trading goods. Maratha Lender refers to lending money conditional on the household borrowing money. These estimations on Maratha trading relationships are probit estimations, where the coefficients reported are the partial derivates of the predicted probability. Terms of payments is an index variable equal to 0 if the trader requires advanced payments; 1 if full payment is required at the time of sale; and 2 if instead payment in installments is acceptable. These are ordered probit estimations. Terms of payments and interest are reported per individual loan, for these estimations regression disturbance terms are clustered at the household and village level for these estimations using the approach of Cameron, Gelbach and Miller (2011). Voted - Personal equals to one if the houshold voted for a candidate due to a personel connection rather than due to the characteristics of the candidate (honesty, good reputation, qualifications). Samples are conditional on voting. The sample of low castes in the voting regressions is SC/STs. Social Capital (AES) is the estimated average effect size of the six variables: Trust, No Cheat, Repair, Donated Cash, Donated Labour, and Agree as defined in the notes of Table 5. Target Village refers to GP funds should be targeted to the village as a whole, compared to poor or low caste individuals. Shared funds refers to GP funds are shared across the village (e.g. for development projects; public goods) compared to going directly to the poor or low status; the rich and high status; or to GP members or other government officials directly. These two estimations are estimated as multinomial logit models.

8.3 Estimations with Maratha Land Holdings

We now report the results from analagous regressions to those estimated in the paper. Here, instead of using a binary variable, MLD which equals 1 if Marathas are the land dominant group, and 0 otherwise, as we did in the paper, we use an alternative source of information on Maratha land holdings from our household surveys. From these 30 households per village we obtain an estimate of the overall proportion of village lands held by Marathas $\widetilde{MLD} \in [0, 1]$ for each village. Refer to Section 10.1 in Appendix D for more details on this variable.

These estimation results are reported in Tables B5 and B6 below. Importantly, the main results discussed in the paper all go through as well in these alternative estimations.

Table B5 - Estimations of GP Measures with Maratha Land Holdings

Dependent Variable	Coefficient (β_1) $\widetilde{MLD}$	Coefficient (β_2) $MPROP$	Coefficient (β_3) $\widetilde{MLD}\cdot MPROP$	Coefficient $\beta_1+\beta_3$	Observations
Maratha Pradhan	0.61 (0.26)**	1.16 (0.22)***	-0.71 (0.40)*	-0.10 (0.23)	275
Revenue (1)	-117.6 (89.6)	-119.4 (83.4)	107.7 (133.2)	-9.87 (84.1)	193
Revenue (2)	-20.7 (9.1)**	–20.5 (8.3)***	49.4 (19.1)***	28.7 (11.2)	291
Expenditure	-19.1 (9.1)**	-19.5 (8.2)***	46.6 (19.0)***	27.5 (11.1)	291
Programs (1)	-1.42 (0.65)**	-1.84 (0.58)***	3.63 (1.09)***	2.21 (0.65)***	7752
BPL Programs (1)	-0.60 (0.25)**	-0.56 (0.21)***	1.46 (0.42)***	0.86 (0.25)***	7752
EGS (1)	-0.16 (0.08)**	-0.05 (0.07)	0.36 (0.12)***	0.20 (0.08)	7725
Income Programs (1)	-1.40 (0.61)	-1.73 (0.54)***	3.53 (1.02)***	2.13 (0.61)***	7752
Programs (2)	-1.70 (0.72)***	-2.00 (0.61)***	3.80 (0.83)***	2.10 (0.70)***	275
BPL Programs (2)	-0.70 (0.28)***	-0.62 (0.21)***	1.50 (0.45)***	0.80 (0.26)***	275
EGS (2)	-0.19 (0.08)***	-0.04 (0.07)	0.39 (0.14)***	0.20 (0.09)	275
Income Programs (2)	-1.67 (0.68)	-1.90 (0.56)***	3.73 (1.10)***	2.07 (0.65)***	275
BDO Meetings	-4.70 (1.64)***	0.64 (2.58)	4.75 (3.61)	0.05 (2.58)	275
MP Meetings	-4.92 (2.86)*	-1.81 (2.30)	5.87 (3.04)**	0.96 (1.72)	275
DC Meetings	-3.94 (1.28)***	-3.51 (1.52)**	5.83 (1.98)***	1.90 (1.10)*	275

Notes: $\widetilde{MLD}$ refers to the proportion of village land that is owned by Marathas constructed from the household level data. All estimations include village-level controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and whether the GP is reserved) and regional fixed effects. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Robust standard errors are in parentheses. Acronyms used are: Maratha popluation proportion (MPROP); Below Poverty Line (BPL); Employment Guarantee Scheme (EGS); Block Development Officer (BDO); District Collector (DC); and Member of Parliament (MP). Revenue (1) refers to data collected from the balance sheets (covers last 24 months) submitted by the GPs (obtained using RTI Act). Revenue (2) and Expenditure are annual per capita values from the 2001 Village Census. Information on programs (Programs (1); BPL Programs (1); EGS (1); Income Programs (1); Non-Income Programs) are reported from household level data, and regression disturbance terms are clustered at the village level for these estimations and household level controls are also included (education, land ownership, and caste identity). Programs (2); BPL Programs (2); EGS (2), Income Programs (2), and Non-Income Programs (2) are variables which aggregate this household level information up to the village level. Estimations are OLS except EGS (1), which is a probit estimation, where the reported coefficients are the partial derivates of the predicted probability.

Table B6 - Estimations of Household Measures with Maratha Land Holdings

Dependent Variable	Coeff. (β_1) $\widetilde{MLD}$	Coeff. (β_2) $MPROP$	Coeff (β_3) $\widetilde{MLD}\cdot MPROP$	Coefficient $\beta_1+\beta_3$	Obs
Insured (1) [Landless]	0.19 (0.08)**	0.11 (0.06)*	-0.31 (0.13)**	-0.13 (0.08)	2579
Insured (2) [Landless]	0.19 (0.08)**	0.09 (0.07)	-0.35 (0.13)***	-0.16 (0.08)	2579
Insured (3) [Landless]	0.20 (0.08)***	0.10 (0.07)	-0.33 (0.13)***	-0.13 (0.09)	2579
Insured (4) [Landless]	0.24 (0.08)***	0.12 (0.07)*	-0.35 (0.14)***	-0.11 (0.09)	2579
Insured (5) [Landless]	0.26 (0.08)***	0.07 (0.07)	-0.37 (0.13)***	-0.11 (0.09)	2579
Insured (6) [Landless]	0.24 (0.08)***	0.08 (0.07)	-0.38 (0.13)***	-0.14 (0.09)	2579
Insurer [Large land owners]	0.18 (0.05)***	-0.02 (0.04)	-0.19 (0.07)***	-0.01 (0.04)	2507
Daily Wage (All Labourers)	-1.65 (0.82)**	-1.84 (0.85)**	6.70 (1.46)***	5.05 (0.97)***	13581
Daily Wage (Low Castes)	-2.25 (0.96)**	-3.25 (1.00)***	8.60 (1.85)***	6.35 (1.31)***	9195
Log Kharif Yields	0.34 (0.18)**	-0.02 (0.21)	-0.47 (0.27)*	-0.13 (0.20)	2323
Maratha Trader [Low Castes]	0.23 (0.08)***	0.17 (0.08)**	-0.11 (0.15)	0.11 (0.10)	3021
Outside Maratha Trader [Low Castes]	0.17 (0.06)***	0.11 (0.06)*	-0.24 (0.11)**	-0.07 (0.07)	2800
Maratha Lender [Low Castes]	0.50 (0.18)***	0.39 (0.17)**	-0.44 (0.32)	0.06 (0.23)	453
Interest Rate on Loan [Low Castes]	-29.1 (10.1)***	2.30 (11.1)	24.7 (19.1)	-4.43 (14.85)	165
Voted-Personal [Low Castes]	0.17 (0.08)**	0.05 (0.07)	-0.20 (0.13)†	-0.04 (0.09)	2116
Social Capital (AES) [Low Castes]	0.10 (0.03)***	0.02 (0.02)	-0.13 (0.05)***	-0.04 (0.03)	4711
Share Water [Low Castes]	0.38 (0.11)***	0.47 (0.11)***	-0.37 (0.18)**	0.01 (0.11)	2942
Target Village [Low Castes]	1.57 (0.57)***	-0.07 (0.62)	-1.93 (1.0)**	-0.36 (0.68)	4883
Shared Funds [Low Castes]	1.17 (0.58)**	0.70 (0.60)	-1.04 (0.99)	0.12 (0.65)	4603

Notes : $\widetilde{MLD}$ refers to the proportion of village land that is owned by Marathas, constructed from the household level data. All estimations include village-level controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and whether the GP is reserved), household-level controls (education, land ownership, and caste identity), and regional fixed effects. Regression disturbance terms are clustered at the village level. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Acronyms used are: Maratha popluation proportion (MPROP) and Average Effect Size (AES). Insured (1): "Would most people in your village help you with some money in times of crisis?". Insured (2): "Would a higher caste member of your village help you with some money in times of crisis?". Insured (3): "Would most people in your village help a lower caste villager with some money in times of crisis?". Insured (4)-(6) are the same questions with "money" replaced by "grain". Insurer: "Suppose a lower caste man asks to borrow a good sum of money from you because someone in his family has fallen ill. He is from the village and has the ability to repay the amount. Would you lend it to him?". Large land owners have $\geq$ than 5 acres. The sample of labourers are all those who work for a daily wage in agriculture. Additional individual controls (gender, age, education) are included in the wage estimations. Regression disturbance terms are clustered at the household and village level for these estimations using the approach of Cameron, Gelbach and Miller (2011). The sample for the yields and profits regressions is all large cultivators ($>$ 5 acres of land). All measures are per acre of land. Kharif yields are the total value of output per acre of land for a given crop, summed over all of the

kharif crops for each household. Kharif profit is yields net of input costs (seeds, fertilizer, irrigation, electricity, pesticides, and labour). Additional crop controls are included in the yields and profits estimations. Maratha Trader is equal to one if the household has traded with a Maratha for any tradeable good (which includes agricultural inputs and outputs, farm enterprise and non-farm enterprise goods) conditional on trading goods. Outside Maratha Trader refers to the trader residing outside of the village conditional on trading goods. Maratha Lender refers to lending money conditional on the household borrowing money. These estimations on Maratha trading relationships are probit estimations, where the coefficients reported are the partial derivates of the predicted probability. Terms of payments is an index variable equal to 0 if the trader requires advanced payments; 1 if full payment is required at the time of sale; and 2 if instead payment in installments is acceptable. These are ordered probit estimations. Terms of payments and interest are reported per individual loan, for these estimations regression disturbance terms are clustered at the household and village level for these estimations using the approach of Cameron, Gelbach and Miller (2011). Voted - Personal equals to one if the houshold voted for a candidate due to a personel connection rather than due to the characteristics of the candidate (honesty, good reputation, qualifications). Samples are conditional on voting. The sample of low castes in the voting regressions is SC/STs. Social Capital (AES) is the estimated average effect size of the six variables: Trust, No Cheat, Repair, Donated Cash, Donated Labour, and Agree as defined in the notes of Table 5. Target Village refers to GP funds should be targeted to the village as a whole, compared to poor or low caste individuals. Shared funds refers to GP funds are shared across the village (e.g. for development projects; public goods) compared to going directly to the poor or low status; the rich and high status; or to GP members or other government officials directly. These two estimations are estimated as multinomial logit models.

8.4 Additional Estimations

Table B7 - Other Outcomes - Landless Sample

Dependent Variable	Coeff. (β_1) MLD	Coeff. (β_2) $MPROP$	Coeff (β_3) $MLD \cdot MPROP$	Observations
Pradhan - Honesty	0.05 (0.12)	-0.28 (0.24)	0.26 (0.28)	2573
Pradhan - Qualifications	0.006 (0.13)	-0.27 (0.22)	0.44 (0.28)	2564
Pradhan - Providing Public Goods	0.08 (0.11)	-0.26 (0.18)	0.45 (0.23)*	2569
Pradhan - Solving Problems	0.001 (0.11)	-0.19 (0.21)	0.43 (0.25)*	2567
Pradhan - Allocating Spending Fairly	0.01 (0.10)	-0.26 (0.24)	0.44 (0.27)	2565
Pradhan - Acquiring Funds	0.16 (0.13)	-0.09 (0.20)	0.29 (0.26)	2566
Pradhan - Not Misusing Funds	-0.14 (0.12)	-0.43 (0.19)**	0.50 (0.26)*	2560
Pradhan - Representing Village to Govt.	-0.06 (0.13)	0.01 (0.17)	0.42 (0.23)*	2566
Share Information	-0.03 (0.03)	-0.07 (0.06)	-0.002 (0.07)	2571
Trust Neighbours	0.01 (0.07)	0.15 (0.13)	-0.22 (0.17)	2582
Trust Own Caste	0.09 (0.08)	0.20 (0.16)	-0.29 (0.18)	2582

Notes: All estimations include village-level controls (latitude, longitude, elevation, presence of river/canal, distance to natural water sources, distance to railways and national roads, soil quality measures, rainfall levels, proportion of the population that is SC/ST, total village population, and whether the GP is reserved), household-level controls (education, land ownership, and caste identity), and regional fixed effects. Regression disturbance terms are clustered at the village level. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. Acronyms used are: Maratha land dominated (MLD) and Maratha popluation proportion (MPROP). The sample is landless. The dependent variables for the Pradhan variables are dummy variables equal to one if respondents answered a low rank in terms of their confidence in their Pradhan with regards to the listed characteristics, and zero otherwise. Estimations are probits, where the coefficients reported are the partial derivates of the predicted probability. "Share Information" refers to a question which asked "Suppose you find employment available at a good wage that others do not know about it, who would you share the information with?", this variable is equal to one if they would share it with everyone in the village as opposed to just their close family and friends. Trust Neighbours is response to: "Would you say that your neighbours can be trusted? 1=Almost none, 2=Some; 3=Majority; 4=Almost. Trust Own Caste is response to: "Would you say that members of your own caste can be trusted? 1=Almost none, 2=Some; 3=Majority; 4=Almost.

9 Appendix C: Additional Theoretical Results

9.1 Wages and Yields Affected by Programs and EGS

Most landless individuals sell their labor to large landowners. Most large landowners have as their largest input cost labor. The way labor relations work in these villages is that the landless people or small landholders who rely mainly on labor income for their livelihood typically work on the farm of a large landowner in a permanent or semi-permanent capacity. Much of what workers need to do can only be partly or very imperfectly supervised, suggesting that asymmetries of information in production may arise. Such permanent working arrangements are coveted by workers, and though there is a spot market for some labor, it seems that workers prefer the permanent working arrangements greatly. The threat of losing such employment disciplines the use of discretionary effort. For individuals primarily relying on labor income for their livelihood the threat of employment loss, which would send them into poverty, provides great incentive for them to keep contributing un- or partially monitored discretionary effort in their employment on large landholders farms. Large landholders grow various crops and their labor needs, timing of application, and other inputs use are largely fixed through the crop cycle. However, the quality of crop obtained depends critically on good labor input and diligence through the production process.

These ingredients suggest an efficiency wage model. Workers are required, by the implicit contract of the landlord, to provide e^* units of labor effort and receive a wage w^*. Landlords imperfectly ascertain, ex post, the effort contribution of their worker and decide whether to rehire them in the next period, or dismiss them from their employ. Since production is largely of a fixed factor variety, we can for simplicity simply characterize the optimal incentive compatible contract (e^*, w^*) offered to each worker by the landlord while letting the landlord's landholding and crop choice (which is a function of the conditions) determine the number of workers required.

In this sort of labor market, even though much of the year sees labor only partially employed or unemployed the activities of the panchayat in providing poverty alleviation programs become significant. In the event that workers are not employed by landlords, they will depend on benefits from the state, or on employment from the state for their livelihood. Thus, we can characterize their reservation utility, $\overline{u}$, as depending positively on the incidence of these programs. For simplicity let this take two values, $\overline{u}(W)$ when W workers control the panchayat and actively seeks out these programs, and $\overline{u}(L)$ when L landlords control it and such programs are shut down. These are taken as given when worker and landlord play the labor/production game.

9.1.1 The Labor/Production game

Given an increasing and concave per worker effort production function, $f(e)$, the landlord chooses the implicit contract parameters (w, e):

$$\max_{e,w} f(e) - w$$

subject to (w, e) being incentive compatible for the worker. That is any pair w, e chosen must satisfy

$$\frac{u(w) - c(e)}{1-r} \geq u(w) + \frac{r}{1-r}\overline{u}(x), \text{ where } x = W \text{ or } L. \tag{14}$$

The term $u(w)$ is increasing and concave, $c(e)$ is increasing and convex, and $\overline{u}(x)$ reservation employment if dismissed, is increasing in probability of obtaining benefits, probability of obtaining EGS employment, and probability of obtaining another job (which we can set equal to zero for simplicity), so that $\overline{u}(W) > \overline{u}(L)$.

Firstly note that any optimal e, w chosen must ensure that (14) binds exactly,

$$\frac{u(w) - c(e)}{1-r} = u(w) + \frac{r}{1-r}\overline{u}(x)$$

implying.

$$w = u^{-1}\left(\frac{c(e)}{r} + \overline{u}(x)\right). \tag{15}$$

Substituting this in, the optimization problem becomes:

$$\max_e f(e) - u^{-1}\left(\frac{c(e)}{r} + \overline{u}(x)\right).$$

With a FOC that implies:

$$f'(e) = u^{-1\prime}\left(\frac{c(e)}{r} + \overline{u}(x)\right)\frac{c'(e)}{r}.$$

This implicitly defines a solution $e^*(\overline{u}(x))$ and from equation (15) the corresponding w^*.

Proposition The optimal implicit contract (w^*, e^*) has wage strictly increasing and effort strictly decreasing in $\overline{u}(x)$.

Proof: At e^* :

$$f^{'}(e^*(\overline{u})) = u^{-1\prime}\left(\frac{c(e^*(\overline{u}, r))}{r} + \overline{u}\right)\frac{c'(e^*(\overline{u}, r))}{r}.$$

Differentiating with respect to $\overline{u}$ yields:

$$\begin{aligned} f^{''}(e^*(\overline{u}, r))\frac{de^*}{d\overline{u}} &= u^{-1\prime\prime}(\cdot)\left(\frac{c'(e^*(\overline{u}, r))}{r}\frac{de^*}{d\overline{u}} + 1\right)\frac{c'(e^*(\overline{u}, r))}{r} \\ &\quad + u^{-1\prime}(\cdot)\left(\frac{c''(e^*(\overline{u}, r))}{r}\frac{de^*}{d\overline{u}}\right), \end{aligned}$$

rearranging:

$$\frac{de^*}{d\overline{u}} = \frac{u^{-1\prime\prime}(\cdot)\frac{c'(e^*(\overline{u}, r))}{r}}{f^{''}(e^*(\overline{u}, r)) - u^{-1\prime}(\cdot)\frac{c''(e^*(\overline{u}, r))}{r} - u^{-1\prime\prime}(\cdot)\left(\frac{c'(e^*(\overline{u}, r))}{r}\right)^2}.$$

Because $u(\cdot)$ is an increasing and concave function, $u^{-1}(\cdot)$ is an increasing and convex function. Then since $c(\cdot)$ is a convex function by assumption it is immediate that the terms on the RHS can be signed as follows:

$$\frac{de^*}{d\overline{u}} = \frac{[+]}{[-] - [+] - [+]} < 0.$$

Differentiating equation (15) with respect to $\overline{u}$ yields:

$$\begin{aligned} sign\left[\frac{dw}{d\overline{u}}\right] &= sign\left[\begin{array}{c} u^{-1\prime\prime}(\cdot)\left(\frac{c'(e^*(\overline{u},r))}{r}\right)^2 + f^{''}\left(e^*\left(\overline{u},r\right)\right) - u^{-1\prime}(\cdot)\frac{c''(e^*(\overline{u},r))}{r} \\ -u^{-1\prime\prime}(\cdot)\left(\frac{c'(e^*(\overline{u},r))}{r}\right)^2 \\ \hline [-] \end{array}\right] \\ &= sign\left[\frac{f^{''}\left(e^*\left(\overline{u},r\right)\right) - u^{-1\prime}(\cdot)\frac{c''(e^*(\overline{u},r))}{r}}{[-]}\right] \\ &> 0. \end{aligned}$$

Prediction Where GPs are controlled by landlords, wages should be lower and effort should be higher across the village. $w(L) < w(W)$ and $e(L) > e(W)$.

9.2 Proofs of propositions

Proof of Proposition 1

Using (3) and (4), vote trading is individually rational for worker j in village k, $U_{jk}(L_i) \geq U_{jk}(W)$, if and only if, $S_i^j \geq \Delta_{wP} + \left(d_j - d_i^j\right)T - x_k$. Since landlords transfer no more than necessary to buy a vote, S_i^j is chosen so that this condition binds. The following conditions are thus also sufficient to ensure incentive compatibility and individual rationality:

$$S_i^j = \Delta_{wP} + \left(d_j - d_i^j\right)T - x_k. \tag{16}$$

Substituting (16) into the landlord's incentive constraint (1) yields:

$$\Delta_{wP} + \left(d_j - d_j^i\right)T - x_k \leq X_i + I_i^j X + x_i^j. \tag{17}$$

There are three specific cases of condition (17) to consider. For a Maratha worker and landlord pair, $c_i = c_j = M$, we have:

$$x_i^j + x_k \quad \geq \quad \Delta_{wP} - (X_M + X) \equiv x_{MM}. \tag{18}$$

For a non-Maratha worker, $c_j = N$, and Maratha landlord, $c_i = M$:

$$x_i^j + x_k \quad \geq \quad \Delta_{wP} - T - X_M \equiv x_{NM}. \tag{19}$$

For Non-Maratha landlords with either type of worker:

$$x_i^j + x_k \quad \geq \quad \Delta_{wP} - X_N \equiv x_{NN} \equiv x_{MN}. \tag{20}$$

Which correspond to the conditions in the statement of the proposition.

Proof of Proposition 2

Let $P_{VT}(k)$ denote the proportion of workers willing to undertake vote trading in village k, for a given x_k we then have:

$$P_{VT}(k) = \sigma_{MMk}\Pr(x_i^j > x_{MM} - x_k) + \sigma_{MNk}\Pr(x_i^j > x_{MN} - x_k) + \sigma_{NNk}\Pr(x_i^j > x_{NN} - x_k) + \sigma_{NMk}\Pr(x_i^j > x_{NM} - x_k)$$

By Assumption 3

$$I_{VTk} = \begin{cases} 0 \text{ if } P_{VT}(k) < 0.5 \\ 1 \text{ if } P_{VT}(k) \geq 0.5. \end{cases}$$

So that $E[I_{VTk}] = \Pr(P_{VT}(k) < 0.5|\sigma_{ijk}) \cdot 0 + \Pr(P_{VT}(k) \geq 0.5|\sigma_{ijk}) \cdot 1 \equiv \Pr(P_{VT}(k) \geq 0.5|\sigma_{ijk})$. Thus, using (8) we obtain:

$$E[v_k|\mathbf{Z_k}] = \alpha_v \Pr(P_{VT}(k) \geq 0.5) + \alpha \mathbf{Z_k} + \mu_{vk}, \tag{21}$$

We now show that if x_k is drawn from a uniform distribution, then the probability of clientelism occurring in village k is given by:

$$\Pr(P_{VT}(k) > 0.5|\,\sigma_{ijk}) = \sigma_{MMk}(X_M + X) + \sigma_{NMk}(T + X_M) + \sigma_{NNk}(X_N) + \sigma_{MNk}(X_N) + C, \tag{22}$$

where C is a constant.

Assume that x_i^j is randomly drawn from a uniform distribution $F(x_i^j)$ on the interval $[0, \overline{x}]$. We see below that the form of this distribution is irrelevant so this is without loss of generality. Given this, the $Pr(x_i^j > z) \equiv \frac{\overline{x}-z}{\overline{x}}$ for any $z \in [0, \overline{x}]$. Applying the law of large numbers and substituting for x_i^j from equations (5), (6) and (7) yields the proportion of vote traders in village k conditional upon candidate quality x_k:

$$\begin{aligned} P_{VT}(k) \;=\; & \sigma_{MMk}\left(\frac{\overline{x} - \Delta_{wP} + X_M + X + x_k}{\overline{x}}\right) + \sigma_{NMk}\left(\frac{\overline{x} - \Delta_{wP} + T + X_M + x_k}{\overline{x}}\right) \\ & + \sigma_{NNk}\left(\frac{\overline{x} - \Delta_{wP} + X_N + x_k}{\overline{x}}\right) + \sigma_{MNk}\left(\frac{\overline{x} - \Delta_{wP} + X_N + x_k}{\overline{x}}\right), \end{aligned} \tag{23}$$

which, since the σ_{ij}s sum to one, rearranges to:

$$P_{VT}(k) = \sigma_{MMk}\left(\frac{X_M + X}{\overline{x}}\right) + \sigma_{NMk}\left(\frac{T + X_M}{\overline{x}}\right) + \sigma_{NNk}\left(\frac{X_N}{\overline{x}}\right) + \sigma_{MNk}\left(\frac{X_N}{\overline{x}}\right) + \frac{\overline{x} - \Delta_{wP} + x_k}{\overline{x}}, \tag{24}$$

Since we assume the law of large numbers within the village, the first moment of the distribution determines the number of vote traders arising from within each of the landlord/worker groups fully. Hence the invariance with respect to the distribution of the x_i^j.

For clientelism in village k, necessarily $P_{VT}(k) > 0.5$, which is equivalent to:

$$x_k > 0.5\overline{x} - \overline{x} + \Delta_{wP} - \sigma_{MMk}(X_M + X) - \sigma_{NMk}(T + X_M) - \sigma_{NNk}(X_N) - \sigma_{MNk}(X_N). \tag{25}$$

Let $G(x)$ denote the uniform distribution of the x_k, with support $[0, 1]$. An additional interval restriction on the variable $\overline{x}$ derives from the assumed support of $G()$ and is required for the likelihood of clientelism

to be a well defined probability. It is that $2(\Delta_{wp} - T - X_M) \geq \overline{x} \geq 2(\Delta_{wp} - 1 - X_N)$. The probability that $P_{VT}(k) > 0.5$, and hence the probability of clientelism, then lies in the unit interval and is given by $(1 - RHS(25))$, which rearranges to equation (22) above with $C = 1 + 0.5\overline{x} - \Delta_{wP}$.

Now substituting using (22) into (21) yields:

$$E[v_k|\sigma_{ijk}, \mathbf{Z_k}] = \alpha_v[C + \sigma_{MMk}(X_M + X) + \sigma_{NMk}(T + X_M) + \sigma_{NNk}(X_N) + \sigma_{MNk}(X_N)] + \alpha\mathbf{Z_k} + \mu_{vk}, \quad (26)$$

which is the statement in the proposition.

Proof of Proposition 3

Under random matching: $\sigma_{MMk} = MLD_k \cdot MPROP_k, \sigma_{NMk} = MLD_k \cdot (1 - MPROP_k), \sigma_{NNk} = (1 - MLD_k) \cdot (1 - MPROP_k)$ and $\sigma_{MNk} = (1 - MLD_k) \cdot MPROP_k$. Substituting these into equation (9) and rearranging yields the expression in the statement of the proposition.

9.3 Generalized Distributions

Here we demonstrate invariance with respect to the distributional assumptions we made for computing the likelihood of vote trading in Section 3.2. Recalling that $P_{VT}(k)$ denotes the proportion of workers willing to undertake vote trading in village k, we have:

$$P_{VT}(k) = S_{MMk}\Pr(x_i^j > x_{MM} - x_k) + S_{MNk}\Pr(x_i^j > x_{MN} - x_k) + S_{NNk}\Pr(x_i^j > x_{NN} - x_k) + S_{NMk}\Pr(x_i^j > x_{NM} - x_k)$$

This again applies the law of large numbers within the village. Substituting for the x_i^j from equations (5), (6) and (7), we have:

$$\begin{aligned} P_{VT}(k) &= \sigma_{MMk}\Pr(x_i^j > \Delta_{wP} - (X_M + X) - x_k) + \sigma_{NMk}\Pr(x_i^j > \Delta_{wP} - T - X_M - x_k) + \\ \sigma_{NNk}\Pr(x_i^j &> \Delta_{wP} - X_N - x_k) + \sigma_{MNk}\Pr(x_i^j > \Delta_{wP} - X_N - x_k). \end{aligned}$$

Substituting for the σ_{ij}s using Assumption 4:

$$\begin{aligned} &P_{VT}(k) = MLD_k \cdot MPROP_k \Pr(x_i^j > \Delta_{wP} - (X_M + X) - x_k) + \\ &MLD_k \cdot (1 - MPROP_k)\Pr(x_i^j > \Delta_{wP} - T - X_M - x_k) + \\ &(1 - MLD_k) \cdot (1 - MPROP_k)\, Pr.(x_i^j > \Delta_{wP} - X_N - x_k) + (1 - MLD_k) \cdot MPROP_k \Pr(x_i^j > \Delta_{wP} - X_N - x_k). \end{aligned}$$

Rearranging:

$$\begin{aligned} P_{VT}(k) &= MLD_k \cdot MPROP_k \left[\Pr(x_i^j > \Delta_{wP} - X_M - X - x_k) - \Pr(x_i^j > \Delta_{wP} - T - X_M - x_k)\right] \\ &+ MLD_k \left[\Pr(x_i^j > \Delta_{wP} - T - X_M - x_k) - \Pr(x_i^j > \Delta_{wP} - X_N - x_k)\right] \\ &+ MPROP_k\,[0] \end{aligned}$$

Without the assumption of uniformity on F and G, but instead persisting with general distributions, we cannot specify the linear form of this equation in Section 3.2, but can instead express this equation directly in terms of the respective CDFs. Substituting from the CDF for F.

$$\begin{aligned} P_{VT}(k) &= MLD_k \cdot MPROP_k \left[1 - F(\Delta_{wP} - X_M - X - x_k) - 1 + F(\Delta_{wP} - T - X_M - x_k)\right] \\ &\quad + MLD_k \left[1 - F(\Delta_{wP} - T - X_M - x_k) - 1 + F(\Delta_{wP} - X_N - x_k)\right] \\ &\quad + MPROP_k [0] \end{aligned}$$

Since, from Assumption 3, we have vote trading in village k if and only if $P_{VT}(k) > 0.5$, then the probability of vote trading is

$$\Pr\left(P_{VT}(k) > 0.5\right) \equiv \Pr\begin{pmatrix} \text{MLD}_k \cdot MPROP_k \left[F(\Delta_{wP} - T - X_M - x_k) - F(\Delta_{wP} - X_M - X - x_k)\right] \\ +\text{MLD}_k \left[F(\Delta_{wP} - X_N - x_k) - F(\Delta_{wP} - T - X_M - x_k)\right] \\ +\text{MPROP}_k [0] \qquad > \qquad 0.5 \end{pmatrix}$$

The law of large numbers removes the idiosyncratic (pairwise between landlord and worker) uncertainty, the remaining uncertainty in this equation (the outer probability) is driven by the stochastic nature of the candidate quality variable, x_k. This is determined by the distribution G. Using the equation before (9) we have:

$$E[v_k|\mathbf{Z_k}] = \alpha_v \Pr(P_{VT}(k) \geq 0.5) + \alpha \mathbf{Z_k} + \mu_{vk}$$

Now insert this object into the estimating equation to obtain, instead of (9):

$$\begin{aligned} E[v_k|\mathbf{Z_k}] &= \alpha_v \Pr\begin{pmatrix} \text{MLD}_k \cdot MPROP_k \left[F(\Delta_{wP} - T - X_M - x_k) - F(\Delta_{wP} - X_M - X - x_k)\right] \\ +\text{MLD}_k \left[F(\Delta_{wP} - X_N - x_k) - F(\Delta_{wP} - T - X_M - x_k)\right] \\ +\text{MPROP}_k [0] \qquad > \qquad 0.5 \end{pmatrix} + \\ &\quad \alpha \mathbf{Z_k} + \mu_{vk} \end{aligned}$$

The outer probability arises from the random draw on x_k. The sign restrictions on coefficients and interpretations obtained from a linear regression as specified in equation (12) or (13) are the same as in the body of the paper. The fact that the distribution of G is no longer uniform does not affect the interpretations, as the precise form of this distribution simply governs the draws of x_k. The effect of the model parameters T, X, X_M and X_N is to simply shift the required cutoff level of x_k (for a given x_i^j) beyond which vote trading is preferred for a landlord/worker pair. The shifts vary by the type of pairing, so that the relative magnitudes of these parameters is identified by observing differences in the outcome variables (the vs) as a function of the distribution of landlord/worker pairings (the σ_{ij} that we infer from $MPROP$ and MLD) in a village.

To see that the sign interpretations are the same when using the equation immediately above, consider each coefficient in turn: The coefficient on $MPROP_k$ is again predicted to be zero. Since this variable is multiplied by zero, its magnitude in the village will not affect the likelihood that the proportion of vote

traders is above a half. For variable v such that $\alpha_v > 0 (< 0)$, the coefficient on MLD_k is again predicted positive (negative) since $T > 0$ and $X_M - X_N > 0$. Intuitively, vote trading is more likely if there are Maratha landlords since they possess two benefits in clientelism relative to non-Maratha landlords – access to trading networks and more powerful social cohesion. Consequently, the larger proportion of Maratha landlords, the larger the number of worker/landlord pairs tipped into vote trading by the existence of these clientelist benefits. The coefficient on the interaction $MLD_k \cdot MPROP_k$ again allows us to identify the relative power of trading network access and social cohesion. It is again positive iff $X_M + X > X_M + T$ or $X > T$. So that, for $\alpha_v > 0, MLD_k \cdot MPROP_k > (<) 0$ implies $X > (<) T$. The sign of $[coeff] MLD_k \cdot MPROP_k + [coeff] MLD_k > 0$ iff $F(\Delta_{wP} - X_N - x_k) > F(\Delta_{wP} - X_M - X - x_k)$, i.e., iff. $X + X_M > X_N$. So the sign of this expression again yields an upper bound for the benefits of Maratha social cohesion relative to social cohesion in non-Maratha groups $X_M - X_N$.

9.4 Workers Less able to Punish Landlords in MLD

Assume now a violation of Assumption 1 where workers in MLD villages are less able to punish a promise violating landlord than in non-MLD villages. In non-MLD villages then continue to assume that x_i^j is drawn from distribution $F(\cdot)$ which continues to be defined over support $[0, \overline{x}]$, but now suppose that x_i^j in MLD villages is drawn from $\widehat{F}(\cdot)$ defined over $[0 - \zeta, \overline{x} - \zeta]$ with $\zeta > 0$. $\widehat{F}(\cdot)$ is first-order stochastically dominated by F. Equation (23) now becomes:

$$\begin{aligned} P_{VT}(k) &= \sigma_{MMk}\left(\frac{\overline{x} - \zeta - \Delta_{wP} + X_M + X + x_k}{\overline{x}}\right) + \sigma_{NMk}\left(\frac{\overline{x} - \zeta - \Delta_{wP} + T + X_M + x_k}{\overline{x}}\right) \\ &\quad + \sigma_{NNk}\left(\frac{\overline{x} - \Delta_{wP} + X_N + x_k}{\overline{x}}\right) + \sigma_{MNk}\left(\frac{\overline{x} - \Delta_{wP} + X_N + x_k}{\overline{x}}\right), \end{aligned} \tag{27}$$

and equation (24) now rearranges to:

$$P_{VT}(k) = \sigma_{MMk}\left(\frac{X_M + X - \zeta}{\overline{x}}\right) + \sigma_{NMk}\left(\frac{T + X_M - \zeta}{\overline{x}}\right) + \sigma_{NNk}\left(\frac{X_N}{\overline{x}}\right) + \sigma_{MNk}\left(\frac{X_N}{\overline{x}}\right) + \frac{\overline{x} - \Delta_{wP} + x_k}{\overline{x}}, \tag{28}$$

Using (28) substituting for the σs and proceeding identically, expression (10) of Proposition 3 now rearranges to:

$$\begin{aligned} E[v_k | MLD_k, MPROP_k, \mathbf{Z_k}] &= \alpha_v C + MLD_k \cdot \alpha_v [T + X_M - X_N - \zeta] + MPROP_k \cdot \alpha_v [0] \\ &\quad + MLD_k \cdot MPROP_k \cdot \alpha_v [X - T] + \alpha \mathbf{Z_k} + \mu_{vk}, \end{aligned} \tag{29}$$

where C is a constant. This expression thus yields identical implications to those previously obtained under expression (10) except that now the predicted coefficient on MLD in the baseline model is biased towards zero by the ζ term. This implies that the interpretation provided by the model for the sign of the coefficient on MLD, given that it is significant and in the direction predicted by the baseline model, is unchanged. The predictions on all other coefficients are also unchanged. There is now, however, a slight difference in interpretation of the magnitude of the coefficient on MLD under this extension of the model – it is reduced

by the ζ which represents less worker cohesion in MLD villages. However, since we have already argued that the normalizations imposed in the model render interpretation of coefficients directly in terms of model parameters meaningless, this has no effect.

9.5 Landlord Candidates Less Appealing in MLD

Assumption 2 of the baseline model asserts that whether landlords are Maratha or not does not affect the (dis)utility that workers obtain from lending political support to the landlord's candidate in Panchayat elections. Consider a violation of this assumption where, in MLD villages, villagers resent landlords more than in non-MLD villages, perhaps due to their always having been the dominant caste, then the disutility workers experience by supporting an otherwise identical landlord's candidate may be worse in the MLD case. A simple modification of the model captures this. Continue to assume that the variable x_k, capturing the cost to workers from voting for a landlord's candidate is drawn from distribution $G(\cdot)$ in all villages. Now, however the realization of the term x_k has subtracted from it the amount (in utility metric) ι in MLD villages. This implies that the same quality draw of candidate from $G(\cdot)$ generates ι less utility for voters if this candidate is drawn in a MLD villages rather than in a non-MLD village. Equation (23) now becomes:

$$\begin{aligned} P_{VT}(k) &= \sigma_{MMk}\left(\frac{\overline{x}-\Delta_{wP}+X_M+X+x_k-\iota}{\overline{x}}\right)+\sigma_{NMk}\left(\frac{\overline{x}-\Delta_{wP}+T+X_M+x_k-\iota}{\overline{x}}\right) \\ &\quad +\sigma_{NNk}\left(\frac{\overline{x}-\Delta_{wP}+X_N+x_k}{\overline{x}}\right)+\sigma_{MNk}\left(\frac{\overline{x}-\Delta_{wP}+X_N+x_k}{\overline{x}}\right). \end{aligned} \tag{30}$$

But notice that this modification of the model is equivalent to the modification introduced in Section 9.4 above and yields an equation that is identical to equation (27) above. This violation of Assumption 2 is thus isomorphic in its effect to the violation of Assumption 1 explored there. The interpretation of coefficients under the model is identical.

9.6 Positive Assortative Matching

Assumption 4 of the baseline model asserts random matching, so that caste wise matching simply reflects population frequencies. Now assume that Maratha landlords are matched with Maratha workers at rate μ times their relative frequency. Hence with the proportion of Maratha workers in the village denoted $MPROP$ the probability of a single Maratha landlord being matched with a Maratha worker is $MPROP \cdot \mu$. This implies that the frequency of Maratha landlord/Maratha worker pairs will be given by $MLD \cdot MPROP \cdot \mu = \sigma_{MM}$ as opposed to $MLD \cdot MPROP$ from the baseline model. Similarly for the other frequencies: $\sigma_{NM} = MLD \cdot (1 - MPROP \cdot \mu)$, $\sigma_{NN} = (1-MLD) \cdot (1-MPROP) \cdot \mu$ and $\sigma_{MN} = (1-MLD) \cdot [1-(1-MPROP) \cdot \mu]$.

Substitute for these values of σ_{ij} into equation (9 to obtain:

$$\begin{aligned} E[v_k|MLD_k, MPROP_k, \mathbf{Z_k}] &= \alpha_v \begin{pmatrix} C + MLD_k \cdot MPROP_k \cdot \mu(X_M + X) + \\ MLD_k \cdot (1 - MPROP_k \cdot \mu)(T + X_M) \\ +(1 - MLD_k) \cdot (1 - MPROP_k) \cdot \mu(X_N) + \\ (1 - MLD_k) \cdot (1 - [1 - MPROP_k] \cdot \mu)(X_N) \end{pmatrix} \\ &+ \alpha \mathbf{Z_k} + \mu_{vk} \end{aligned}$$

Rearranging:

$$\begin{aligned} E[v_k|MLD_k, MPROP_k, \mathbf{Z_k}] &= \alpha_v C + MLD_k \cdot \alpha_v [T + X_M - X_N] + MPROP_k \cdot \alpha_v [0] \\ &+ MLD_k \cdot MPROP_k \cdot \mu \cdot \alpha_v [X - T] + \alpha \mathbf{Z_k} + \mu_{vk} \end{aligned} \tag{31}$$

This is equivalent to the expression in the baseline model (equation 10) without assortative matching up to the positive multiplicative term μ multiplying the coefficient on the interaction term $MLD \cdot MPROP$. Clearly this term does not affect the sign of the coefficient on the interaction term, and therefore does not affect our interpretation of the relative sizes of X and T.

10 Appendix D: Independence of MLD_k and $MPROP_k$

Our two key measures are Maratha population numbers, $MPROP_k$, and landholdings, MLD_k, both at the village level. In what follows we discuss how these measures were obtained, and argue that both of these measures are historically pre-determined, and importantly not endogenous to our outcome variables.

10.1 Dominance Measures

Both Maratha population ($MPROP_k$) and landholdings (MLD_k) were collected in the village surveys. These surveys were conducted on a focus group discussion model, which included key villagers such as members of the GPs and upper level governments (particularly the *Gram Sevak and Talathi*) as well as school teachers and health care workers. The *Gram Sevak* represents the development wing of the state government and is well versed with the village population, since all of the benefit applications go through him. He, or members of the GP, readily provided the population numbers by caste group in the villages. The *Talathi*, who is from the Revenue department, is responsible for keeping and updating all land records. It was typically the Talathi who provided us with a ranking of total land ownership by caste group (at the sub-caste or *jati* level) in the villages. Both the Talathi and Gram Sevak are members of the higher levels of government and do not usually reside in the surveyed village.

We can create an alternative measure of Maratha land dominance ($\widetilde{MLD}_k$) using our household surveys, where we collected information on land ownership (refer to Section 8.3 in Appendix B). We can aggregate this data up to the village level to obtain a measure of land distribution by caste group at the village level. Since only 30 households per village were surveyed, these measures are quite noisy. Nevertheless, if we

construct a Maratha land dominance variable from this household level data, it matches our village level data (obtained from the Talathi) for 85% of villages. For those 15% of villages which did not match, the total land ownership of the top two ranked (in terms of land ownership) castes was very close using the household level data. In these cases, according to our village level data, Marathas were typically the second ranked caste in terms of land ownership. In other words, these were villages where two castes were fairly close in terms of their total land ownership, and this explains why the noisy household level data did not match up perfectly to the village level data.

In our baseline empirical analysis we use the village level data to construct our measure of Maratha land dominance (MLD_k). Results are robust to instead using the alternative aggregate measure ($\widetilde{MLD}_k$), constructed from our household level data, as we have reported in Appendix B.

10.2 Distribution of Caste Groups

We have no direct proof that caste population numbers are historically pre-determined at the village level, and not the consequence of any of our subsequent outcome variables, because no historical records reside at the village level on caste population numbers. However, at the district level, others have exploited the temporal invariance of caste numbers, and used caste composition measures from the historical census to predict outcomes today (Banerjee and Somanathan 2007). The assumption of time invariant caste distributions at the village level has also been exploited in other states of India (see Anderson (2011) for Uttar Pradesh and Bihar). Similarly, using the historical censuses of India (1891 - 1931), we can compare the relative population distribution of Marathas across the districts in our sample to the distribution in our current data. Despite our sampling of only non-tribal rural areas, the historical census variation closely matches the current variation found in our data. Of particular note is the virtual absence of Marathas in the most eastern part of the state (East Vidarbha). This part of the state was part of the Central Provinces in colonial times, a region where the Rajput caste were traditionally dominant.[34] All of the empirical results that we report are robust to excluding this region in our estimations.

In 41% of villages, Marathas form the majority of the population, but we see considerable village level variation in caste population numbers, which is the norm in India. Villages are typically multi-caste and rarely identical in either the number of castes or in the numerical strength (Srinivas 1987, Mandelbaum 1970, Marriott 1955). In general, Indian village anthropological studies link the origins of distributions at the village level back hundreds of years (Srinivas 1987, Mandelbaum 1970, Marriott 1955), and the Marathas are no exception.[35] The early settlement of the original tribes that grew into the prominent caste groups

[34]The present state of Maharashtra came into being in 1960. The state unites the Marathi speaking people (who have existed for centuries). During British rule, Marathi speakers were geographically divided between Bombay Presidency, Central Provinces and Berar, and the Nizam's state of Hyderabad. After Independence (1947) they continued as respective parts of these states until the formation of the bilingual state of Bombay in 1956 (two languages Marathi and Gujarati). The unilingual state of Maharashtra was formed in 1960.

[35]For example, in the case of Palanpur, a village in western Uttar Pradesh, events which took place some two hundred years ago explain the dominance of an upper caste group (Dreze et. al. 1999). Another village level study in northwest Uttar Pradesh

in Maharashtra dates to the 6th century BC (Kosambi 1955). The prominence of Marathas in the region dates back to at least the fourth century AD (Altekar 1927, Kosambi 1969).[36]

In our Village survey, we asked directly about the historical origin of caste groups in the village. In more than 95% of cases, the caste groups were reported to have resided in the village since well before independence. A possible concern is the possibility of migration in response to contemporary governance and economic outcomes, which would in turn directly alter village level caste composition. At the individual or household level, these concerns are not warranted. Firstly, this is almost unheard of in our sample of villagers. Secondly, given the strict rules governing hereditary caste rankings, there is virtually no mobility of individuals across different caste groups. Moreover, there is very little migration in India as a whole; see Munshi and Rosenzweig (2005) for an extensive analysis. This seems to be primarily because of reliance on sub-caste networks of mutual insurance that do not seem to cross village boundaries. At the caste level as a whole, there is no evidence of large scale migrations that could explain the variation in caste population dominance that we observe today.[37]

10.3 Land Ownership of Caste Groups

Marathas own the most land in 59% of the villages of our survey. Throughout history, Marathas have been the dominant land owners in Maharashtra and their prominence has been traced back to at least the fourth century AD when major chieftainships were under their control (Altekar 1927, Kosambi 1969). With respect to landholdings, their documented dominance of landowning extends back to at least the fourteenth century. Prior to independence, under either foreign rule or during their own Maratha empire, Marathas were the dominant land owners.[38] Under both Muslim and British regimes, land was allocated to Marathas by outside rulers to buy the loyalty of dominant lineages, and in return for supply of armies (Altekar 1927, Kosambi 1969, Drekmeier 1962, Dahiwale 1995). During colonial rule, the regions of present-day Maharashtra fell under different administrative units and systems of land revenue collection.[39] However,

dates the origins of present caste composition to more than 600 years ago (Danda 1987).

[36]Basic elements of the village organization, the *balutedari* system, were developed by the fourth century AD. This system was a reciprocal arrangement between the hereditary farming and artisan castes (OBCs in today's classification), service castes (SCs), and the higher landholding castes.

[37]With the exception of the movement of a small population of Brahmins from rural to urban areas in the early 20th century. They are less than 1% of our sample, and it is this exodus of Brahmins from rural areas that further solidified the dominance of Marathas in this region.

[38]Under the leadership of Chhatrapati Shivaji, the Maratha Empire was founded in 1674. At it's height in the 18th century, the empire extended from present-day Pakistan to Orissa in the east and from Punjab to central Karnataka in the south. It also included Tamil Nadu. The vast empire was in decline by 1818 when Maharahstra had fallen to the British East India Company, however remnants of it lasted until Independence in 1947.

[39]In particular, Western Maharashtra was part of the Bombay Presidency which had a *ryotwari* (cultivator-based) system of land revenue collection. Eastern Vidarbha was part of Central Provinces which had a *zamindari* (landlord-based) system. Western Vidarbha was a part of Berar, formerly part of the princely state of Hyderabad, which was given to the East India Company as a debt payment in 1860 and made into a ryotwari region at that time. Marathawada never fell under British rule and remained a part of the princely state of Hyderabad until Independence in 1947. Land there was divided between government and feudal ownership. The former was run similarly to the ryotwari system whereas the latter was more similar to the landlord system. Refer to Banerjee and Iyer (2004) who analyse the impact of these different land systems on outcomes

irrespective of the land revenue system used, Marathas continued to own the large majority of agricultural land. (This is documented in the *Imperial Gazetteers of India* which report the break down of caste land ownership patterns at the district level.[40]).

Upon Independence, Indian states legislated large scale land reforms. In Maharashtra, the Tenancy and Agricultural Lands Act of 1948 placed a ceiling on all landholdings and transferred ownership rights to tenant cultivators. These acts effectively redistributed land from large land owners to their former permanent tenants ("Other Backward Castes" or OBCs under today's classification). This lead to a dramatic change in ownership (but not cultivation) patterns.[41] These land reforms thus represent a striking break with the past. They gave rise to a new class of landowners drawn from a previously non-landowning caste. The land reforms thus fully account for villages where a non-Maratha caste are the largest landowning group in our sample.

Since the reforms, other changes in land ownership and distribution have been almost entirely due to the process of inheritance and partition (land is typically divided amongst sons), with the combined ownership of each dynasty remaining fairly constant. Formal sales of land are rare. In our sample less than 2% had bought or sold land within the past 5 years (almost all distress sales) and 86% of our sample of landowners report that they inherited their land. Almost 12% report that they purchased some of their land, but this was, in almost all cases, a purchase from a relative or co-caste member.

This settlement history, and the fact that land reforms managed the redistribution of large landlord holdings ensures a distinct pattern of caste and land ownership in Maharashtrian villages today.[42] In villages where few Marathas reside, the dominant land-owning caste *can* be a low caste (OBCs, former tenants). In villages where Marathas are populous, although the lower castes typically also own some land, Marathas are highly likely to constitute the dominant landowning caste.

Marathas may own the highest quality land today because they historically chose to reside in the high quality land villages, and ran the lower quality land as absentee landlords. This is an issue we already discussed in Section **??**. There it was demonstrated, in Table 1, that there are no significant differences in village land use patterns and soil quality measures across Maratha land dominated villages compared to others. Table 1 also demonstrates no significant differences across Maratha and lower (OBC) caste land dominated villages in key demographic and geographic variables. There are no significant differences in terms of total population numbers, proportion SC (the lowest ranked caste group), cultivatability of the

today. Our estimation results include regressors which control for these different land revenue systems.

[40]The relevant publications are *Imperial Gazetteer of India, Provincial Series (1909)* for *Bombay Presidency; Hyderabad State; Central Provinces;* and *Berar*.

[41]Maharashtra is one of the few states where the agricultural lands acts were comprehensively and successfully implemented, effectively granting of ownership rights to former tenants. Land ceilings were sometimes circumvented via transfers to extended family members, but land redistributions away from intermediaries and absentee landlords were highly effective. (Kamat 1980).

[42]Anderson (2011) similarly treats land dominance by caste groups at the village level as pre-determined using data from Uttar Pradesh and Bihar. The empirical strategy used here is also related to Besley and Burgess (2000) who estimate the impact of state-level land reforms on outcomes today.

land, rainfall patterns, and also distance to exogenous (to the GP responsibilities) measures of amenities.[43] These include distance to a national main road, major rivers, and the nearest town. We checked these differences using our own household and village level data and also using the Village Amenities data from the Census of India 2001.

References

[1] Altekar, A. S. (1927) *A History of Village Communities in Western India*. Bombay: Oxford University Press.

[2] Anderson, Siwan (2011) "Caste as an Impediment to Trade" *American Economic Journal: Applied Economics*, 3(1), 239-63.

[3] Banerjee, A. and L. Iyer (2005) "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India". *American Economic Review*, 95(4), 1190-1213.

[4] Banerjee, Abhijit and Somanathan, Rohini (2007). "The political economy of public goods: Some evidence from India," *Journal of Development Economics*, 82(2), 287-314.

[5] Besley, Timothy and Robin Burgess (2000). "Land Reform, Poverty Reduction, And Growth: Evidence From India" *Quarterly Journal of Economics*, 115(2), 389-430.

[6] Danda, Ajit K. (1987) "Ethnic Status: Identity and hierarchy" *Journal of Indian Anthropological Society*, 22(1), 52-64.

[7] Dahiwale, S. M. (1995) "Consolidation of Maratha Dominance in Maharashtra" *Economic and Political Weekly*, 30(6), 336-342.

[8] Drekmeier, Charles (1962) *Kingship and Community in Early India*. Stanford: Stanford University Press.

[9] Dreze, Jean, Peter Lanjouw, and Naresh Sharma (1999) "Economic development in Palanpur, 1957-93" in *Economic Development of Palanpur Over Five Decades*, ed. Peter Lanjouw and Nicholas Stern, 114-242. Oxford University Press.

[10] Kamat, A. R. (1980) "Politico-Economic Developments in Maharashtra: A Review of the Post-Independence Period" *Economic and Political Weekly*, 15(39), 1627-1630.

[11] Kosambi, D. D. (1955) "The basis of ancient Indian History, I" *Journal of the American Oriental Society*, 74(4), 330-337.

[43] The SC group, Scheduled Castes, are the lowest ranked castes, formerly known as the untouchable castes. They are ranked lower than the OBC category which refers to the backward caste groups. OBCs are the traditional farming and artisan castes while SCs traditionally performed menial tasks not directly related to agriculture, and hence they were not direct recipients of land during reforms.

[12] Kosambi, D. D. (1969) *Ancient India: A History of its Culture and Civilization.* New York: World Publishing Company.

[13] Mandelbaum, David, G. (1970) *Society in India*, University of California Press.

[14] Marriott, M. (1955) *Village India: Studies in Little Community.* Bombay, Asia Publishing House.

[15] Munshi, K. and M. Rosenzweig (2005) "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth" BREAD Working Paper No. 97.

[16] Srinivas, M.N. (1987) *The Dominant Caste and Other Essays.* New Delhi, Oxford University Press.