

# Parochial Politics: Ethnic Preferences and Politician Corruption

Abhijit V. Banerjee and Rohini Pande\*

## Abstract

This paper examines how increased voter ethnicization, defined as greater voter preference for the party representing her ethnic group, affects legislator quality. In situations where parties and politicians cannot commit to policies prior to the election, ethnicization reduces average winner quality for the pro-majority party, with the opposite true for the minority party. Overall, the average winner-loser quality gap reduces. These effects increase with greater numerical dominance of the majority and are absent in jurisdictions with equal-sized voter groups. Empirical evidence from a survey on politician corruption in North India is remarkably consistent with our theoretical predictions.

## 1 Introduction

We vote and you rule, this won't work anymore.

*-Campaign slogan of BSP, an Indian low caste party*

This paper sets out to make a rather elementary point: If voters care about politicians' group identity, be it gender, class, or ethnicity, either for itself or because of expectations

---

\*The authors are from MIT and Harvard respectively. We are grateful to Rasika Duggal and, especially, Santosh Kumar and Bhartendu Trivedi for organizing the survey. We thank Alberto Alesina, Esther Duflo, Michael Greenstone, Seema Jayachandran, Phil Keefer, David Laitin, Dominic Leggett, Peter Rosendorff, Ashutosh Varshney, members of PIEP, NBER political economy group and numerous seminar participants for comments. Pande thanks NSF for financial support under grant SES-0417634

about the policies they will select, then the party associated with the numerically dominant group in a jurisdiction enjoys a competitive advantage. Its candidates will win even when along other dimensions – competence, probity, or what, for want of better word, we will call quality – they are not quite as good.

This simple observation has an important corollary: In settings with a numerically dominant population group, a strengthening of the influence of group identity on citizens’ political preferences worsens the quality of political representation. This is for two reasons: first, the probability that the winner is from the party representing the dominant group goes up; second, the quality threshold at which s/he can win goes down.

In this paper, we develop a simple model of political competition with two parties to formalize this argument and test its predictions using micro-data on the corruption record of Indian politicians. In doing so, we provide an explanation for why, as seen in Figure 1, ethnic politics and a relatively high incidence of politician corruption dominate the political landscape of many democracies.<sup>1</sup>

In our model, all citizens prefer higher quality candidates but differ in their policy preferences, which we assume reflects differences in group identity. There are two parties. Each party has a fixed supply of candidates who are associated with the same policy. Therefore, each party selects the highest quality candidate. However, the highest quality available to a party in a particular jurisdiction at a given point in time is a random variable, and the two parties may end up with candidates of differing quality. As a result, to have a candidate who delivers a specific policy, voters might have to sacrifice quality. We examine how this trade-off changes as we vary two parameters of the model – the numerical dominance of the majority population group (*bias of the jurisdiction*), and the extent of own-group preference among voters (given our empirical application, we call this *voter ethnicization*).

---

<sup>1</sup>Our data on ethnic politics comes from 2007 Freedom House country reports. A country gets a “1” ranking if no ethnic political parties, or ethnic-based discrimination was mentioned. The ranking reduces as ethnic politics become more important with the lowest ranking of “0.2” given if ethnic-political ties dominate political parties. Ethnic politics exist in 52 of the 137 countries in our sample, and are more prevalent among low income countries.

We show that an increase in either own-group preference or bias worsens the quality of winners belonging to the numerically dominant group (the opposite is true of minority party winners) and reduces the average winner-loser quality gap. Further, the effects of own-group preference on both winner quality and the winner-loser quality gap are absent in jurisdictions with no bias, i.e. with equally sized voter groups. In contrast, in jurisdictions with positive bias, the effect of own-group preference increases with the extent of bias.

To test these predictions, we use panel data on politician quality from over a hundred jurisdictions in India's largest state, Uttar Pradesh (UP). We measure the bias of a jurisdiction by its ethnic composition and a politician's quality by his/her propensity to engage in corrupt practices. Using these data, we examine whether the impact of voter ethnicization on politician corruption differs with the bias of the jurisdiction.

We employ a difference-in-difference-in-difference (DDD) empirical strategy which exploits three sources of variation. The first is the dramatic rise in ethnic politics and voter ethnicization (meaning caste in this case) between 1980 and 1996. The second is significant cross-jurisdiction demographic variation, which implies that different caste groups (numerically) dominate in different jurisdictions. The third source of variation differs by outcome of interest. For winner quality, we exploit variation in whether the winning party in a jurisdiction represents the numerically dominant group (the theory predicts different outcomes for majority, and minority, party winners), and for the winner-loser quality gap, we exploit variation in who wins. All regressions include jurisdiction and year fixed effects to control for unobserved jurisdiction characteristics and over-time trends. Further, the winner quality regressions control for party and ethnic-group specific time trends, while the winner-loser gap regressions include fully flexible jurisdiction specific time effects.

The sharpest source of identification, however, comes from the non-linearity that our model generates: For both winner quality and the winner-loser gap in quality, we predict that voter ethnicization should not affect outcomes in jurisdictions with (roughly) equal sized groups since competitive pressure remain unchanged. On the other hand, we expect

similar effects when a group is dominant, *irrespective of the group's identity*. By directly testing for and finding strong support for this, we show that our results are not simply an artifact of party-specific time effects which trend (linearly) with jurisdiction demographics.

Our results strongly support the proposed theory, and the magnitude of the identified effects of increased ethnicization on politician corruption are relatively large. At least along some dimensions, the entire increase in legislator corruption in our sample jurisdictions between 1980 and 1996 can be attributed to increased corruption among legislators from the party that shared the ethnic identity of the dominant population group in that jurisdiction. Further, the increase in corruption is largely concentrated in jurisdictions with substantial high-caste or low-caste domination.

The structure of this paper is as follows: Section 2 positions this research within the existing literature. Section 3 uses a model of political competition to identify how voter ethnicization reduces politician quality. Section 4 describes the institutional context, data-sets, and empirical strategy. Section 5 provides the results, and Section 6 concludes.

## 2 Related Literature

In the canonical one-dimensional political economy model, varying the extent of preference diversity in the population does not alter a party's electoral incentive to choose the median voter's preferred outcome. However, as a number of recent papers show, this logic does not extend to multiple dimensions. Below we describe how our paper builds on and extends the insights afforded by these papers. We conclude the section by relating our paper to the broader empirical literature on ethnic diversity and quality of governance.

A first set of papers examine the case where voters pay taxes that are used to finance alternative forms of spending that may favor different groups. Alesina et al. (1999) show that if, one, citizens first vote on how much taxes to collect and then on how to spend it and, two, all citizens face identical taxes, then greater preference diversity lowers public good provision. However, if we drop the requirement of identical taxes, then one large

population group may use directed transfers to sustain a coalition with a smaller group with very different preferences. This creates, in effect, a continuum of policy options that specifies who pays how much in taxes and what the taxes are used to buy (Levy, 2005; Fernandez and Levy, 2008; Lizzeri and Persico, 2005). Fernandez and Levy (2008) show that if, relative to the extent of potential diversity in voter preferences, the set of available policies is small, then increased diversity first increases targeted redistribution (because there is someone to target). But beyond a critical level of diversity, there is a move back to general redistribution (targeting becomes too inefficient). In Lizzeri and Persico (2005), voters are identical, but diversity emerges endogenously as parties position themselves (in terms of how they will spend tax money) to segment the voting population into targeted and non-targeted groups. Here, as the number of parties goes up, the provision of the non-targeted public good decreases monotonically.

Here, increasing the numerical dominance of any single group while holding the number of groups (i.e., diversity) constant will tend to improve provision by increasing the likelihood that the median voter for taxes and public good provision is drawn from the same group (except in Lizzeri and Persico (2005), where citizens are identical). Our model, instead, predicts that politician quality is maximized when population groups are equally sized and, consequently, parties can only compete on the quality dimension.

A second, related set of papers only allows anonymous income tax and transfer schemes but introduce a non-economic dimension (Roemer, 1998; Smith and Wallerstein, 2006; Besley and Coate, 1997). If politicians can commit to policies, then the voter coalition targeted by parties will vary with the salience of the non-economic dimension (Roemer, 1998). In the absence of policy commitment, as in citizen candidate models (Besley and Coate, 1997; Osborne and Slivinski, 1996) and partisan politics models (Alesina and Rosenthal, 1989), the relevant force is that different politicians embody alternative policies while candidate supply is fixed.<sup>2</sup> Hence, making the non-economic dimension more salient may change the equilibrium choice of the winning candidate (or party) and thereby change

---

<sup>2</sup>Caselli and Morelli (2004) endogenize the supply of politicians and show that one reason for limited supply is that the most competent citizens may often not enter politics.

policy-making. We build on insights offered by this class of models.

Our paper also compliments a small but growing literature that directly examines the economic effects of reduced political competition (potentially due to relatively greater salience of other, non-economic dimensions). Besley et al. (2008) show that a reduction in partisan bias can enhance political competition on the economic policy dimension, and increase growth. They provide supportive evidence from the United States. Svaleryd and Vlachos (2008) explore similar ideas in the context of Swedish municipalities and find that greater political competition reduces rents.

Our contribution is to emphasize, and test, an important comparative static prediction: The adverse impact of increased salience of ethnicity on politician quality is more pronounced as the numerical dominance of a group increases and is absent when groups are equally sized.

The prediction that greater numerical dominance can worsen outcomes in an ethnicized population also resonates with the political science literature on ethnic politics. Horowitz (1985) likens elections in ethnicized populations to a racial census, and using African and Indian case studies respectively, Bates (1983) and Chandra (2004), argue that voter ethnicization may complement patronage politics and lower the quality of governance.

In contrast, empirical papers in the economics literature have largely focused on the negative reduced form relationship between the extent of ethnic diversity and the quality of governance.<sup>3</sup> A recent exception is Munshi and Rosenzweig (2008), who demonstrate that, in electorates that are small enough to be dominated by a single ethnic group, the group's internal disciplining mechanisms can act as an effective check on politician misbehavior.<sup>4</sup> They exploit the specific structure of reservation of seats for backward

---

<sup>3</sup>Cross-sectional evidence from regions within a country or across countries suggests that greater ethnic diversity, is correlated with reduced GDP (Alesina et al., 1999), lower GDP growth (Easterly and Levine, 1997), worse private provision of public goods (Miguel and Gugerty, 2004; Khwaja, 2008) and increased corruption (Mauro, 1995). However these results do not usually survive the inclusion of a region/country fixed effect (Banerjee and Somanathan, 2007). Most of these papers use standard measures of diversity such as ethnolinguistic fragmentation and the Esteban and Ray (1994) polarization measure, which are typically maximized with equal-sized groups (this is easiest seen in the two group case).

<sup>4</sup>This argument is also made in Miguel and Gugerty (2004).

groups in Indian village councils to test the predictions of the model.

Here, we emphasize political competition at a macro level where the electorate is too large for any significant sized group to directly discipline politicians. In this environment, moving away from equal-sized population groups towards one-group dominance worsens outcomes (measured by politician quality) because competition becomes less intense.<sup>5</sup> This is corroborated by our data. That said, we posit that this negative trend with respect to one group dominance may be reversed when a group is sufficiently large so that it is electorally viable for multiple parties to represent the same ethnic group. In that setting, the main axis of competition would be within the dominant ethnic group (and therefore on quality), leading to improved policy outcomes.

### 3 A Model of Political Competition

We develop a model of political competition where voters have two-dimensional preferences. Voters prefer honest politicians, but differ in the extent to which they share the majority group's preferences. Given our empirical application, this is naturally thought of as an ethnically targeted policy. Below, we describe the political environment and then derive predictions on how increased voter ethnicization can affect politician quality.

#### 3.1 The Political Environment

We consider a single jurisdiction with a large population of voters. Each voter is characterized by a scalar  $\tilde{\lambda}$ , which represents her degree of identification with the majority ethnic group in the jurisdiction. (The way we define preferences below, higher values of  $\tilde{\lambda}$  represent weaker association with the majority group.) We are interested in the effects of changes in the distribution of  $\tilde{\lambda}$ : We capture this by assuming that  $\tilde{\lambda} = \mu\lambda$ , where  $\lambda$  is a random variable distributed with a uniform density function on  $[\lambda_0, \lambda_1]$ ,  $\lambda_0 < 0 < \lambda_1$ ; and  $\mu > 0$  is a parameter that adjusts the spread of the distribution of  $\tilde{\lambda}$ . A higher  $\mu$

---

<sup>5</sup>This is also Munshi and Rosenzweig (2008)'s interpretation for why their results differ from ours.

represents a more spread out distribution. We will assume that  $\frac{\lambda_0 + \lambda_1}{2} \leq 0$ , so those with  $\tilde{\lambda} \leq 0$  are the majority in the population.

Two parties,  $L$  and  $R$ , representing the two possible policy positions,  $L$  and  $R$ , compete. In each election, party  $i$  gets a random draw of a particular candidate. The candidate's policy position is defined by his party's position. However, his quality level,  $Q_i$ , is randomly drawn from a uniform distribution on  $[0, Q]$ . The draws for two parties are independent of each other and of the outcomes of all other draws across elections.

Voters evaluate candidates on the basis of their position and quality: A voter  $\tilde{\lambda}$  assigns a value  $Q_L - \tilde{\lambda}P$  to a politician of quality  $Q_L$  from party  $L$ , and a value  $Q_R + \tilde{\lambda}P$  to a politician of quality  $Q_R$  from party  $R$ , where  $P$  is a fixed positive number. Therefore, all else equal, voters with  $\tilde{\lambda} \leq 0$  prefer party  $L$ , while those with  $\tilde{\lambda} > 0$  prefer party  $R$ . In this sense, party  $L$  is the pro-majority, and party  $R$  the pro-minority party. Moreover, parties are symmetric in the sense that party  $L$  is as pro-majority (i.e., favorable to those with  $\tilde{\lambda} < 0$ ) as party  $R$  is pro-minority (i.e., favorable to those with  $\tilde{\lambda} > 0$ ).

The median value of the  $\lambda$  distribution,  $\beta = \frac{\lambda_0 + \lambda_1}{2}$  is a natural measure of the *bias* of the underlying distribution: For any given  $\mu$ , a distribution with a more negative  $\beta$  will have greater fraction of pro-majority voters. For a fixed  $\beta$ , a higher  $\mu$  corresponds to more polarized preferences in the sense of there being fewer voters in the middle of the  $\lambda$  distribution. Since a polarized distribution of preferences has a specific formal meaning in the literature (Esteban and Ray, 1994), we use the term *ethnicization* to refer to increases in  $\mu$  (this is also consonant with our interpretation of  $\tilde{\lambda}$  as the dimension of ethnic preferences in our empirical analysis).

In this setting, sincere voting is optimal. The candidate with the most votes wins unless both candidates tie, in which case the winner is chosen randomly.

### 3.2 Analysis

Define  $Q_R(Q_L)$  to be the minimum value of  $Q_R$  that just gives party  $R$  a chance of winning for any given value of  $Q_L$  if such value exists in  $[0, Q]$  and  $Q$ , otherwise. Then



$Q_R(Q_L) = \min\{Q, Q_L - 2\tilde{\lambda}_m P\}$ , where  $\tilde{\lambda}_m = \beta\mu$  represents the median value of  $\tilde{\lambda}$ .

Assume  $2\tilde{\lambda}_m P + Q > 0$ , so that if party  $R$  has the best possible candidate ( $Q_R = Q$ ) and party  $L$  has the worst ( $Q_L = 0$ ), the party  $R$  candidate will win. The average quality of a party  $L$  winner, which we denote as  $Q_L^{\bar{W}}$ , is given by

$$Q_L^{\bar{W}} = \frac{\int_0^Q Q_L \int_0^{Q_R(Q_L)} dQ_R dQ_L}{\int_0^Q \int_0^{Q_R(Q_L)} dQ_R dQ_L}.$$

**Lemma 1: The average quality of a party  $L$  winner is increasing in  $\tilde{\lambda}_m$ , i.e.,  $\frac{dQ_L^{\bar{W}}}{d\tilde{\lambda}_m} > 0$ . The converse is true for a party  $R$  winner.**

The proof, which is in the Appendix, follows from the fact that whenever  $\tilde{\lambda}_m < 0$ , then  $Q_R(Q_L) < Q_L$ . That is, a party  $L$  candidate may win even when he has lower quality. Increasing  $\tilde{\lambda}_m$  (i.e., making it closer to 0) shrinks the range over which party  $L$  wins with a worse candidate. Identical reasoning applies to minority-party winners, but in reverse. Ethnicization and a greater pro-majority bias makes it harder for minority party candidates to win. Hence, the quality of winning minority party candidates will be higher.

As long as  $\mu \neq 0$ , then  $\frac{dQ_L^{\bar{W}}}{d\tilde{\lambda}_m} > 0$ . Since  $\frac{dQ_L^{\bar{W}}}{d\tilde{\lambda}_m} > 0$ , and  $\tilde{\lambda}_m = \mu\beta$ , then  $\frac{dQ_L^{\bar{W}}}{d\mu} = \beta \frac{dQ_L^{\bar{W}}}{d\tilde{\lambda}_m} > 0$  unless  $\beta = 0$ . Combining these gives our first result:

**Result 1: Assuming there is bias ( $\beta < 0$ ) and voter ethnicization ( $\mu > 0$ ) :**

1. **Majority-party winners will, on average, be lower quality than minority-party winners.**
2. **The quality gap will be larger in more majority-biased jurisdictions (keeping the degree of ethnicization fixed) and in jurisdictions with more ethnicized voters (keeping the degree of bias fixed)**

The observation that  $\frac{dQ_L^{\bar{W}}}{d\mu} \approx 0$  when  $\beta \approx 0$  provides an additional prediction: Absent bias, the quality of the majority party winner is invariant to an increase in ethnicization.

More generally, if bias is not too extreme, the effect of an increase in voter ethnicization (i.e., an increase in  $\mu$ ) on the median value,  $\tilde{\lambda}_m$ , increases with the extent of bias (i.e.,  $\beta$ ).

However, once bias is strong enough that party L is guaranteed to win, then an increase in ethnicization no longer affects who wins. More formally,  $\frac{dQ_L^W}{d\tilde{\lambda}_m} \rightarrow 0$  as  $Q + 2\tilde{\lambda}_m P \rightarrow 0$ . Hence, for sufficiently negative values of  $\beta$ , we expect  $\frac{dQ_L^W}{d\mu} \rightarrow 0$ . When we directly compute  $\frac{d^2 Q_L^W}{d\beta d\mu}$ , we see that the expression is positive when  $\beta = 0$  and negative when  $Q + 2\beta\mu P = 0$ .<sup>6</sup> Therefore,  $\frac{d^2 Q_L^W}{d\beta d\mu}$  becomes negative for values of  $\beta$  somewhere between these limits. Direct computation shows that there is only one switch, which occurs when  $Q \approx -3.5\tilde{\lambda}_m P$ . For values of  $\beta$  such that  $Q + 3.5\tilde{\lambda}_m P \geq 0$  (i.e., quality is important enough to ensure that the highest quality party R candidate can defeat a party L candidate who is at the 43<sup>rd</sup> percentile in the distribution of Q),  $\frac{d^2 Q_L^W}{d\beta d\mu}$  is positive. The reverse argument holds for the minority party winners. This gives us:

**Result 2: An increase in voter ethnicization:**

1. **Does not affect the quality of either the majority or minority party winner if the jurisdiction has no bias.**
2. **Increases the quality gap between the majority and minority party winner by more in more biased jurisdictions, as long as  $Q + 3.5\tilde{\lambda}_m P \geq 0$ .**

We now consider the winner-loser gap in quality, denoted by *WGap*, and given as

$$\begin{aligned}
 WGap &= \frac{1}{Q^2} \left[ \int_0^{Q+2\tilde{\lambda}_m P} \left[ \int_{Q_L-2\tilde{\lambda}_m P}^Q (Q_R - Q_L) dQ_R + \int_0^{Q_L-2\tilde{\lambda}_m P} (Q_L - Q_R) dQ_R \right] dQ_L \right] \\
 &\quad + \frac{1}{Q^2} \left[ \int_{Q+2\tilde{\lambda}_m P}^Q \int_0^Q (Q_L - Q_R) dQ_R \right] dQ_L
 \end{aligned}$$

---

<sup>6</sup>The expression is

$$\begin{aligned}
 &\frac{2}{3} \frac{P}{(4P^2\beta^2\mu^2 + 4PQ\beta\mu - Q^2)^3} \times \\
 &(64P^6\beta^6\mu^6 + 192P^5Q\beta^5\mu^5 + 176P^4Q^2\beta^4\mu^4 - 96P^3Q^3\beta^3\mu^3 - 12P^2Q^4\beta^2\mu^2 + 4PQ^5\beta\mu - Q^6)
 \end{aligned}$$

When  $\beta = 0$ , this expression is  $\frac{2}{3}P > 0$ . When  $Q + 2\beta\mu P = 0$ , the expression is  $\frac{3}{16}P < 0$ .

**Lemma 2: The winner-loser gap is increasing in  $\tilde{\lambda}_m$ , as long as  $\tilde{\lambda}_m < 0$ .**

$$\frac{dWLgap}{d\tilde{\lambda}_m} = -\frac{1}{Q^2} [8P^2\tilde{\lambda}_m] (Q + 2\tilde{\lambda}_mP).$$

The proof is in the Appendix. The result follows from the same forces that implied Lemma 1: The maximum quality deficit that the majority party can overcome, by virtue of being aligned with the majority, increases when  $\tilde{\lambda}_m$  becomes more negative since the weight put on party identity relative to quality increases on average. In other words, better candidates will lose, and worse candidates will win.

Since  $\tilde{\lambda}_m = \beta\mu$ ,  $\tilde{\lambda}_m < 0 \rightarrow \beta \neq 0$ , and  $\mu \neq 0$ , it follows from Lemma 2 that

$$\begin{aligned} \frac{dWLgap}{d\mu} &= -\frac{1}{Q^2} [8P^2\beta^2\mu] (Q + 2\beta\mu P) < 0, \\ &\text{and} \\ \frac{dWLgap}{d\beta} &= -\frac{1}{Q^2} [8P^2\beta\mu^2] (Q + 2\beta\mu P) > 0. \end{aligned}$$

This gives us:

**Result 3: Assuming the jurisdiction exhibits bias ( $\beta < 0$ ) and voter ethnicization ( $\mu < 0$ ), the average winner-loser gap increases with bias (lower  $\beta$ ) and voter ethnicization (higher  $\mu$ ).**

A further prediction follows from the observation that  $\frac{dWLgap}{d\mu} \approx 0$  when  $\beta \approx 0$ . In other words, in the absence of bias, an increase in the degree of ethnicization does not affect the winner-loser gap. With little bias, the two parties are (approximately) symmetrically situated. As a result, irrespective of the degree of ethnicization, the better candidate almost always wins. It is bias that gives one party an initial advantage, which is then enhanced by ethnicization.

Once again, the effect of ethnicization on the winner loser gap is not monotonically increasing with bias. If bias is so large that party  $L$  always wins, then ethnicization does not alter the winner-loser gap. If we instead start with a more modest bias, then an

increase in ethnicization favors party  $L$  and reduces winner quality.<sup>7</sup> This gives us:

**Result 4: An increase in voter ethnicization:**

1. **Does not affect the winner loser gap in a jurisdiction with no bias.**
2. **Reduces the winner-loser gap by more in more biased jurisdictions as long as  $Q + 3P\beta\mu > 0$ .**

### 3.3 Extensions

*Multiple Parties and Party Entry* We have considered the two party case, with fixed party positions. Our average winner quality results are robust to allowing more parties, *per se* (i.e.,  $m$  party  $L$ s and  $n$  party  $R$ s, with the quality of the best party  $L$  and  $R$  candidate (across all  $L$  and  $R$  parties respectively) uniform on  $[0, Q]$  and  $[0, Q']$  respectively). The winner-loser gap results would need to be slightly modified (to allow for the possibility that both the winner and loser belong to, say, party  $L$ s), but nothing will really change.

This extension assumes that bias and ethnicization do not affect the set of parties. To see how things play out when we drop this assumption, consider the case where, in addition to the two existing parties (one on the left, one on the right), a third party contemplates entry. Entry involves some sunk cost (e.g., building a reputation, achieving visibility, etc.), and occurs before candidate identity is known. Only a party with a sufficiently high expected chance of winning will enter, and with sunk entry costs no party exits. In a jurisdiction with  $\tilde{\lambda}_m$  very close to zero and just the original two parties, the *ex ante* probability that either wins is approximately 50%. A third party will only enter in position  $L$ , since for any realization of candidate quality, it does (weakly) better with a pro-majority candidate. Because  $\tilde{\lambda}_m \approx 0$ , the three parties will be equally likely to win. Assume that a 1/3 probability of winning does not make entry worthwhile. Now consider

---

<sup>7</sup>The exact condition can be seen by differentiating the expression for  $\frac{dWLgap}{d\mu}$  with respect to  $\beta$  to get

$$\frac{d}{d\beta} \left[ \frac{dWLgap}{d\mu} \right] = -16 \frac{P^2}{Q^2} \beta \mu (Q + 3P\beta\mu)$$

a jurisdiction with  $\tilde{\lambda}_m$  significantly below zero. The likelihood of winning for a party that enters in position  $L$  will exceed  $1/3$  as  $\tilde{\lambda}_m$  falls further below zero (in the limit, as  $\tilde{\lambda}_m$  goes to  $-Q/2P$ , party  $R$  essentially never wins, and the two  $L$  parties each win with probability  $1/2$ ). Hence, third party entry will occur in jurisdictions with extreme  $\tilde{\lambda}_m$  values. Such entry will improve party  $L$  winner quality since the new party  $L$  winner will be the better of the two party  $L$  candidates. And since the party  $R$  candidate faces higher quality opponents, the average party  $R$  winner quality will also increase. Consequently, winner quality might increase with ethnicization and be higher in more biased jurisdictions.

While this may sound initially surprising, notice that a more negative  $\tilde{\lambda}_m$  implies greater homogeneity: In the limit, when  $\tilde{\lambda}_m$  is very negative, the entire population belongs to the majority group and party  $R$  is irrelevant. Instead, multiple party  $L$ s compete and the highest quality candidate will win. This suggests that the effects we highlight may reverse once  $\tilde{\lambda}_m$  becomes sufficiently negative and party entry is allowed. In this sense, our predictions are potentially consistent with the now conventional idea that sufficiently homogenous populations have a more benign political economy. In our empirical analysis, we are unable to test this prediction as we do not have a sufficient number of very biased jurisdictions. We, therefore, leave this for future work.

*Politician Quality* We assume that candidate choice is dictated by chance. However, in reality candidate quality may be endogenous: Different individuals may join as candidates in jurisdictions where the party represents the dominant group, or parties may choose, or invest in, more or less qualified candidates. It is plausible that the increased dominance of the party makes it more attractive to low quality candidates (criminals, for example) and that they end up taking over the party machine in those jurisdictions and running for election. While outside our model, the basic logic is very similar – criminals favor the majority party because they expect its candidates to win.

A previous version of the paper considered the case where each party gets a random draw of candidates but candidates could differ on both  $P$  and  $Q$  (we assumed that one party always gets candidates with  $P < 0$ , and the other candidates with  $P > 0$ ). We

showed that a party’s vote-maximizing choice of candidate did not depend on  $\beta$  or  $\mu$ , and hence, it essentially reduced to the case analyzed here.

A more complex possibility is if parties have a fixed candidate supply and there are multiple jurisdictions. In allocating candidates across jurisdictions, the party that represents the numerically dominant group overall may adopt a very different strategy from the party that almost always represents the minority. Specifically, it may reserve its best candidates for jurisdictions where minorities are most numerous (since it expects to win the remaining jurisdictions quite easily), and force the minority party to put its best candidates where it is relatively strong. The predictions of such a model may diverge from our model. However, at least in the context we are studying, candidates are, with few exceptions, always local politicians, and switching candidates across jurisdictions is rare.

Finally, we have assumed identical quality distribution in the two parties. We obtain similar results if  $Q_L$  is uniform on  $[0, Q]$  while  $Q_R$  is uniform on  $[0, Q']$ . However, a uniform quality distribution is more than an assumption of convenience. To see this, consider Result 1. In general, for any  $Q_R$  realization a more negative  $\tilde{\lambda}_m$  lowers the quality threshold party  $L$  needs to win – a force in the direction predicted by Result 1. However, this does not ensure that the quality of the average party  $L$  winner worsens. For instance, suppose the  $Q_L$  distribution is such that the fall in threshold quality mostly affects who wins for very high realizations of  $Q_R$ . This occurs if, at all other levels of  $Q_R$ , the likelihood of a party  $L$  candidate just below the threshold is very small. Then a fall in  $\tilde{\lambda}_m$ , by increasing the probability that high quality party  $L$  candidates (who can beat high quality party  $R$  candidates) win, may increase average party  $L$  winner quality.

Our assumption of a uniform distribution rules out this possibility —the probability of someone being near the threshold is the same at all levels; more generally, we need the distribution to be not too unlike the uniform for Results 1 and 2 to hold.

Result 3 does not rely on assumptions about the quality distribution (see online appendix). With a non-uniform distribution, a decline in  $\tilde{\lambda}_m$  may introduce different party  $L$  winners (who may be better than the old winners – which is why Results 1 and 2 do not

hold), but the average quality of winners always decline relative to the average quality of losers. While, in general, the effect of greater bias (a more negative  $\beta$ ) on  $\frac{dWLgap}{d\mu}$  cannot be signed in this case, we know that absent bias ( $\beta = 0$ ) ethnicization does not affect the winner loser gap since the better candidate always wins. Therefore  $\frac{dWLgap}{d\mu} = 0$  when  $\beta = 0$  and hence  $\frac{d^2WLgap}{d\mu d\beta} = 0$ . When  $\beta < 0$ , on the other hand, we know that  $\frac{dWLgap}{d\mu} < 0$ . Therefore, it is easily shown that for  $\beta < 0$  but not too small,  $\frac{d^2WLgap}{d\mu d\beta} > 0$ .

## 4 The Context and Empirical Strategy

We describe the setting of our empirical analysis, our data-sets, and empirical strategy.

### 4.1 Ethnic Politics in Uttar Pradesh

We focus on politics in India’s most populous state, Uttar Pradesh (UP).<sup>8</sup> UP has a population of 166 million, over 80% of which is Hindu by religion. Ethnic politics in UP, as in much of India, is closely linked to the Hindu caste system. Historically, the caste system divided Hindu society into a hierarchically ordered set of endogamous groups, with groups lower in that hierarchy – which are categorized by the present political system as Scheduled Castes (SC) and Other Backward Castes (OBC) – facing significant discrimination. In recent years, the sense of hierarchy has been breaking down, with lower castes more likely to see themselves as just another ethnic group (though they often, with some reason, demand restitution against past discrimination).

At Independence, the Congress Party dominated UP politics. The leadership of this party in UP has historically been upper caste (Jaffrelot (2003)). While non-Congress parties (briefly) came to power in the late 1960s and between 1977-1980, the Congress hegemony in UP was largely unchallenged until after 1984. Until this point, the main opposition parties were also relatively upper caste dominated, and low caste legislators were mainly

---

<sup>8</sup>India is a federal country; legislature elections use plurality rule (with single-member jurisdictions).

confined to reserved jurisdictions, where only SC candidates could stand for election.<sup>9</sup> In 1984, an explicitly SC party, the Bahujan Samaj Party (BSP), was formed. The party campaign slogans make clear its ethnic nature ((The upper castes of) *Brahmins, Thakurs and Baniyas are thieves, the rest belong to the oppressed group*) and its ambition to use the numerical dominance of lower castes as a justification for its quest for power (*85% living under the rule of 15%, this will not last, this will not last, and The highest number has to be the best represented*). A second low caste party which mainly targeted OBC voters, the Samajwadi Party (SP), was formed in 1992. Since the early 1990s, one (or both) of these two parties have been a part of the elected UP state government.

Survey data for the 1996 elections shows that over 70% of voters reported voting for the parties which represented their caste group (the Congress and BJP for upper caste voters, and SP and BSP for lower caste voters). While comparable survey data is unavailable for the 1980s, a time-series analysis of electoral outcomes is consistent with greater voter ethnicization. Table 1 compares electoral outcomes in a representative sample of UP jurisdictions in 1980 and 1996. We distinguish between jurisdictions depending on whether low castes constitute a population majority. Relative to a jurisdiction where low castes are a minority, the probability that a majority low caste jurisdiction had a non-low caste party legislator fell by 38% between 1980 and 1996 (Section 4.2 describes the low caste population share measure and party coding).

This rise of voter ethnicization has been widely documented, with considerable debate over the reasons for it. Yadav (2000) argues that ethnicity was made politically salient in the 1980s by the growth of popular low caste movements spearheaded by charismatic individuals who went on to form low caste parties. Chandra (2004) suggests that voter ethnicization can be traced to affirmative action in favor of low castes and agricultural growth. These created a class of middle-class, low-caste citizens who demanded political

---

<sup>9</sup>Upper caste Hindus also dominated the main opposition party, Jan Sangh. During the 1960s, communist and socialist parties constituted the third and only major block that attempted to represent lower caste interests and to cultivate lower caste leaders. Of these, the most important party (electorally) was Bhartiya Kisan Dal (BKD), a pro-peasant party which was briefly part of a coalition government in the late 1960s and early 1970s (when the socialists and BKD merged to form the Bhartiya Lok Dal).



recognition and social change, and caste-based quotas led them to organize politically along ethnic lines. In a similar vein, Jaffrelot (2003) argues that the political use of affirmative action, especially by socialist parties, was a key cause of voter ethnicization. Affirmative action is also probably responsible for the simultaneous hardening of the political stances of upper caste Hindus along both caste and religious lines and was reflected in the growing political influence of Bharatiya Janata Party (BJP). Here, we take this rise of ethnicization as given and examine whether voter ethnicization affected politician quality in UP. In Section 4.3, we discuss how our empirical strategy addresses the concern that the causes of voter ethnicization, rather than voter ethnicization *per se*, drives the results. The basic idea is that our theory predicts very different impacts of voter ethnicization across jurisdictions that vary in their demographic composition.

Finally, while our empirical analysis focusses on a single Indian state, the extent of voter ethnicization in UP in 1996 resembles reported levels of ethnic voting in many other low-income settings. The classic treatise on ethnic politics remains Horowitz (1985), who quotes the examples of Guyana, Trinidad and Ghana where surveys during the 1960s found that parties often received 80-90 percent of their votes from one ethnic group. Posner (2007) provides more recent evidence: In the Zambian and Kenyan parliamentary elections in the early 1990s a candidate running on the ticket of the party identified with the majority linguistic and ethnic group respectively had a 50 percentage points higher probability of winning than a candidate running on the ticket of the wrong party. Using Afrobarometer data for twelve countries between 1999 and 2001 Norris and Mattes (2003) shows that, on average, the party representing the largest language group in the country enjoys a seventeen percentage point electoral advantage.

## 4.2 Data

The data used in this paper comes from multiple sources, which we describe below.

*Politician Quality* We measure a politician's quality by his record of corrupt and illegal behavior as identified in a field survey which we conducted in 2003. We collected these

data for politicians who either won, or were the runner-up, in the 1980 and 1996 election in a random sample of 102 UP jurisdictions.

Our respondent sample consists of journalists and politicians. For each district and election year, we randomly selected two journalists from the pool of prominent journalists who covered that election, and two politicians elected from non-sample jurisdictions in the district. Each respondent answered questions about a random sub-sample of three politicians in the district. Roughly 90% of the respondents lived in the district about which they were questioned during the relevant election. Respondents for both election years had known the politicians for roughly the same number of years at the time of the election they were interviewed about. In both years roughly 20% of the respondents shared the caste identity of the politician they were questioned about (the number is similar for those sharing the politician's party identity).

Our premise that politicians and journalists knew a lot about other politicians of their own era was evidenced in their ability to answer detailed questions on the politicians. This allows us to construct, and use, multiple measures of a politician's propensity to engage in illegal activities. Table 2 describes our three main measures.

Our first measure is the corruption rank of the politician. Each respondent ranked politicians on a 1-10 corruption scale, where 10 is the most corrupt. On the same scale she also ranked three hypothetical politician vignettes, termed X, Y and Z. The vignettes were clearly distinguished in their corruption performance, with X the least, and Z the most, corrupt. We combine a respondent's ranking of actual and hypothetical politicians to construct an ordinal ranking: If the respondent gave the politician a corruption rank below what she gave vignette X, then we assign the politician a corruption rank of one. We assign the politician a rank of two if the respondent gives the politician the same rank as she gives vignette X, three if she gives a ranking that lies between the rank of vignette X and Y, and so on (on the construction of such ordinal ranks see King et al., 2004). This gives us a seven-point scale which, by construction, controls for respondent specific biases in what constitutes corruption.

The second measure is an index of the economic gain enjoyed by the politician after entering politics. The index is the equally weighted average of four outcomes: whether the politician used political office for personal gain, whether he or his family saw a significant improvement in their economic position, whether they started or expanded a business, and whether they started or expanded contracting activity.<sup>10</sup> Our third, and final, measure is whether the politician has a criminal record. Table 2 shows a significant decline in politician quality across all three measures between 1980 and 1996.

A first concern with using these measures is the veracity of subjective reports on corruption (Olken, 2006). We should emphasize that, given that the measures vary systematically with jurisdiction bias and ethnicization in the manner predicted by our model, the key concern is that respondents understand and base their responses on our theory (even though it is false). Specifically, they systematically (and incorrectly) state that, relative to minority politicians, majority politicians have become more corrupt over time and that this effect is strongest in jurisdictions where the population is more biased towards the majority. While this seems unlikely, to check whether our measures correlate with objective measures of politician quality we obtained criminal records for a random sample of 75 politicians in our 1996 sample from the Local Intelligence Unit cell of the district police. The match rate across the subjective reports and objective verification was 84% for the politicians where all respondents agreed, and 75% otherwise.<sup>11</sup> We therefore always report two specifications: One, which includes all reports (the Allsample), and a second, which includes a single observation for each politician (the Agreesample). In the second specification, the variable of interest (a dummy variable) takes a positive value only if all respondents agree that the politician has engaged in the activity being asked about, and is zero otherwise.

A related concern is that the respondent sample changed over time in different ways in

---

<sup>10</sup>We get similar results if in our regressions we estimate effects for the four outcomes using Seemingly Unrelated Regressions, and then average across these; this follows Kling et al. (2007)

<sup>11</sup>We do not expect a 100% match. For one, respondents may have differing views about how to classify criminal charges that were accepted by one court but overturned by a higher court on a technicality.

different jurisdictions. Table 2, however, suggests no significant difference in the average corruption rank for the three hypothetical politicians across 1980 and 1996. This is also true when we allow the effect to vary across jurisdictions with different low caste population share (i.e. we include the interaction of our measure of bias with the year dummy). The Allsample regressions control for respondent characteristics – respondent age, college education, occupation, and finally, whether he shares the politician’s party affiliation, caste, or is a friend or relative (and cluster our standard errors by politician).

Finally, a caveat on interpretation: our data summarizes a politician’s life up to now and potentially captures, in part, the consequence of getting elected. That said, since our main regressions compare across winners, we do not expect this to be a source of bias.

*Demographic and Party data* We measure a jurisdiction’s bias by its low caste population share, as measured by the 1931 census: LOshare.<sup>12</sup> This was the last census to collect caste-wise data (Banerjee and Somanathan, 2007). To account for subsequent population growth, we scale low caste population share by the 1991 Hindu population share. An important advantage of using 1931 census data is that caste population shares are defined prior to the rise of voter ethnicization. However, one may worry that 1931 data provides a very noisy measure of current population share. Fortunately, for one group of low castes, the Scheduled Castes, the census continued to collect population data, and a comparison of 2001 and 1931 census data shows that these data are highly positively correlated. Further, our survey asked respondents to identify the politically dominant groups in the jurisdiction. The correlation between LOshare and political dominance by low castes as reported in our survey exceeds 80%.

As a proxy for the degree of voter ethnicization, we rely on the widely shared claim (also supported by our data on voting patterns) that ethnic identification in the voter population rose significantly between 1980 and 1996. While we lack a direct measure of people’s preferences, survey and electoral data (see Table 1) strongly suggest a significant

---

<sup>12</sup>Low castes include castes and tribes which are officially classified as Scheduled Castes, Other Backward Castes and Scheduled Tribes.

rise in caste-based voting over this period.

Finally, we use the nature of party campaigns, membership, and especially leadership, to code parties ethnic affiliation. By this metric, we identify the Congress and BJP parties as the main non-low caste parties in the state.<sup>13</sup>

### 4.3 Estimation

Given the absence of a jurisdiction-level measure of voter ethnicization, we use a time-effect – the change between 1980 and 1996 – to capture it. While we have discussed evidence that supports this interpretation, it is clear that our estimation strategy needs to account for the fact that other things also changed between 1980 and 1996.

To account for unobservable differences across jurisdictions our regressions include jurisdiction fixed effects. Jurisdictions which differ in ethnic composition may exhibit different time trends with respect to politician selection. Our regressions therefore include either a jurisdiction-specific time trend (when we compare winners and losers within a jurisdiction) or a time trend which varies with ethnic composition of the jurisdiction (when we compare winners across jurisdictions). Finally, since candidate supply may exhibit different trends across parties, all regressions include party-specific time trend.

Our regressions can include these trends because we predict that, holding party identity constant, trends in politician quality differ depending on whether the party represents the majority population group in the jurisdiction (Results 1 and 3). Moreover, the extent of quality divergence is increasing in the dominance of the majority population group (Results 2 and 4). For reasons we discuss below, we believe testing the latter set of results afford a more credible difference-in-difference-in-difference (DDD) empirical strategy. Here, our first source of difference is over time, since voter ethnicization increased

---

<sup>13</sup>We follow the political science literature in focussing on the ethnic (caste) affiliation of the party, rather than the candidate. “Ethnically aware voters have understood that presenting a multiethnic slate is an exigency of political life, even for an ethnic party, and have accordingly voted for the ethnic party rather than for or against the ethnic identity of the individual candidates. When voters elect minority members of their ethnic party, it is wrong to regard this as non-ethnic voting. Quite the contrary: it is party and not candidate ethnic identification that counts.” Horowitz (1985)

between 1980 and 1996. The second is cross-jurisdiction variation in demographic composition (captured by LOshare). The third varies at the jurisdiction-level and by outcome. For average winner quality, it is the winner's party identity, and for the winner-loser gap in quality, it is whether the candidate won.

Below, we describe in more detail the mapping from the testable predictions offered by our model of political competition to the corresponding empirical tests.

*Average Winner Quality* We predict that, holding LOshare constant, an increase in voter ethnicization between 1980 and 1996 reduces winner quality more in jurisdictions where s/he is from the majority party relative to jurisdictions where his/her party represents the minority. Further, in the 1996 cross-section, the quality gap between majority and minority party winners will be greater in more biased jurisdictions (Result 1).

Testing the pure time series prediction presumes similar time-trends in candidate quality across jurisdictions which differ in their propensity to elect a majority-party candidate. This presumption is incorrect if, for instance, the majority group gains political influence in periods of high growth and a more successful economy increases opportunities for corruption. Or conversely, an increase in voter awareness in some jurisdictions improves both the electoral fortunes of the minority party and the quality of the candidate pool. Similarly, testing the pure cross-sectional prediction requires us to assume that high bias jurisdictions where minority candidates win differ from low bias jurisdictions where minority candidates win *only because jurisdictions differ in bias*. This assumption fails if, for example, minority candidates are equally likely to win in any low bias jurisdiction but only enjoy electoral success in high bias jurisdictions with a strong independent media (that causes the entire candidate pool to be less corrupt).

We therefore focus on testing Result 2 which introduces a further level of difference by stating that the impact of voter ethnicization on the majority and minority winner quality gap will depend on the extent of the bias in the jurisdiction. Specifically, *voter ethnicization should not affect winner quality in jurisdictions with limited or no bias. Outside this range, the effect should be increasing in jurisdiction bias. Hence, the decline*

*in the quality of non-low caste party winners (relative to a low caste party winner) should be smaller in high LOshare jurisdictions than in low LOshare jurisdictions.*

In other words, we predict that in jurisdictions with little or no bias, voter ethnicization should not affect the gap between the quality of majority and minority party winners. Rather, the entire change in the gap should be concentrated at relatively high and low values of LOshare. Multiple axes of heterogeneity in the world (unlike in our model) imply that the exact empirical counterpart of zero (or high) bias is unclear. Given that caveat, we approximate zero bias as having 50% low caste population and normalize the LOshare variable so that it equals zero at 50% low caste population share. We estimate the following cubic polynomial:

$$Y_{irjt} = \alpha_j + \sum_{k=1,2,3} \gamma_k P_i \times LO_j^k \times POST + \sum_{k=1,2,3} \beta_k LO_j^k \times POST + \sum_{k=1,2,3} \delta_k P_i \times LO_j^k + \phi_1 P_i \times POST + \phi_2 POST + \phi_3 P_i + \mu X_r + \epsilon_{irjt},$$

where  $r$  denotes respondent,  $i$  winner,  $j$  jurisdiction, and  $t$  year.  $P_i$  is a dummy which equals one if the politician belongs to a non-low caste party,  $LO$  is the normalized low caste population share (LOshare), and  $POST=1$  if the year is 1996. The parameters of interest are  $\gamma_k$ , which are the coefficients on the triple interaction of the non-low caste party dummy, the  $POST$  dummy, and (separately) the first three powers of LOshare.

For reasons discussed above the regressions also include a time trend ( $\phi_2$ ), a party-specific time trend ( $\phi_1$ ) and a time trend which varies with the first three powers of LOshare ( $\beta_k$ ). We control for time-invariant jurisdiction characteristics with a jurisdiction fixed effect  $\alpha_j$ . The regressions which use reports of all respondents (the Allsample) include a full vector of respondent characteristics  $X_r$ .<sup>14</sup>

Given our normalization,  $\gamma_3$  is the slope coefficient when LOshare equals 50%. The prediction we then test is that the slope of  $\sum_{k=1,2,3} \gamma_k P_i \times LO_j^k \times 1996$  with respect to

---

<sup>14</sup>These respondent characteristics are: respondent age and dummies for whether the respondent has a college education, is a journalist, shares politician's party affiliation, shares politician's caste and whether is a friend or relative of the politician. We cluster standard errors by politician.

$LO$  is negative when (the normalized)  $LOshare$  is either sufficiently negative or positive, and zero in the neighborhood of zero. The latter requirement implies that  $\gamma_1 \approx 0$ . In addition, a point of inflection at  $LOshare = 0$ , suggests  $\gamma_2 \approx 0$  and  $\gamma_3 < 0$ .

*Winner-Loser Quality Gap* Result 3 states that the winner-loser gap should fall between 1980 and 1996. However, the possibility that time trends differ systematically across winners and losers (even within a jurisdiction) makes it difficult to directly test this prediction. For instance, if vote “buying” has become easier and more corrupt candidates have an advantage in buying votes, then we may observe a spurious time trend in the winner-loser gap.

Therefore, as with average winner quality, we focus on a DDD strategy which examines whether, in line with Result 4, the effect of voter ethnicization on the winner-loser quality gap is mediated by  $LOshare$ . We make the working assumption that jurisdictions with  $LOshare$  between 45-55% are low bias jurisdictions, while recognizing the *ad hoc* nature of this assumption. We predict

*The decline in the winner-loser quality gap between 1980 and 1996 should be increasing in  $LOshare$  in jurisdictions with  $LOshare > 55\%$  and decreasing in  $LOshare$  in jurisdictions with  $LOshare < 45\%$ . There will be no change in low bias jurisdictions.*

The prediction that the change in the winner-loser gap depends on the extent of this bias distinguishes it from a generic time-trend. Moreover, even if one can identify reasons (other than the one we propose) for why the trend in the winner-loser gap in quality might be correlated with  $LOshare$ , it would be hard to imagine reasons for such a particular non-monotonic shape.

To test this theory, we define a variable  $bias$  which equals  $LOshare$  if  $LOshare > 0.55$ ,  $1-LOshare$  if  $LOshare < 0.45$ , and zero, otherwise. We also define an indicator variable  $lowbias=1$  if  $LOshare$  is between 0.45 and 0.55, and zero, otherwise. We estimate

$$Y_{irjt} = \alpha_{jt} + \gamma_1 W_{ijt} \times bias \times 1996 + \gamma_2 W_{ijt} \times lowbias \times 1996 + \gamma_3 W_{ijt} \times bias \\ + \gamma_4 W_{ijt} \times lowbias + \gamma_5 W_{ijt} \times R_j \times 1996 + \gamma_6 W_{ijt} \times R_j + \gamma_7 X_r + \epsilon_{irjt}.$$



$W_{ijt}$  is a dummy which equals one if the politician won the election. By including jurisdiction\*year fixed effect (denoted as  $\alpha_{jt}$ ), we control flexibly for time varying jurisdiction-specific variables.  $\gamma_1$  and  $\gamma_2$  capture the change in the winner-loser quality gap within a biased and low bias jurisdiction between 1980 and 1996.

We allow for a differential effect in jurisdictions that are reserved for Scheduled Caste candidates ( $R_j = 1$  if the jurisdiction is reserved).<sup>15</sup> While we assume that parochial voters respond to the ethnic identification of the party, it is plausible that candidate ethnic identity also matters. In reserved jurisdictions all candidates are, by law, Scheduled Castes. In our model, this can be captured by assuming a smaller absolute value of  $P$  when candidates share the same ethnic identity (the difference remains  $2P$ ). It is easily checked that this implies a smaller change in winner-loser quality in reserved jurisdictions (less variation in  $P$  implies more competition on the quality dimension).

Finally, we also report specifications for the 1996 cross-section – an election year with significant ethnicization. Since our identification uses within-jurisdiction variation, we can control for jurisdiction fixed effects and examine whether, in line with our theory, the winner-loser quality gap in 1996 was smaller in more biased jurisdictions (we cannot undertake this exercise for average winner quality since those regressions rely on cross-jurisdiction variation in the winner’s party identity). This specification also helps allay concerns that our panel regression results are driven by the choice of base year.

## 5 Results

As discussed earlier, we measure politician quality by his propensity to engage in corrupt and illegal behavior.

---

<sup>15</sup>Between 1980 and 1996 the reservation status of jurisdictions remained fixed.

## 5.1 Average Winner Quality

Column (1) in Table 3 considers the winner’s ordinal corruption rank. The first three rows report the terms on the cubic polynomial, which captures how the change in corruption rank of a non-low caste party politician between 1980 and 1996 varied with LOshare.<sup>16</sup> As predicted,  $\gamma_1$  is indistinguishable from zero and  $\gamma_3$  is significantly negative.  $\gamma_2$  has a negative, but relatively small and insignificant, point estimate. The overall shape of the relationship is exactly as expected (see Figure 2, Panel A): Consistent with a cubic polynomial shape, the estimated relationship is flat in the middle and has steeper slopes at very low and high values of LOshare. We also report the slope at the 5th and 95th percentile values of LOshare (LOshare takes values between 0.4% and 76% in our data); in both cases, the relationship is significantly negative. Moreover the predicted change in corruption rank between 1980 and 1996 is almost identical for non-low caste and low caste winners for all LOshare values between 20% and 60%, whereas at 10% LOshare the difference in change in rank is 2 (the scale goes from 1 to 7 and the average increase for the 1980-1996 period is 0.2) and at 70% LOshare it is -2 (-4 at 75% LOshare). In other words the effects are large but concentrated high bias jurisdictions.

Columns (2) and (3) show an identical pattern for economic gain. The point estimates are similar for the All and Agreed samples, though more noisily estimated for the All sample.<sup>17</sup> Figure 2, Panel B shows that in this case, as well, the estimated relationship is steeper at high and low values of LOshare. Finally, we examine the likelihood that the winner has a criminal record. In Columns (4) and (5), we find that voter ethnicization increased this likelihood by significantly more in jurisdictions where the politician’s party ethnic identity reflected that of a larger fraction of the population and less so in other jurisdictions. Once again, a graphic illustration of the estimated relationship shows a non-linear relationship between LOshare and criminality (Figure 2, Panel C).

---

<sup>16</sup>For rank, which is an ordinal measure, there is no obvious default measure when respondents disagree. We, therefore, only report results for the Allsample.

<sup>17</sup>This pattern of results holds up for each separate measure of economic gain

## 5.2 Winner-Loser Quality Gap

Next, in Table 4, we examine the link between voter ethnicization and the winner-loser gap in quality. In Panel A, we directly examine the link between jurisdiction bias and winner-loser quality gap in the 1996 data. In Panel B, we examine whether the patterns observed in the 1996 cross-section are consistent with panel estimates where we examine the change in winner loser quality gap between 1980 and 1996.

Our first two measures of politician quality – corruption rank and economic improvement – show a worsening of winner quality relative to loser (see Columns (1) and (2), Table 4). This is true in both the 1996 cross-section (Panel A) and the panel regressions (Panel B). That is, relative to the runner-up in the jurisdiction, the winner’s quality, as measured by corruption rank or propensity to benefit economically, worsened by 1996. This decline in relative quality was, however, absent in low bias jurisdictions. It was also absent in reserved jurisdictions, where the forces of parochialism were potentially muted by the fact that candidates shared the same ethnic identity.

Interestingly, Columns (4)-(5) show no significant effect of voter ethnicization on the overall winner-loser gap in criminality (though we continue to observe a differential effect in reserved jurisdictions). A potential explanation is that while opportunities for economic corruption are mainly realized when holding political office, criminal activities are readily engaged in even when outside office (and, indeed, most criminal records are acquired before entering politics). Further, relative to losers, winners are likely to be better able to prevent criminal charges being filed against them when in office.

## 5.3 Robustness checks

While our results support our theory, it is useful to examine alternative explanations. A first possibility is that our results reflect differential respondent bias which may, in turn, reflect media bias. However, for this to generate our results, the media must have become systematically more biased against the party associated with the dominant group in each

jurisdiction, with the bias absent in jurisdictions with equal sized groups (and increasing in LOshare outside it). Moreover, throughout this period upper castes controlled the national and state media, and if they were biased, it was against low caste parties everywhere.

More generally, for each candidate, we have respondents from different professions who are unlikely to share the same biases. It is therefore plausible that respondent agreement that a particular politician was corrupt reflects the undeniable nature of his corruption rather than a shared bias against him. It is reassuring that we obtain similar results across the All and Agreed sample.<sup>18</sup>

A related, but distinct issue is the extent to which these measures correlate with actual corruption. For instance, if politicians' salaries increased over time, then honest politicians may have become wealthier. This would be a concern if such a trend is correlated with LOshare *and* party identity – for instance, if politicians from non-low caste party saw a relatively greater salary increase in jurisdictions where they form a population majority. In general, it is harder to imagine reasons for why trends in these variables will vary by party and jurisdiction demographics. Importantly we do not expect these trends to be absent in lowbias jurisdictions and increasing in LOshare everywhere else.

Another concern is that our corruption measures may reflect other, more positive aspects of candidates. For example, more visible candidates may be considered more corrupt, simply because their name is invoked more often and majority party winners are more visible. For this to constitute a problem, the gap between perception and reality must have increased over time. To check that this doesn't underlie our results, Table 5 considers alternative measures of politician quality as reported by our respondents (for brevity, we report results for the All respondent sample). We focus on the winner sample

---

<sup>18</sup>A different concern is that our survey measures lifetime corruption which reflects a politician's type and available opportunities. To check for this we obtained the criminal records from the affidavits filed by a candidate as part of the paperwork required for standing for election (filing criminal record became mandatory only in 2004 which limited us to a cross-sectional analysis). We found that a non low caste party candidate who wins from a high LOshare jurisdiction was relatively less likely to have a criminal record with the converse true for low LOshare jurisdictions. Within a jurisdiction, the winner was more likely to have a criminal record.

since many of these quality measures relate to performance when in office. We estimate Equation (1); Table 5 reports the terms for the cubic polynomial in LOshare.

Columns (1)-(3) consider measures which should be strongly correlated with visibility but do not necessarily have anything to do with corruption. These include whether the politician was associated with setting up or expanding schools, known for development activities, and whether he held a party or ministerial position. The patterns we found for corruption measures do not show up here. In Columns (4)-(7), we consider more ambiguous quality measures. Columns (4) and (5) ask whether the politician was associated with business groups or criminals. As we may expect, the fact that between 1980 and 1996 politicians' propensity to engage directly in business and criminal activities increased is also reflected in a stronger association with these groups. Columns (6)-(7) examine whether the politician used his political influence to benefit his party and own social group. We see a similar, but completely insignificant trend in winners' use of political influence for party or social gain.

In Columns (8)-(10), we examine politician ability to deliver public goods. We use administrative data on the provision of three public goods – number of kilometers of road built, number of schools constructed, and number of villages electrified. For comparability, we normalize all three variables and estimate these regressions within a SUR framework. We find no evidence that more corrupt candidates better provide for their constituents.

## 6 Discussion

Our empirical results strongly support the hypothesis that voter ethnicization creates substantial opportunities for corrupt politicians: The average change in corruption between 1980 and 1996 is dwarfed by the increases in the corruption of winners from the favored parties in high bias jurisdictions. Moreover, these effects are absent in jurisdictions with no to very low levels of bias. In other words, it is jurisdictions with a more biased caste distribution which show substantial increases in corruption. The results also demonstrate

that voters recognize corruption as something undesirable: Non-low caste candidates had to show themselves as remarkably uncorrupt in order to have a chance of winning in jurisdictions dominated by low castes, and vice versa. Equally, the data provide no support for the view that corrupt politicians are good at pork-barrel politics.

The sharp trade-off between ethnic loyalties and quality reflects the absence of enough good candidates who are credible representatives of their ethnic group. This could change over time as more good candidates invest in also being seen as a representative of a specific ethnic group and as competition among them drives out the corrupt candidates.

While our theory does not directly rely on the reason why voters favor ethnic parties, it does affect the interpretation, especially in welfare terms. At one extreme, if the support for ethnic parties comes from their ability to redistribute effectively, then their presence provides real value to some voters, and our valuation of ethnic politics depends on how we weigh the preferences of the beneficiary groups relative to the losers. On the other hand, if all voters get from an ethnic party is the assurance that they would be protected from its rapacity, which would be directed towards other ethnic groups (Myerson, 1993; Miquel, 2007), then we would expect the electoral victory of a more honest politician, who does not extract resources for his personal benefit, to improve welfare. Yet another possibility is that politicians do very little for their supporters, either because they are too busy doing things for themselves or because they cannot really target very effectively. A voter might still favor his own ethnic party for historical, social, or symbolic reasons, but we would not expect changes in the politician's identity to substantially alter redistribution among groups.<sup>19</sup>

Finally, while our empirical evidence is for a large Indian state, the phenomenon of voter ethnicization has been noted in many democracies. Our results serve as a warning against investing excessive hope in the power of democracy to discipline politicians, especially in ethnically divided societies, and in extreme cases, might argue in favor of restricting government ability to target specific ethnic groups.

---

<sup>19</sup>If this were the case, and voters were rational in holding these preferences, we expect the effects of group dominance on politician quality would be relatively small.

## References

- Alesina, A., R. Baqir, and W. Easterly (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics* 114(4), 1243–1284.
- Alesina, A. and H. Rosenthal (1989). Partisan Cycles in Congressional Elections and the Macroeconomy. *American Political Science Review* 83(2), 373–398.
- Banerjee, A. and R. Somanathan (2007). The Political Economy of Public Goods: Some Evidence from India. *Journal of Development Economics* 82(2), 287–314.
- Bates, R. (1983). *Essays on the Political Economy of Rural Africa*. Cambridge Press.
- Besley, T. and S. Coate (1997). An Economic Model of Representative Democracy. *Quarterly Journal of Economics* 112(1), 85–114.
- Besley, T., D. Sturm, and T. Persson (2008). Political Competition and Economic Performance: Theory and Evidence from the United States. mimeo, LSE.
- Caselli, F. and M. Morelli (2004). Bad Politicians. *Journal of Public Economics* 88.
- Chandra, K. (2004). *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. Cambridge: Cambridge University Press.
- Easterly, W. and R. Levine (1997). Africa’s Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112(4), 1203–1250.
- Esteban, J. and D. Ray (1994). On The Measurement of Polarization. *Econometrica* 62(4), 819–851.
- Fernandez, R. and G. Levy (2008). Diversity and Redistribution. *Journal of Public Economics* 92, 925–943.
- Horowitz, D. L. (1985). *Ethnic Groups in Conflict*. Berkeley: UC Press.

- Jaffrelot, C. (2003). *India's Silent Revolution: The Rise of the Lower Castes in North India*. London: Hurst and Company.
- Khwaja, A. I. (2008). Can Good Projects Succeed in Bad Communities? mimeo, Harvard.
- King, G., C. Murray, J. Salomon, and A. Tandon (2004). Enhancing the validity and cross-population comparability of measurement in survey research. *American Political Science Review* 98, 567–583.
- Kling, J., J. Liebman, and L. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 83–119.
- Levy, G. (2005). The Politics of Public Provision of Education. *Quarterly Journal of Economics* 120(4), 1507–1534.
- Lizzeri, A. and N. Persico (2005). A drawback of electoral competition. *Journal of the European Economic Association* 3(4), 1318–1348.
- Mauro, P. (1995). Corruption and Growth. *Quarterly Journal of Economics* 110, 681–712.
- Miguel, E. and M. K. Gugerty (2004). Ethnic Divisions, Social Sanctions, and Public goods in Kenya. *Journal of Public Economics*.
- Miquel, G. P. (2007). The Control of Politicians in Divided Societies: The Politics of Fear. *Review of Economic Studies* 74, 1259–1274.
- Munshi, K. and M. Rosenzweig (2008). The Efficacy of Parochial Politics: Caste, Commitment, and Competence in Indian Local Governments. *Mimeo*.
- Myerson, R. B. (1993). Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems. *American Political Science Review* 87, 856–869.
- Norris, P. and R. Mattes (2003). Does Ethnicity Determine Support for the Governing Party? *Afrobarometer Paper No. 26*.



- Olken, B. (2006). Corruption Perceptions vs. Corruption Reality. NBER.
- Osborne, M. and A. Slivinski (1996). A Model of Political Competition with Citizen-Candidates. *Quarterly Journal of Economics* 111(1), 65–96.
- Posner, D. N. (2007). Regime Change and Ethnic Cleavages in Africa. *Comparative Political Studies*.
- Roemer, J. (1998). Why Don't the Poor Expropriate the Rich? *Journal of Public Economics* 70, 399–424.
- Smith, D. A. and M. Wallerstein (2006). Redistribution and Affirmative Action. *Journal of Public Economics* 90, 1789–1823.
- Svaleryd, H. and J. Vlachos (2008). Political Rents in a Non-corrupt Democracy. *Journal of Public Economics*.
- Yadav, Y. (2000). *Understanding the Second Democratic Upsurge: Trends of Bahujan Participation in Electoral Politics in the 1990s*.

## 7 Appendix: Proofs

**Lemma 1**  $\frac{dQ_L^{\bar{W}}}{d\lambda_m} > 0$

$$Q_L^{\bar{W}} = \frac{\int_0^Q Q_L \int_0^{Q_R(Q_L)} dQ_R dQ_L}{\int_0^Q \int_0^{Q_R(Q_L)} dQ_R dQ_L}$$

which is the same as

$$\frac{\int_0^{Q+2\lambda_m P} Q_L \int_0^{Q_L-2\lambda_m P} dQ_R dQ_L + Q \int_{Q+2\lambda_m P}^Q Q_L dQ_L}{\int_0^{Q+2\lambda_m P} \int_0^{Q_L-2\lambda_m P} dQ_R dQ_L + Q \int_{Q+2\lambda_m P}^Q dQ_L}$$

Taking derivatives with respect to  $\lambda_m$  gives us

$$\frac{d\bar{Q}_L^W}{d\lambda_m} = P(Q + 2P\lambda_m) \times \frac{\int_0^{Q+2\lambda_m P} (2Q_L - Q - 2\lambda_m P) \int_0^{Q_L-2\lambda_m P} dQ_R dQ_L + Q \int_{Q+2\lambda_m P}^Q (2Q_L - Q - 2\lambda_m P) dQ_L}{\left[ \int_0^{Q+2\lambda_m P} \int_0^{Q_L-2\lambda_m P} dQ_R dQ_L + Q \int_{Q+2\lambda_m P}^Q dQ_L \right]^2}$$

The denominator is positive. In the numerator, the term  $Q \int_{Q+2\lambda_m P}^Q (2Q_L - Q - 2\lambda_m P) dQ_L$  is positive since it is defined on the range where  $Q_L$  (and hence  $2Q_L$ ) is always bigger than  $Q + 2\lambda_m P$ . Next take the term  $\int_0^{Q+2\lambda_m P} (2Q_L - Q - 2\lambda_m P) \int_0^{Q_L-2\lambda_m P} dQ_R dQ_L$ . Note that

$$\int_0^{Q+2\lambda_m P} 2Q_L dQ_L = (Q + 2\lambda_m P)^2 = \int_0^{Q+2\lambda_m P} (Q + 2\lambda_m P) dQ_L$$

Hence,  $\int_0^{Q+2\lambda_m P} (2Q_L - Q - 2\lambda_m P) \int_0^{Q_L-2\lambda_m P} dQ_R dQ_L$  is of the form

$$\int_0^{Q+2\lambda_m P} \left( 2Q_L - \int_0^{Q+2\lambda_m P} 2Q_L dQ_L \right) H(Q_L) dQ_L$$

$H(Q_L)$  is an increasing function and, therefore, always positive. Identical reasoning applies to party  $R$  winner, but in reverse.

**Lemma 2** The winner loser gap is increasing in  $\lambda_m$ , i.e.  $\frac{dWLgap}{d\lambda_m}$  as long as  $\lambda_m < 0..$

$$\begin{aligned} WLgap = & \frac{1}{Q^2} \left[ \int_0^{Q+2\lambda_m P} \left[ \int_{Q_L-2\lambda_m P}^Q (Q_R - Q_L) dQ_R + \int_0^{Q_L-2\lambda_m P} (Q_L - Q_R) dQ_R \right] dQ_L \right. \\ & \left. + \frac{1}{Q^2} \left[ \int_{Q+2\lambda_m P}^Q \int_0^Q (Q_L - Q_R) dQ_R \right] dQ_L \right] \end{aligned}$$

Differentiating this with respect to  $\lambda_m$  gives us

$$\begin{aligned} \frac{dWLgap}{d\lambda_m} = & \frac{1}{Q^2} \int_0^{Q+2\lambda_m P} \left[ \frac{d}{d\lambda_m} \left[ \int_{Q_L-2\lambda_m P}^Q (Q_R - Q_L) dQ_R \right] + \int_0^{Q_L-2\lambda_m P} (Q_L - Q_R) dQ_R \right] dQ_L \\ & + 2P \frac{1}{Q^2} \left[ \int_0^Q (Q + 2\lambda_m P - Q_R) dQ_R \right] - 2P \frac{1}{Q^2} \left[ \int_0^Q (Q + 2\lambda_m P - Q_R) dQ_R \right] \end{aligned}$$

which simplifies to

$$\frac{1}{Q^2} \int_0^{Q+2\lambda_m P} [-8P^2 \lambda_m] dQ_L.$$

Therefore

$$\frac{dW_{Lgap}}{d\lambda_m} = -\frac{1}{Q^2} [8P^2 \lambda_m] (Q + 2\lambda_m P).$$

This expression is positive as long as  $\lambda_m < 0$ .

# ONLINE APPENDIX: NOT FOR PUBLICATION

## Changing the Quality Distribution

For the case where  $Q_L$  has a distribution function  $G_L(Q_L)$  on  $[0, Q]$  with an associated density  $g_L(Q_L)$  and  $Q_R$  has a distribution function  $G_R(Q_R)$  and a density  $g_L(Q_L)$  (also on  $[0, Q]$ ), the expression for the winner-loser gap will be

$$\begin{aligned} \frac{dWLGap}{d\lambda_m} &= \frac{1}{Q^2} \left[ \int_0^{Q+2\lambda_m P} \left[ \int_{Q_L-2\lambda_m P}^Q (Q_R - Q_L) g_R(Q_R) dQ_R + \right. \right. \\ &\quad \left. \left. \int_0^{Q_L-2\lambda_m P} (Q_L - Q_R) g_R dQ_R \right] g_L(Q_L) dQ_L \right] + \\ &\quad \frac{1}{Q^2} \left[ \int_{Q+2\lambda_m P}^Q \int_0^Q (Q_L - Q_R) g_R(Q_R) dQ_R dQ_L \right] g_L(Q_L) dQ_L \end{aligned}$$

Differentiating this with respect to  $\lambda_m$

$$\begin{aligned} &\frac{1}{Q^2} \int_0^{Q+2\lambda_m P} \left[ \frac{d}{d\lambda_m} \left[ \int_{Q_L-2\lambda_m P}^Q (Q_R - Q_L) g_R(Q_R) dQ_R + \int_0^{Q_L-2\lambda_m P} (Q_L - Q_R) g_R(Q_R) dQ_R \right] g_L(Q_L) dQ_L \right. \\ &\quad \left. + 2P \frac{1}{Q^2} \left[ \int_0^Q (Q + 2\lambda_m P - Q_R) g_R(Q_R) dQ_R \right] g_L(Q + 2\lambda_m P) \right. \\ &\quad \left. - 2P \frac{1}{Q^2} \left[ \int_0^Q (Q + 2\lambda_m P - Q_R) g_R(Q_R) dQ_R \right] g_L(Q + 2\lambda_m P) \right] \end{aligned}$$

which works out to be

$$-\frac{8\lambda_m P^2}{Q^2} \int_0^{Q+2\lambda_m P} g_R(Q_L - 2\lambda_m P) g_L(Q_L) dQ_L$$

Therefore the

$$\frac{dWLGap}{d\mu} = -\beta \frac{8\lambda_m P^2}{Q^2} \int_0^{Q+2\lambda_m P} g_R(Q_L - 2\lambda_m P) g_L(Q_L) dQ_L \leq$$

$$\frac{dW Lgap}{d\mu} = -\mu \frac{8\lambda_m P^2}{Q^2} \int_0^{Q+2\lambda_m P} g_R(Q_L - 2\lambda_m P) g_L(Q_L) dQ_L \geq 0$$

It follows that Result 3 immediately generalizes to this setting. To look at the effect of increasing  $\beta$  on this expression, we need to differentiate once again with respect to  $\lambda_m$ . This has three distinct effects: It increases  $\frac{8\lambda_m P^2}{Q^2}$  and raises the limit  $Q + 2\lambda_m P$  which also, *ceteris paribus*, increases the expression  $\int_0^{Q+2\lambda_m P} g_R(Q_L - 2\lambda_m P) g_L(Q_L) dQ_L$  (since the integrand is non-negative). Both of these effects work in the same direction, which is to moderate the fall in the winner-loser gap (because of the minus sign)

## Data Appendix

**Country Rankings for Figure 1** The graph uses data on democratic countries, defined as countries with a "Political Rights Score" of 5 or lower in the 2007 Freedom House Country Reports. For these countries, the country reports were used to create an "Ethnic Party Ranking":

- Ranking of 1: If there was no mention of ethnic political parties or ethnic-based discrimination in the report on a country.
- Ranking of 0.8: If there was a sizable minority ethnic group with its own political party or if a sizable ethnic group was described as facing discrimination (the assumption being that the exclusion of certain minority ethnic groups from the political process is a signal that dominant ethnic groups have utilized the political party system to promote their own ethnic party interests by organizing along ethnic lines). An ethnic group was considered "sizable" if it was not explicitly described as being "very small" or its population number was not given in the thousands.
- Ranking of 0.6: If significant minority group political parties were present in a country, if ethnic discrimination affected a significant percentage of the population (the threshold was at approximately 10%), or if ethnic ties were mentioned as influencing national politics.
- Ranking of 0.4: If ethnic ties played an important role in party politics in a given country but were not the only consideration in voting.

- Ranking of 0.2: If ethnic-political ties were the dominant force in political party organization for a given country.

**Selection of Sample Jurisdictions** We started with the 1991 UP districts (the administrative unit below state) and combined districts with below five jurisdictions. This gave us a sample of 51 districts with an average of 7.5 jurisdictions per district. We randomly sampled two jurisdictions per district (a third was used for substitution).

**Respondent Selection for Survey** To identify journalists as respondents, we used newspaper circulation figures to select four state-level and two district-level newspapers in each district in the two election years. In these districts, we identified prominent journalists associated with these newspapers who are still alive and randomly selected two as respondents. To identify politician respondents, we divided still alive politicians into candidates from the electorally most successful party in that year, and others. For each year and jurisdiction, we randomly selected one politician from each of these groups as respondents. If all winners from either party grouping were dead, then we substituted the first runner up, and so on.<sup>20</sup>

**Caste data** The last detailed caste enumeration was done by the British during the 1931 census. These data are available district-wise for each province under British rule and for semi-autonomous princely states. For jurisdictions from which national legislators are elected caste figures were obtained by weighing caste figures by area. We use data on Hindu castes that form more than 1% of the population of each state or province in 1931, and define LOshare as the fraction 1931 Hindu population that was OBC or Scheduled Caste or Tribe. We use the most current state-specific government lists to identify these groups.

---

<sup>20</sup>We substituted for 38 politicians, and no journalists. Six politicians were non-traceable and we were unable to get appointments with other 32 (either they refused, were in jail or politically too important to contact).

Table 1: Jurisdiction Demographics and Non-Low Caste Party Legislators: 1980 and 1996

	Low caste population (LOshare)	
	Below 50%	Above 50%
1980	0.72 (0.09)	0.80 (0.04)
1996	0.69 (0.09)	0.39 (0.05)

Notes:

1. The sample consists of the 102 jurisdictions included in the politician survey. Each cell reports the fraction of jurisdictions in which a non-low caste party candidate was elected legislator.
2. Standard errors are reported in parentheses.

Table 2: Measures of Politician Quality

	Mean, 1980 (1)	Mean, 1996 (2)	Difference (3)
<b>I. Corruption rank:</b> This is an ordinal rank on a scale 1-7: A politician receives a rank of 1 if the respondent ranks the politician as more honest than vignette X, 2 if politician is ranked the same as vignette X, 3 if ranked between vignette X and vignette Y, and so on with politician ranked 7 if he is ranked as less honest than vignette Z.	3.33 (0.05)	3.53 (0.06)	0.20 (0.10)
<i>Vignettes used to create ordinal rank (scale 1-10, where 1 is most honest)</i>			
<b>Vignette X:</b> Used political position to benefit party, but not himself. His lifestyle reflected his honestly earned income.	2.82 (0.06)	3.00 (0.06)	0.04 (0.12)
<b>Vignette Y:</b> Used political position to benefit party, own social group and family. His lifestyle was better than he could afford on his honestly earned income.	5.92 (0.07)	5.94 (0.07)	0.02 (0.12)
<b>Vignette Z:</b> Used political position to benefit party, own social group and family. Known for taking money from business groups and associating with criminals. His lifestyle far exceeds his honestly earned income.	9.45 (0.04)	9.44 (0.04)	0.00 (0.07)
<b>II. Economic Gain Index:</b> Equally weighted index of below four measures; each is a dummy variable which equals 1 if positive response	0.33 (0.01)	0.44 (0.01)	0.11 (0.02)
<b>Economic improvement:</b> Own/family economic situation improved a lot after entering politics.	0.30 (0.01)	0.40 (0.01)	0.10 (0.03)
<b>Personal Gain:</b> Used political influence for personal benefit.	0.40 (0.02)	0.54 (0.02)	0.14 (0.04)
<b>Business:</b> New/ expansion of business activity since entering politics.	0.27 (0.02)	0.40 (0.02)	0.13 (0.04)
<b>Contracting:</b> New/ expansion of contracting activity since entering politics.	0.21 (0.02)	0.27 (0.02)	0.06 (0.03)
<b>III. Criminal record:</b> Has a criminal record.	0.08 (0.01)	0.16 (0.01)	0.09 (0.03)

Notes:

1. Standard error in parentheses. Standard errors in column (3) are corrected for clustering at the candidate level.
2. All variables are from the politician survey. We report averages for the sample of winners and losers.

Table 3: Voter Ethnicization and Politician Quality

	Corruption Rank	Economic Gain Index		Criminal Record	
	All	All	Agreed	All	Agreed
	(1)	(2)	(3)	(4)	(5)
Non-low caste party*	0.36	0.36	1.20	1.48	0.38
LOshare*POST	(3.48)	(0.85)	(1.00)	(0.86)	(0.68)
Non-low caste party*	-19.74	-7.78	-7.84	-14.00	-13.93
LOshare <sup>2</sup> *POST	(12.93)	(2.84)	(4.01)	(4.08)	(4.88)
Non-low caste party*	-91.42	-22.44	-22.46	-40.41	-29.47
LOshare <sup>3</sup> *POST	(43.78)	(9.84)	(12.30)	(12.59)	(13.14)
<i>Slope coefficient at</i>	<i>-22.50</i>	<i>-3.07</i>	<i>-2.27</i>	<i>-4.73</i>	<i>-0.69</i>
<i>12% LOshare</i>	<i>(9.81)</i>	<i>(1.82)</i>	<i>(2.41)</i>	<i>(2.10)</i>	<i>(1.98)</i>
<i>Slope coefficient at</i>	<i>-19.29</i>	<i>-5.74</i>	<i>-5.31</i>	<i>-9.53</i>	<i>-7.70</i>
<i>72% LOshare</i>	<i>(8.17)</i>	<i>(1.72)</i>	<i>(2.42)</i>	<i>(2.65)</i>	<i>(3.38)</i>
N	655	664	233	626	220

Notes:

1. The sample consists of reports on winners. The All sample includes all respondent reports and the Agreed sample a single report per politician (for the economic gain index, we use a single report for each of the four components of the index). The report equals 1 if all respondents gave a positive response, and otherwise zero. The dependent variables are as defined in Table 2.

2. The non-low caste party is a dummy variable which equals 1 if the politician belongs to Congress or BJP parties, and zero otherwise. LOshare is the fraction low caste population share in the jurisdiction, normalized to equal zero at 0.5. POST is a dummy which equals 1 if the year is 1996.

3. All regressions include as controls: (i) the interactions (separately) of POST and non-low caste party with LOshare, LOshare<sup>2</sup>, and LOshare<sup>3</sup>; (ii) Non-low caste party (iii) POST and (iv) Non-low caste party\*POST. All regressions include jurisdiction fixed effects. The All sample regressions include as respondent controls: respondent age and dummies for whether the respondent has a college education, is a journalist, shares politician's party affiliation, shares politician's caste and whether is a friend or relative of candidate. Standard errors in All sample regressions are clustered by politician. The Agreed sample regressions control for number of reports per politician.



Table 4: Voter Ethnicization and the Winner-Loser Corruption Gap

	Corruption	Average Economic		Criminal record	
	rank	Gain			
	All	All	Agreed	All	Agreed
	(1)	(2)	(3)	(4)	(5)
<b>PANEL A: 1996 Cross-section</b>					
winner*bias	0.37 (0.15)	0.14 (0.03)	0.11 (0.04)	0.05 (0.05)	0.00 (0.06)
winner*lowbias	0.26 (0.51)	0.11 (0.12)	0.12 (0.20)	0.18 (0.21)	-0.19 (0.37)
winner*reserved	-0.38 (0.26)	-0.14 (0.05)	-0.11 (0.04)	-0.03 (0.10)	-0.08 (0.16)
N	598	560	217	559	206
<b>PANEL B: 1980 &amp; 1996</b>					
winner*bias* POST	0.62 (0.21)	0.09 (0.04)	0.08 (0.05)	-0.01 (0.05)	0.02 (0.06)
winner*bias	-0.25 (0.15)	0.04 (0.03)	0.03 (0.03)	0.05 (0.03)	-0.02 (0.02)
winner*lowbias*POST	-0.71 (0.72)	0.07 (0.15)	0.16 (0.23)	0.19 (0.23)	-0.17 (0.35)
winner*lowbias	0.94 (0.53)	0.03 (0.08)	-0.05 (0.12)	0.00 (0.10)	-0.02 (0.02)
winner*reserved* POST	-0.72 (0.39)	-0.27 (0.06)	-0.22 (0.07)	-0.03 (0.11)	-0.09 (0.16)
winner*reserved	0.35 (0.28)	0.13 (0.05)	0.11 (0.06)	0.02 (0.04)	0.01 (0.01)
N	1186	1210	435	1139	412

Notes:

1. The sample includes reports on winners and losers. The All and Agreed samples are as defined in notes to Table 3. Dependent variables are as defined in Table 2 and LOshare and POST in Table 3.

2. Bias is a variable which equals LOshare if LOshare>0.55; equals 1-LOshare if LOshare<0.45 and equals zero between 0.45 and 0.55. Lowbias is a dummy=1 if LOshare is between 0.45 and 0.55. Winner is a dummy variable=1 if the politician won the election, and zero otherwise. Reserved is a dummy=1 if the jurisdiction is reserved for SC candidates and POST is a dummy=1 if the year is 1996.

3. The 1996 cross-section regressions use data for only 1996 and include jurisdiction fixed effects (Panel A). The regressions in Panel B include 1980 and 1996 data and include jurisdiction\*year fixed effects. Standard errors for regressions using the All sample are clustered by politicians and include the respondent controls listed in Notes to Table 3. The Agreed sample regressions control for number of reports per politician.

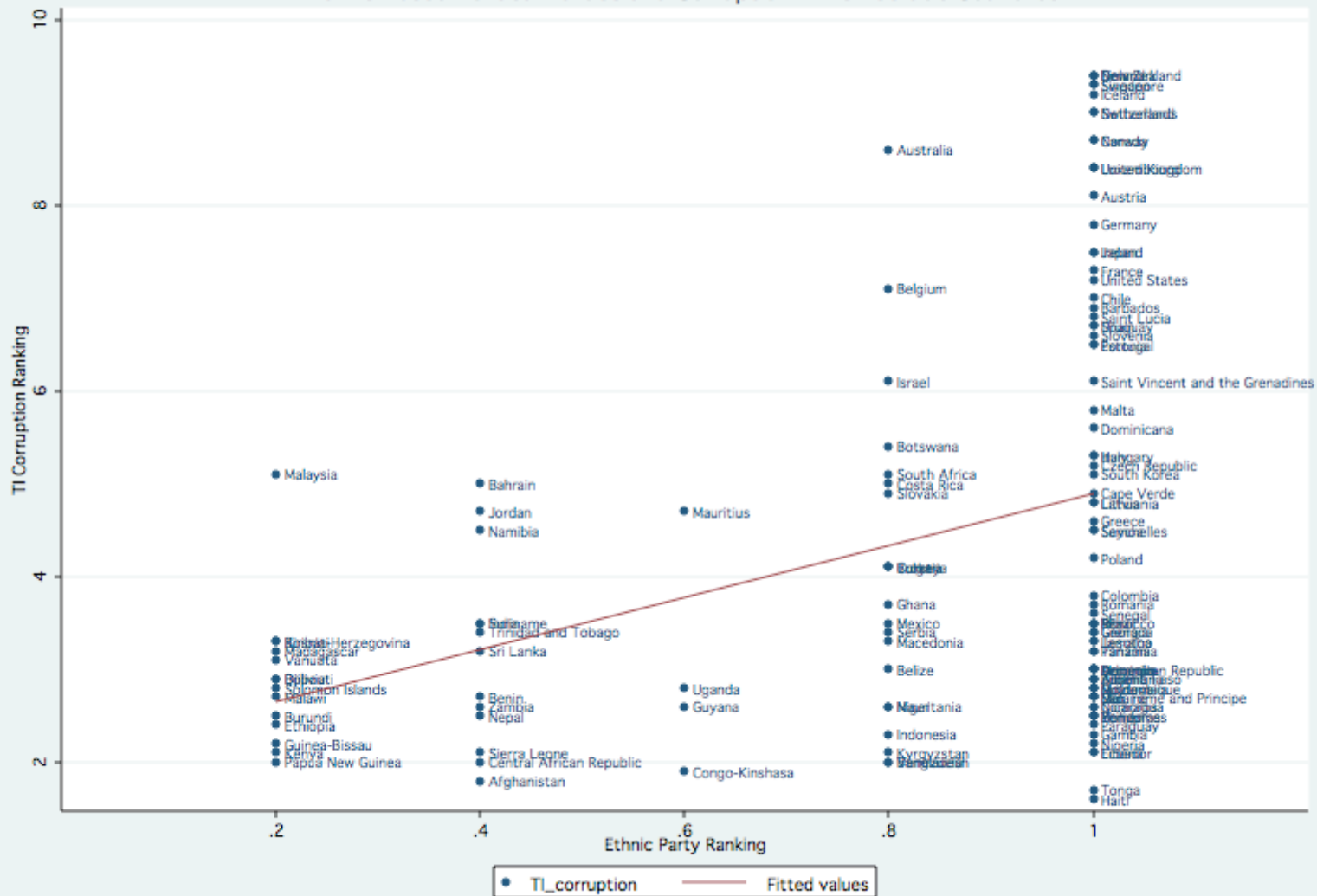
Table 5: Robustness Checks: Other Correlates of Politician Quality and Public Good Provision

	Survey Measures of Politician Performance						Infrastructure Provision			
	Built Schools/ Hospital	Known for Development	Party/ Minister Position	Associated with		Used political influence for		Roads	Schools	Electrified Villages
				Business	Criminals	Party	Social group			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
Non-low caste party*	-0.43	-1.82	1.80	1.29	2.26	0.62	-0.57	4.45	2.78	0.39
LOshare*POST	(1.28)	(1.26)	(1.60)	(1.19)	(1.09)	(1.27)	(0.87)	(3.87)	(4.29)	(1.93)
Non-low caste party*	-0.17	-2.93	-4.29	31.58	-13.91	-2.85	-6.10	-10.66	4.61	2.07
LOshare^2*POST	(5.68)	(4.89)	(6.52)	(16.34)	(5.06)	(5.32)	(4.36)	(13.26)	(16.63)	(6.32)
Non-low caste party*	2.53	8.35	-18.56	-149.64	-47.42	-11.00	-10.06	3.03	16.42	19.40
LOshare^3*POST	(17.92)	(15.62)	(21.27)	(61.98)	(15.29)	(16.85)	(13.07)	(42.76)	(62.24)	(25.70)
N	647	638	608	589	625	625	664	225	225	225

Notes:

1. The regressions include all the independent variables listed for regressions in Table 3.
2. The dependent variables in Columns (1)-(7) are dummy variables which equal 1 if the politician is reported to have undertaken that activity. The sample includes All respondent reports, and we cluster the standard errors at the politician level.
3. In columns (8)-(10), the dependent variables are the total kilometers of roads constructed in the district, the total number of primary and secondary schools in the district, and the number of villages electrified in the district, respectively. For comparability, we create and use a normalized measure for each public good and use SUR estimation to obtain covariance. Standard errors are clustered by district.

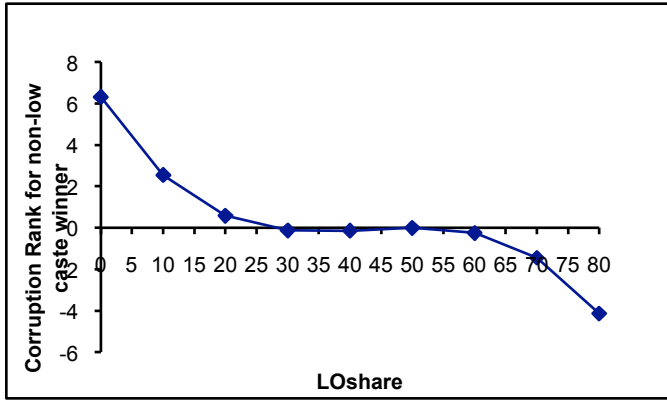
Ethnic-Based Political Parties and Corruption in Democratic Countries



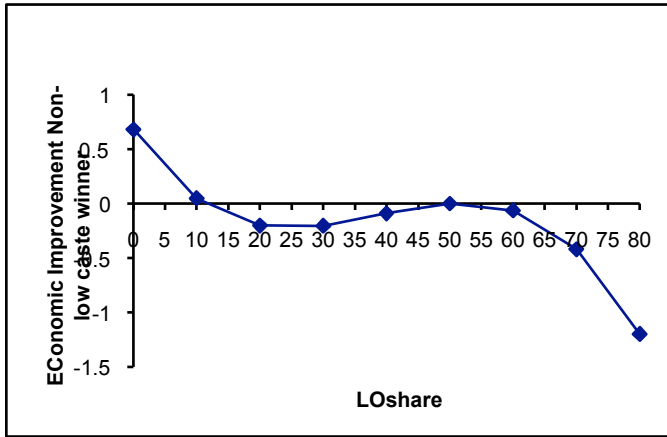
Notes: Ethnic party rankings are for the 137 democracies with 2007 Freedom House Country Reports where 1 is no mention of ethnic parties, and 0.2 is main political organization is along ethnic lines. The Corruption Ranking is the Transparency International 2007 Ranking.

Figure 2: Voter Ethnicization and Politician Quality

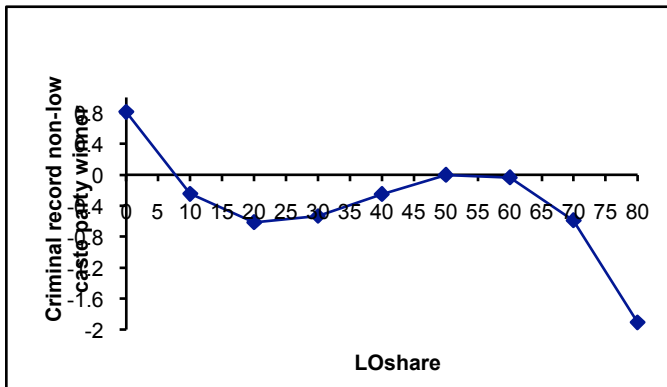
PANEL A: Change in Corruption Rank for Non-low caste winner



PANEL B: Change in Economic Gain for Non-low caste winner



PANEL C: Change in Likelihood of Criminal Record for Non-low caste winner



Notes: Each panel graphs the cubic polynomial in LOshare using the Table 3 coefficients, column 1 (for Panle A); column 2 (for Panel B) and column 3 (for Panel C). In each panel the vertical axis represents the change in corruption measure for a non- low caste winner between 1980 and 1996, relative to what you would have expected based on general trends, trends for non-low caste candidates and trends for jurisdictions with that level of LOshare.