How Do Indian Voters Respond to Candidates with Criminal Charges : Evidence from the 2009 Lok Sabha Elections

Dutta, Bhaskar and Gupta, Poonam

Warwick University, National Institute of Public Finance and Policy (NIPFP), New Delhi

26 April 2012

How Do Indian Voters Respond to Candidates with Criminal Charges : Evidence from the 2009 Lok Sabha Elections

Bhaskar Dutta and Poonam Gupta[1]

April 2012

Abstract

This paper examines the response of voters to candidates who have reported that they have criminal charges against them, within the framework of a simple analytical model which assumes that criminal charges give rise to some *stigma* amongst the electorate, and result in a negative effect on vote shares. Campaigning, the cost of which is borne from candidates' wealth, helps a candidate to increase his or her expected vote share by winning over the "marginal" voter. A criminal candidate gets an additional benefit since he can use the campaigning to convince voters of his innocence, and so reduce the negative effects of the stigma associated with criminal charges. We test the implications of the model using data for the 2009 Lok Sabha elections in India, and find support for all the implications of the model. Our empirical results show that voters *do* penalise candidates with criminal charges; however, this negative effect is reduced if there are other candidates in the constituency with criminal charges; besides, the vote shares are positively related to candidate wealth, with the marginal effect being higher for the candidates with criminal charges.

1. Introduction

It is now well-known that the nexus between Indian politicians and criminals has assumed alarming proportions. Roughly a fourth of the members of the current Lok Sabha (the lower house of the national parliament) face pending criminal charges.[2] A similar situation prevails in the various state assemblies. Many of the members of the national parliament or states assemblies have been indicted with serious charges including murder. Not surprisingly, this has attracted increasing attention in both the media as well as in academic research. It has also attracted official attention with the appointment of an independent commission to analyse the phenomenon and suggest remedial measures.[3]

It would be misleading to suggest that there is a complete absence of legal measures to prevent the influx of criminals into parliament and the state assemblies. In fact, the Representation of People's Act, 1951 specifies that candidates will be barred from contesting an election on conviction by a court of Law. The period of disqualification is for six years from the date of conviction, or from the date of release from prison, depending on the severity of the charge. Unfortunately, this law hardly has any bite because of the well-known infirmities in the Indian judicial system. In particular, governments typically drag their feet when it comes to prosecuting "local elites". Even when cases are registered, inordinate judicial delay implies that these cases drag on, seemingly indefinitely.

This is why the Election Commission had proposed in 2004 that the Representation of the People Act, 1951 should be amended to disqualify candidates accused of offences which carry sentences of five years or more *as soon as* a court deems that charges can be framed against the person. However, the Lok Sabha itself would be required to pass appropriate legislation to implement the Election Commission's suggestion. Obviously, such legislation is against the interests of a large number of politicians, and so it is not surprising that the Election Commission's proposal has not been implemented.

A landmark judgement of the Supreme Court in 2002 required every candidate contesting state and national elections to submit a legal affidavit disclosing his or her personal educational, financial, and criminal records. The court also stipulated that wide publicity should be given to

---

[2] That is, courts have decided that these charges have sufficient credibility for judicial proceedings to be initiated. However, this does not mean that these charges have culminated in convictions.
[3] See the Vohra Commission Report, 1995.

the contents of the affidavits so that the electorate can take an informed decision about who to elect to the assemblies and parliaments. Unfortunately, the Supreme Court's order does not seem to have had much impact in so far as the influx of legislators with criminal indictment is concerned.[4]

The continuing entry of large numbers of candidates with criminal records into Indian legislatures raises several questions. First, why do parties nominate such candidates? Given the huge demand for party tickets, the nomination of candidates with criminal records suggests that such candidates must possess some electoral advantage. We discuss some hypotheses which have been suggested to explain this electoral advantage. Second, what is the economic effect of electing candidates with a criminal record? Third, what is the response of voters to candidates who have reported that they have criminal charges against them?

While the first two issues have been discussed in the literature, the third issue has not been scrutinised rigorously. A somewhat cursory look at the data by simply looking at the ratio of winning candidates to number of contesting candidates amongst the criminal and non-criminal groups suggests that criminal candidates have a higher probability of winning. Perhaps, this has given rise to the feeling that criminals have an electoral advantage. The following from Aidt et al (2011) is representative of the prevailing view: "Criminals, we show, boast an extraordinary electoral advantage in India."

We examine this issue within the framework of an analytical model which assumes that criminal charges do give rise to some *stigma* amongst the electorate. This stigma has a negative effect on vote shares since voters are less likely to vote for candidates who have criminal charges levied against them. Campaigning, the cost of which is borne from candidates' wealth, helps a candidate to increase his or her expected vote share by winning over the "marginal" voter. A criminal candidate gets an additional benefit since he can use the campaigning to convince voters of his innocence, and so reduce the negative effects of the stigma associated with criminal charges. This is plausible since the candidates have not been convicted, but only charged with some criminal offence. We look at a Nash equilibrium of a game in which the only strategic variable is the amount of campaign expenditure.

---

[4] However, the judgement has been of immense help to several researchers who have exploited the information contained in the affidavits. Apart from the present paper, see, for instance, Aidt et al (2011), Chemin (2008), Paul and Vivekananda (2004), Vaishnav (2011).

We test the implications of this simple model using data for the 2009 Lok Sabha elections. We find that the data supports all the implications of the model. We briefly describe the principal results.

First, voters *do* penalise candidates with criminal charges. That is, all else being equal, the vote share of a candidate with criminal charges is lower than that of the one who does not have any such blemish. However, this negative effect is reduced if there are other candidates in the constituency with criminal charges. Notice that the negative effect of criminal charges on vote shares seems to contradict the prevalent view that candidates with criminal charges have an electoral advantage.

We do not have data on campaign expenditure of candidates. However, our model predicts that (i) the higher the wealth of a candidate, the greater will be his campaign expenditure, (ii) campaign expenditure has a positive effect on expected vote share, the marginal effect possibly differing across the two categories of candidates - those with criminal charges, and those with an unblemished record. Putting these together, the model prediction is that expected vote shares should be positively related to candidate wealth, with the marginal effect perhaps being different across the two categories of candidates. The regression results corroborate both conclusions.

Since voters penalise candidates with criminal charges, why do political parties still nominate them when so many candidates without criminal charges fight to get their party's nomination? A plausible explanation starts from the premise that candidates facing the threat of criminal convictions are more keen to contest the elections. Their enthusiasm is easily explained. Apart from the usual benefits which accrue to *all* successful candidates, candidates with criminal indictments look forward to an additional benefit. In particular, successful candidates (particularly those belonging to parties in the government) can with high probability either use coercion or influence to ensure that the local administration does not pursue the case(s) against them with any vigour.

Moreover, the data suggest that criminal candidates are wealthier than those without criminal charges.[5] Also, they are probably willing to contribute a higher fraction of their wealth to the party, or perhaps they ask for less resources from the party. This simply reflects the higher

---

[5] Vaishnav (2011) also finds that potentially criminal candidates have higher wealth.

price or value that they place on a party ticket. So, criminal candidates generate positive externalities to candidates of their own party since their additional contributions release party funds which can be used in other constituencies. This explains why parties may nominate candidates with criminal backgrounds even if they are (partially) penalised at the polls.

Several recent papers offer explanations of why parties choose candidates with a dubious background. Banerjee and Pande (2009) start with the observation that voters may have a preference for candidates belonging to their own ethnic group. This implies that a politician belonging to the ethnically dominant group in a constituency may win even if he is of lower *quality*. Banerjee and Pande (2009) assume that parties do want to select candidates of the best quality. However, the quality of candidates available to a party in any constituency is a random variable. They show that an increase in the relative size of the ethnically dominant group or an increase in voters' preferences for candidates belonging to their own group can worsen the quality of the winning candidate. Banerjee and Pande test the predictions of their model by using panel data on politician quality in 102 jurisdictions in the state of Uttar Pradesh.[6]

Of course, the Banerjee-Pande hypothesis does not explain why so many candidates with a criminal background *contest* elections. But, it does provide at least a *partial* explanation of why there is an increasing number of *successful* legislators in state assemblies as well as the Lok Sabha with criminal background.

Vaishnav (2011) studies elections to 28 state assemblies between 2003 and 2009. He finds that personal wealth of candidates is positively associated with criminal status where a candidate is defined to be a criminal if he has been charged with a "serious" crime. The basic result is subjected to a variety of robustness checks. This leads him to offer the same explanation that we have mentioned earlier - parties nominate criminal candidates simply because they contribute larger sums to the party coffers.

Aidt. el al (2011) offer a theoretical model where they assume that criminal candidates have some electoral advantage, although parties also incur some reputational cost in nominating them. They "are agnostic about the sources of this advantage", but speculate that the electoral advantage of criminals could arise because they can intimidate prospective voters of rival parties into staying away from the polls. Notice that this would imply voting turnout should be

---

[6] They measure a politician's quality by his record of illegal and corrupt behaviour as identified in a field survey.

negatively correlated with number of criminals in a constituency. We show that this is not true in the 2009 Lok Sabha elections.

So, parties face a trade-off between the reputational cost of nominating candidates with criminal charges and their electoral advantage. This trade-off implies that parties would be more willing to incur the reputational cost in constituencies which are likely to witness close contests since the electoral advantage is more attractive in these constituencies. Conversely, a party would be unlikely to field a tainted candidate in a constituency where the party is very likely to win. Similarly, candidates with criminal indictments are more likely to be fielded in constituencies where the cost is lower – for instance, in constituencies where voters are poorly informed about the characteristics of the contesting candidates.

These theoretical predictions are plausible enough given the specified model. Unfortunately, there are some questionable issues in their empirical exercise. Perhaps, the most problematic is that they use literacy as the proxy for the cost of fielding a tainted candidate. Their rationale for doing so is that illiterate voters are less likely to be aware of the criminal background of the contesting candidates. Even if this is accepted at face value, there are at least two problems with using literacy as an explanatory variable. First, the only available data on literacy is from the 2001 Census, although their electoral data are for the 2004 and 2009 Lok Sabha elections. Second, census data are available only for administrative districts which do not coincide with political constituencies. Clearly, literacy data at the constituency level for 2004 and 2009 simply do not exist!

Aidt et al measure competitiveness by the percentage difference between the vote shares of the winning candidate and her closest rival in the *same* election. This clearly raises serious endogeneity problems since the individual candidate characteristics (whether of criminal background or not) presumably has some influence on vote shares and hence on the measure of competitiveness used by the authors!

Chemin (2008) studies state elections and observes that bureaucratic corruption is lower in constituencies which elect criminal representatives. He also finds that poverty is higher in these constituencies. However, the mechanisms through which these effects operate is not spelt out in any detail.

The rest of the paper is organized as follows. In section 2 the theoretical framework is laid out and the testable hypotheses are spelt out. The econometric specification and the details on the data and the different data sources used in the paper are described in section 3. Results from the empirical exercise are discussed in section 4, and the last section concludes.

2. The Theoretical Framework

In this section, we outline a simple model of electoral competition which provides a rationalization for the regression equation(s) that we use in the paper. Before setting out the formal model, we briefly outline its basic features. Fix any constituency. Since we want to focus on how criminal charges affect the electoral fortunes of different candidates in the constituency, we do not consider how candidates choose their policy platforms. Instead, we assume that every candidate i in the constituency has a *fixed* policy or electoral platform. An alternative interpretation is that policy platforms are chosen in the first stage. Given the vector of policy platforms, candidates decide how much to spend on campaigning in the second stage. The main focus of the theoretical model is on how candidates decide on the amount of campaign expenditure, and how this affects expected vote shares.

Voters take into account the vector of policy platforms as well as candidate characteristics such as education, their past record in public service and party characteristics in deciding which candidate to support. A particular candidate characteristic that we will emphasize in the paper is *criminal record*. That is, some candidates may have a certain number of criminal charges levied against them. Such criminal charges result in some *stigma* associated to the candidates.[7]

Campaign expenditures benefit candidates in two ways. First, campaigning helps each candidate to influence voters that his or her electoral platform and individual characteristics are superior to that of the rivals. Second, candidates with criminal charges can campaign to convince voters that the charges leveled against them are baseless.[8] Voters base their voting decisions on the policy platforms, as well as the stigma attached to the different candidates.

---

[7] We use the term "stigma" to refer to the negative feeling experienced by voters about the candidate. All other things being equal, voters will prefer to vote for the candidate with a lower level of stigma.

[8] There is anecdotal evidence that in several states, candidates do *institute* false criminal charges against their opponents. Any such charge results in criminal proceedings being started. Given the inordinate delays in completing judicial proceedings in India, there is ample scope for a particular candidate to convince voters that the charges are false.

Finally, candidates choose the amount of campaign expenditure taking into account their expected vote share and its cost.

We now describe the model in greater detail.

Suppose there are n candidates in the constituency. For each candidate i, the *exogenous* characteristics are given by (p_i, c_i, w_i) where p_i represents i 's electoral platform as well as all relevant individual characteristics other than criminal record, c_i is a dummy variable which takes value 1 if i is a "criminal" and 0 otherwise, while w_i refer to the wealth of i. Each candidate i has to decide on the amount of campaign expenditure, which is financed out of the candidate's wealth. Let e_i denote the amount of expenditure of candidate i spent in order to convince the "marginal voter" to vote for him. Since all candidates participate in this activity, this resembles a contest. Let $h_i(e_i, e_{-i})$ describe the extent to which candidate i is succesful in winning over marginal voters when he spends e_i, while the campaign expenditure of others (for this purpose) is $e_{-i} = (e_1, \ldots, e_{-i}, e_{i+1}, \ldots e_n)$. Then, $h_i(e_i, e_{-i})$ is a Tullock contest function. We assume that $h_i(e)$ is a strictly increasing, strictly concave function in e_i for all e_{-i}, and strictly decreasing in e_j. So, the higher the campaign expenditure of candidate i, the larger is the expected number of votes that i can hope to win over. However, the marginal benefit of additional expenditure is decreasing in e_i. On the other hand, campaigning by other candidates eats into the vote share of i. Also, we assume that

(1) $$\lim_{e_i \to 0} \frac{\partial h_i(e_i, e_{-i})}{\partial e_i} = \infty$$

and that $\sum_{i=1}^{n} h_i(e) = 0$, where $e = (e_1, \ldots \ldots, e_n)$. An example of such a function is

$$h_i(e) = \frac{e_i}{\sum_{j=1}^{n} e_j} - 1/n$$

The criminal cases attract some *stigma* to i. Tainted candidates can campaign in order to convince voters that the charges against him are politically motivated and baseless. Let v_i denote the level of expenditure incurred for this purpose. Then, letting $S_i(c_i, v_i)$ denote the stigma attached to candidate i, we assume that S_i is decreasing and strictly concave in v_i. Also, we assume that

(2) $$\lim_{v_i \to 0} \frac{\partial S_i(1, v_i)}{\partial v_i} = -\infty$$

and

(3) $\qquad S_i(1, v_i) > 0$ for all v_i

So, all tainted candidates have an incentive to spend some strictly positive amount in reducing stigma, but they cannot wipe away the stigma completely. Of course, $S_i(0, v_i) = 0$ - no stigma is attached to candidates without any stigma. Such candidates will set $v_i = 0$.

We assume that for all i, total campaign expenditure cannot exceed the candidate wealth, so that $e_i + v_i \leq w_i$. Let p denote the vector of candidate platforms $(p_1, \dots \dots \dots, p_n)$. Similarly, e, v, c, w denote the corresponding vectors. Hence, the profile of candidate characteristics in the constituency is denoted (p, c, e, v, w).

Fix the profile of candidate characteristics (p, c, e, v, w). Candidate i 's *expected vote share EV_i* is

(4) $\quad EV_i(p, c, e, w) = K_i(p) + h_i(e) - S_i(c_i, v_i) + \sum_{j \neq i} g_{ij} (S_j(c_j, v_j)).$

Equation 4 has the following interpretation. Suppose no candidate has any criminal charges against them so that there is no stigma attached to any candidate. Also, assume first that no candidate does any campaigning. Then, $K_i(p)$ specifies i 's expected vote share corresponding to the vector of policy platforms p chosen by the competing candidates. Although we have not specified voters' behavior in detail, notice that K_i is very general. For instance, suppose P is the policy space, with voters' ideal points being distributed over P according to some distribution. Then, as in Downsian models of electoral competition, a voter will vote for the candidate whose policy platform is closest to his ideal point. Notice that we have made no assumption either about the structure of P or the distribution of voters' ideal points. We assume that

$$\sum_{i=1}^{n} K_i(p) = 1$$

As we have remarked earlier, $h_i(e)$ represents the expected increase in vote share due to campaigning.

Suppose now that candidate i has criminal charge(s) levied against him. Then, the function S_i comes into play. We assume that candidate i's stigma reduces his own expected vote share.

What is the effect on candidate i's expected vote share if some other candidate j has criminal charges instituted against him? Suppose first that candidate i is tainted. The fact that there are other candidate(s) with criminal charges lowers the stigma attached to i, and this increases i's expected vote share. Also, since the stigma attached to j makes every other candidate seem "better" in the eyes of each voter, and so increases their vote share. Assume that

(5) $$g_{ij}(S_j(c_j, v_j)) = \frac{1}{n-1} S_j(c_j, v_j)$$

Notice that for all j, $S_j(c_j, v_j) = \sum_{i \neq j} g_{ij} S_j(c_j, e_j)$. Since $\sum_{i=1}^{n} h_i(e) = 0$ and $\sum_{i=1}^{n} K_i(p) = 1$, we have

(6) $$\sum_{i=1}^{n} EV_i = 1$$

Each candidate's objective is to maximize expected vote share, net of the disutility associated with campaign expenditure. Let the disutility be represented by $d(e_i, w_i)$. We assume that

(7) $$\frac{\partial d(e_i + v_i, w_i)}{\partial e_i} > 0, \frac{\partial^2 (e_i + v_i, w_i)}{\partial e_i^2} > 0, \frac{\partial^2 (e_i + v_i, w_i)}{\partial e_i \partial w_i} < 0$$

The latter assumption means that marginal disutility is decreasing in wealth. This is a reasonable assumption and mirrors the usual assumption of decreasing marginal utility of wealth.

The only strategic variable for the candidates is the level of campaign expenditure. A Nash equilibrium is a vector (e^*, v^*), such that for each i,

(8) (e^*, v^*), maximizes $EV_i(p, c, (e_i, e_{-i}^*), (v_i, v_{-i}^*)) - d(e_i, w_i)$

Consider any tainted candidate i. His choice of (e_i^*, v_i^*) must satisfy the first order conditions.

(9) $$\frac{\partial h_i(e_i, e_{-i}^*)}{\partial e_i} = \frac{\partial d(e_i^* + v_i^*, w_i)}{\partial e_i^*}$$

(10) $$\frac{\partial S_i(1, v_i^*)}{\partial v_i} = \frac{\partial d(e_i^* + v_i^*, w_i)}{\partial v_i^*}$$

The term on the left hand side of equation 9 is the increase in expected vote share from additional campaign expenditure arising because candidate i is better able to convince voters that her policy platform is superior to that of others. So, the left hand side represents the marginal benefit arising from additional campaign expenditure. The right hand side is the marginal

disutility arising from additional campaign expenditure. So, the equation represents the familiar condition that marginal benefit should be equate to marginal disutility in equilibrium. Equation 10 is the requirement that expected marginal benefit from expenditure to reduce stigma must equal marginal disutility arising from additional campaign expenditure.

These conditions follow because our assumptions on $h_i(e)$ and $S_i(1, v_i)$ ensure an interior equilibrium; that is, $e_i^* > 0$ and $v_i^* > 0$ for each tainted candidate. Candidates who have no criminal charges against them set $v_i^* = 0$ and so the only relevant first-order condition for them is equation 9.

Given the assumptions we have made so far, a Nash equilibrium must exist. Moreover, for each (e_{-i}, v_{-i}), there is a unique pair (e_i, v_i) solving i's first order condition. Since each i's best response is unique, there can only be pure strategy Nash equilibria.

Lemma 1 : Consider any two candidates i and j such that $c_i = c_j$ and $w_i > w_j$. Then, at any Nash equilibrium (e^*, v^*), $e_i^* > e_j^*$. Moreover, if $c_i = c_j = 1$, then $v_i^* > v_j^*$

Proof : Choose i, j such that $c_i = c_j$ and $w_i > w_j$. Suppose the lemma is wrong and that there is some Nash equilibrium where $e_i^* \leq e_j^*$. From equations 9 and 10, and the fact that h_i and $S_i(1, v_i)$ are strictly concave, this implies that $v_i^* \leq v_j^*$. Then,

$$\frac{\partial h_i(e_i^*, e_{-i}^*)}{\partial e_i} \geq \frac{\partial h_j(e_j^*, e_{-j}^*)}{\partial e_j}$$

and

$$\frac{\partial d(e_i^* + v_i^*, w_i)}{\partial e_i} < \frac{\partial d(e_i^* + v_i^*, w_j)}{\partial e_j}$$

the latter following from equation 7 and $w_i > w_j$.

But, then either i or j is not satisfying the first order condition.

A similar proof establishes that $v_i^* > v_j^*$ when $c_i = c_j = 1$

This contradiction establishes the lemma.

We only have data on the wealth of candidates and not on their campaign expenditure. Fortunately, the previous lemma shows that there is a monotonically increasing relationship between wealth and campaign expenditure *within* each of the two category of candidates – that

11

is, the "tainted" candidates with criminal charges, and those who do not have any criminal charges. This monotonic relationship is used to establish the following very simple proposition. Fix any Nash equilibrium (e^*, v^*) corresponding to the exogenous vectors of characteristics(p, c, w). Let $\emptyset^* = (\emptyset_1^*, \ldots, \emptyset_n^*)$denote the expected vote shares of the candidates at this Nash equilibrium.

Proposition: The expected vote share vector $\emptyset^*$ satisfies the following

(i) For any pair of candidates i and j, if $w_i = w_j$ and $c_i = 1$, $c_j = 0$, then $\emptyset_j^* - \emptyset_i^* > K_j(p) - K_i(p)$.

(ii) Ceteris paribus, criminal charges against candidate i's rivals have a positive effect on i's vote share.

(iii) For any two candidates i and j , if $c_i = c_j$ and $w_i > w_j$, then $\emptyset_i^* - \emptyset_j^* > [K(p) - K(p)]$.

Proof : (i) Consider two candidates i and j such that $c_i = 1$and $c_j = 0$. Also, assume that $w_i = w_j$. We first show that

$$h_j(e^*) > h_i(e^*)$$

To see this, we need to show that $e_j^* > e_i^*$. Given the assumptions we have made, $v_i^* > 0$ since $c_i = 1$. Suppose $e_i^* \geq e_j^*$. Then, given $w_i = w_j$,

$$\frac{\partial d_i(e_i^* + v_i^*, w_i)}{\partial e_i^*} < \frac{\partial d_j(e_j^*, w_i)}{\partial e_j^*}$$

But,

$$\frac{\partial h_i(e^*)}{\partial e_i^*} < \frac{\partial h_j(e^*)}{\partial e_j^*}$$

The latter two inequalities show that either i or j is not satisfying equation 9.

Hence, $e_j^* > e_i^*$ and so $h_j(e^*) > h_i(e^*)$. Moreover, $S_i(1, v_i^*) > 0$, and so this too reduces i's expected vote share. From equation 4, it follows that

$$\emptyset_j^* - \emptyset_i^* > K_j(p) - K_i(p)$$

(ii) This follows straightaway from the specification of the model. If $c_j = 1$, then $S_j(1, v_j^*) > 0$ and hence $g_{ij}(S_j(1, v_j^*) > 0$.

12

(iii) Suppose $w_i > w_j$ and $c_i = c_j$, $K_i(p) = K_j(p)$. Then, we know from lemma 1 that

$$e_i^* > e_j^*$$

Moreover, if $c_i = c_j = 1$, then

$$v_i^* > v_j^*$$

It follows straightaway from equation 4 that

$$\emptyset_i^* - \emptyset_j^* > K_i(p) - K_i(p)$$

This concludes the proof of the proposition.

We discuss briefly the implications of the proposition for our regression exercise. Consider part (i) of the proposition. Essentially, this says that once we have controlled for wealth and policy platforms, then expected vote share will be lower for a tainted candidate. We will attempt to verify this in the regression exercise by checking whether the criminal dummy has a negative coefficient[9]. The implication of part (ii) of the proposition is straightforward - the coefficient on the variable representing rival candidates with criminal charges should be positive. Finally, part (iii) requires that the coefficient on the wealth variable should be positive. Notice that the proposition leaves open the possibility that wealth has a differential impact on vote shares of tainted and non-tainted candidates.

3. Econometric Specification and Data

We derive our regression equation from the model in the previous section. We are interested in explaining the vote share of each candidate i. Our model specifies that the vote share should depend negatively upon the dummy for criminal charges. Since the stigma attached to a tainted candidate is decreasing in the number of other tainted rivals, we also include the number of other candidates with charges within the constituency interacted with criminal charge dummy as an explanatory variable. The nature of the dependence of vote share on wealth is more nuanced. The wealth of candidate i himself should have a positive impact, while the wealth of other candidates should have a non-positive effect since vote shares add up to one. It therefore makes sense to use the relative wealth of candidate i as an explanatory variable. Morever, the wealth effect could differ across the two categories of candidates with criminal charges and those who do not. We accommodate this possibility by including wealth of the candidate interacted the dummy for

---

[9] In the theoretical model, the only determinants are (p, c, e, v). In the regression exercise, we will have additional controls.

criminal charges as an explanatory variable. Finally, we include other candidate characteristics such as the level of education of the candidate, dummy for the incumbent candidates seeking reelection to Lok Sabha, and a dummy for the candidates contesting as members of the state incumbent party(ies) in the regression equation.

$$Y_i = \alpha + \beta_c \; Criminal_i + \beta_w \; Relative \; Wealth_i + \beta_{cw} Criminal_i * Wealth_i + \; \gamma_s \; Number \; of$$
$$Candidates \; with \; Charges_i \; *Criminal_i + \; \beta_n \; Incumbency_i \; + \beta_{ns} \; State \; Incumbent_i \; +$$
$$\beta_e \; Education_i + \gamma \; Constituency \; Fixed \; Effects + \lambda \; Party \; Fixed \; Effects + \varepsilon_i \quad (11)$$

The vote share that we want to explain takes value between 0 and 1 and thus is bounded between these limits. Thus we transform the variable by calculating the log odds ratio for vote share of each candidate and estimate the model by ordinary least squares, with heteroskedasticity corrected standard errors. The dependent variable Y_i is thus calculated as $\log\left(\frac{Vote\;share_i}{1-vote\;share_i}\right)$. In all our regressions we include constituency fixed effects and party fixed effects to control for omitted variables, such as the varying policy platforms of the candidates belonging to different political parties. In robustness tests, we include fixed effects varying over state-party combinations instead of the other fixed effects.

We describe the data used in the paper and also discuss the broad patterns observed in the data. Our empirical results are presented in the next section.

In 2003, the Supreme Court in India decreed that all candidates contesting an election for the Lok Sabha, Rajya Sabha, or state assemblies in India had to file an affidavit with the Election Commission of India containing information on their assets (and liabilities), criminal charges and education. We derive the data on these variables directly from the affidavits of the candidates- these are available on the election commission's website as well as from a website maintained by the Association for Democratic Reforms (ADR), http://myneta.info.

The data on percent of votes obtained, age, and gender of the candidates are obtained from the election commission's website. Information on candidate incumbency has been gathered using various sources including searching through reports in the newspapers or on various internet sites. We define a party as an incumbent in a state if it was in power in the state (or was a major coalition partner), from 2008 up to the elections in 2009. The state level

incumbency information has been put together using the information contained in various articles in the Economic and Political Weekly and elsewhere (see Appendices for details). The state level data on crime has been obtained from the National Crime Bureau Reports. Appendix A1 provides the data sources from where the data on various variables have been obtained, while Appendix A2 provides the summary statistics of the variables.

India has 28 states and 7 Union Territories (UTs) in all. Among the UTs, only Delhi has its proper local administration with its own Chief Minister, while the remaining UTs are administered by the center. Therefore, we include Delhi as a "state" in our sample while excluding the remaining six UTs from the analysis. We follow Gupta and Panagariya (2011) and exclude the eight northeastern states since they have a special status with deep involvement of the center in their development process, as well as the state of Jammu and Kashmir. This leaves us with a total of 20 states including Delhi. These states account for 506 out of the total of 543 parliamentary seats across the country.

Using the data from the affidavits, we define five categories for the education status of the candidates, assigning them values from 0 to 4; 0 represents the lowest category and 4 the highest. Relative wealth is calculated as the ratio of the wealth of the candidate to the average wealth of the rest of the candidates. In one set of regressions, we exclude all Independent candidates. In this case, relative wealth is calculated as the ratio of the candidate's wealth to the average wealth of the other non-Independent candidates in the constituency.

Each candidate's affidavit has to contain information on whether the candidate faced any criminal charges, as well as the sections of the Indian Penal Code (IPC) under which the charges if any have been framed. In addition, the candidate has to declare whether he or she has ever been convicted. Thus, in principle, the data is available on the number of criminal cases that a candidate faces, the specific sections of the IPC under which the candidate faces these charges and whether the candidate has ever been convicted. The ADR further divides the charges into the charges for serious and non serious offences, by examining the sections of the IPC under which the candidates face the charges. The conviction rate of candidates facing charges is very low—out of the 1,155 candidates in the 2009 Lok Sabha elections who faced at least one criminal charge, only 15 candidates were convicted.

It is sometimes claimed that the data on criminal charges is misleading since the charges are sometimes initiated by political rivals. Moreover, some of the charges are associated with involvement in political activities. In order to clean the data of such "spurious" charges, we specify a value of one to the criminal dummy only when a candidate faces more than one charge. This adjustment takes care of some obvious cases of frivolous charges or charges arising out of political activities.[10] Henceforth, we will use the term *tainted candidate* to denote a candidate who has two or more criminal charges against them.

Consider now the patterns of criminal charges across candidates, states, and parties, and at their correlates with other candidate specific factors for the 20 states that are included in our regression analysis. Table 1 shows that it is the national and recognized state parties which field a substantially higher proportion of tainted candidates. In fact, roughly one in seven candidates fielded by state parties have at least two criminal charges levied against them. The corresponding number for national parties is over one in ten candidates. This, together with the fact that a substantially higher number of *winning* candidates come from the national and state parties, explains why the win-ratio (the ratio of the number of successful candidates to the number of contesting candidates) is substantially higher for tainted candidates. This is documented in Table 2.

Table 3 shows the distribution of constituencies by the number of candidates who faced at least two charges. On average, about 15 candidates contested the election in each constituency in the 2009 Lok Sabha elections. Despite the large number of candidates, an overwhelming number of constituencies - over 75 per cent – had no tainted candidates. In other words, there was a concentration of tainted candidates in some constituencies. In fact, states like Bihar, Jharkand, Kerala had a concentration of tainted candidates.

Table 4 shows that on average, tainted candidates were wealthier, more likely to be incumbents and obtained a much larger percent of the votes. Somewhat surprisingly, the average age and education level of tainted candidates is also higher. Indeed, the difference in averages of these variables are statistically significant at the 1 percent level.

---

[10] As robustness checks, we choose alternative specifications where (i) the criminal dummy takes value one if a candidate has three or more criminal charges; or (ii) the *number* of criminal charges instead of a criminal dummy is used as an explanatory variable.

4. Regression Results

We have pointed out in the last section that tainted candidates have a higher *win ratio*. At first sight, this might suggest that voters do not care at all about the criminal records of their elected representatives. This is misleading because of several reasons – tainted candidates are more likely to be nominated by established political parties, and they have higher wealth. These are obvious factors which influence vote shares and hence the win ratio. So, a more detailed analysis is required because any conclusions can be drawn about the response of voters to tainted candidates. This provides a strong motivation for our regression exercise.

We have two parallel sets of basic regressions. In the first, we estimate our regressions using the data for all candidates in the twenty states that we have included in our analysis. We then run the same regression on a smaller sample which includes only the candidates affiliated with some political party, thus dropping the observations for "Independent" candidates. We drop the Independent candidates since the majority of these candidates obtained only negligible vote shares.[11] Almost all the results are invariant with respect to the two samples.

Table 5 reports the basic regression results. Column I contains the results for our benchmark specification. In column II, we interact the crime level in states with the criminal dummy for the candidates in order to get some idea about how the level of criminality in the state affects voters' attitudes to tainted candidates. For instance, voters in crime-prone states may be accustomed" to a certain level of crime. If this is the case, they may be ignore the criminal records of tainted candidates. Alternatively, voters may be incensed at the level of crime, and this may make them "punish" tainted candidates. Columns III and IV of the table correspond to columns I and II respectively, but when the Independent candidates are dropped from the data.

The variables we are particularly interested in are the criminal dummy variable, relative wealth, as well as the interaction of the criminal dummy with wealth and with the number of other candidates with criminal charges in the constituency. Table 5 shows that our results are remarkably consistent with the theoretical model of Section 2. Thus, the negative coefficient on the criminal dummy shows that tainted candidates lose vote share relative to the others. Relative

---

[11] There were 3825 independent candidates with an average vote share of about 0.80 percent. Only 10 Independent candidates won in the 2009 election.

wealth has a positive effect on vote shares. The coefficient of relative wealth interacted with the criminal charge dummy is positive, implying that the loss in vote share is smaller for a wealthier candidate. Similarly, the coefficient of the interaction between the number of other tainted candidates with criminal charges and the criminal dummy is positive and significant in the regressions for all the candidates. This implies that the stigma attached to being a tainted candidate declines if there are other tainted candidates in the constituency.

Among other results, the education status of the candidates has a positive effect on vote share; and the stigma of criminal charges increases as the crime reported in the state increases. This suggests that voters in crime-prone states tend to punish tainted candidates. We also find that incumbency at the candidate level as well as at the party level in the state increases the vote share of the candidates.[12] Most of these results are robust to the exclusion of Independent candidates from the sample.

We now report on some robustness checks. Since the primary purpose of the paper is to throw light on voter response to tainted candidates, we conduct a key robustness test by constructing the dummy for criminal charges in an alternative way. This dummy takes value 1 if the candidate faces at least *three* criminal charges (instead of *two* in the earlier specification), and zero otherwise. Construction of the dummy in this way reduces the possibility of labeling a candidate as tainted if the charges against him are politically motivated or perhaps arising from violations of the law while undertaking political activities. The specifications in different columns in Table 6 are exactly the same as in the previous table. The results are also qualitatively similar to the ones obtained earlier for most of the variables. The coefficients of the criminal dummy and the interaction between wealth and criminal dummy are however larger than before, thus indicating that the loss of vote share is larger for a candidate who faces three or more charges than for the candidates with at least two charges. For such candidates, additional wealth helps in reducing the stigma by a larger amount as well.

Table 7 reports some robustness checks. Column I in Table 7 includes the interaction of state incumbent and criminal dummy, while column II includes age and gender of the candidate

---

[12] Gupta and Panagariya (2011) also come to the same conclusion. Of course, the positive effect of incumbency may be due to the fact that the coalition of parties constituting the UPA retained power in 2009. There are other instances where the ruling party or coalition has been defeated in the election- in such cases, we would not expect to observe a positive effect of incumbency on vote shares.

in the regressions. In column III, we replace the dummy for criminal charges by the *number* of criminal charges as an explanatory variable, using the latter with a logarithmic transformation. Similarly, we interact the logarithm of the number of criminal charges against each candidate with the candidate's wealth. Finally, in the last column we include the number of all candidates with criminal charges in a constituency rather than only against the top four candidates by vote share, interacted with the dummy for criminal candidates.

The results show that the coefficients of the main variables of interest—wealth or relative wealth, criminal dummy and the interaction of wealth and criminal dummy, retain their significance. The only variable which loses significance in some of the specifications is the interaction of the number of charges against other candidates with criminal dummy.

Some other robustness tests are reported in Table 8. In column I, relative wealth is calculated as the ratio of the candidate's own wealth to the sum of the wealth of candidates who received at least 3 percent of the total votes. Similarly, the number of candidates with charges also includes the data for only these candidates. In column II we estimate the regressions using the data only for the constituencies reserved for candidates from the scheduled castes and schedule tribes. In the next column we estimate the regressions only for the constituencies which are not reserved for the candidates of the schedules or scheduled tribes. In the last column we include a dummy for serious charges (which takes a value 1 if there are at least two charges against a candidates and the ADR identifies that at least one of the charges against the candidate is for a serious crime). Again all our main results hold— the criminal dummy has a negative coefficient, wealth or relative wealth has a positive coefficient, and the interaction of wealth and criminal dummy has a positive coefficient. The coefficient of other candidates with charges is mostly positive, but insignificant in some of the specifications.

We have conducted two more robustness tests, but do not report the results. In one, we drop one state at a time and estimate our benchmark specification with the rest of the data. All of our results hold with minor variations in the coefficients or the significance levels. This robustness test confirms that our results are not driven by any outlier state. Second we estimate regressions similar to those in Table 7 by eliminating the Independent candidates from the sample. The qualitative results remain unchanged.

These results seem to leave very little doubt that voters do punish tainted candidates –
this conclusion remains true irrespective of the specification chosen by us, and also remains true
when we leave Independents out of the regression exercise. However, this raises the obvious
question. Why do political parties nominate so many tainted candidates when they have so many
other aspiring candidates fighting for a party ticket? As we have mentioned earlier, Aidt et al
(2010) construct a theoretical model which assumes that tainted candidates have some *electoral
advantage* which induces political parties to nominate them despite some reputational cost. They
do not specify the nature of the electoral advantage, but mention in passing that it could be the
power of criminal candidates to intimidate voters who are likely to vote for their rivals. If this
were the case, then one would expect voter turnout to be lower the greater is the number of
tainted candidates. Table 9 negates this hypothesis – the data seem to show no negative
relationship between voter turnout and the number of tainted candidates in a constituency.

An alternative hypothesis advanced by Vaishnav (2010) is that tainted candidates are
wealthier. In fact, he find empirical support for this hypothesis in his data set which consists of
elections in various state assemblies. As Table 10 shows, this seems to be true even in our
sample. So, it seems plausible to argue that tainted candidates use their greater wealth to "buy"
their tickets. They can use their wealth to campaign more intensively, and perhaps also
contribute to party funds. Unfortunately, we have no data (other than the self-reported wealth of
the candidates) to empirically verify any other hypothesis.

5. Conclusion

Our   main empirical results suggest that voters do punish candidates who have criminal
charges against them. However, these tainted candidates are able to overcome this electoral
disadvantage because they have greater wealth, and wealth plays a significant role in increasing
vote shares.   The most plausible channel through which wealth affects vote shares is of course
through campaign expenditures, which are likely to be positively related to wealth.

There is now a fair body of evidence suggesting that voters who have information about
the corrupt ion of incumbent politicians do punish the latter. For instance, Ferraz and Finan
(2008) use detailed Brazilian electoral and audit data to show that new information about

political corruption reduces the probability of reelection for corrupt incumbents. Bobonis et al (2011) find that publicly available pre-election municipal audits significantly reduce the level of corruption in Puerto Rican municipalities. Closer home, Banerjee et al (2011) conclude, on the basis of a field experiment conducted before the Delhi state legislative elections, that voters who had access to information about incumbent performance punished worse performing incumbents and those facing better quailed challengers – these incumbents then received significantly fewer votes.

Our empirical results, along with this body of evidence suggests that it is important for voters to be better informed about candidate characteristics. The mere requirement that candidates file affidavits with the Election Commission about their characteristics is of limited use if voters do not have access to this information. Perhaps, the Election Commission needs to play a more active role in disseminating this information. The Commission must also think seriously about enhancing the existing ceilings on campaign expenditure since practically no candidate or party adheres to the current limits on expenditure. However, the Commission must ensure that all candidates adhere to the enhanced (but realistic) ceiling. This will then at least reduce the "wealth advantage" enjoyed by tainted politicians.

References

Aidt, Toke, Miriam Golden and Devesh Tiwari, 2011, "Incumbents and Criminals in the Indian National Legislature." Typescript. Department of Political Science, University of California-Los Angeles.

Banerjee, Abhijit V. and Rohini Pande, 2009, "Parochial Politics: Ethnic Preferences and Political Corruption." Unpublished paper.

Banerjee, Abhijit V. and Rohini Pande, Selvan Kumar and Felix Su, 2011, Do informed voters make better choices? Experimental Evidence from Urban India.

Bobonis Gustavo J. and Luis R. Cámara Fuertes, 2009, Does Exposing Corrupt Politicians Reduce Corruption?.

Chemin, Matthieu, 2008, "Do Criminal Politicians Reduce Corruption? Evidence from India." Working Paper 08-25, Department of Economics, University of Quebec at Montreal.

Ferraz, C. and F. Finan (2008). Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. Quarterly Journal of Economics 123 (2), 703–745.

Gupta, Poonam and Arvind Panagariya, 2012, Growth and Election Outcomes in a Developing Country.

Paul, Samuel and M. Vivekananda, 2004, "Holding a Mirror to the New Lok Sabha." *Economic and Political Weekly* (November 6).

Vaishnav, Milan, 2010, "The Market for Criminality: Money, Muscle and Elections in India." Unpublished paper, Columbia University.

Appendix A1: Description and Data Sources of Variables

Variable	Source	Description
Percent of Votes Obtained	Election Commission	
Age	Election Commission	In years
Gender dummy	Election Commission	Dummy takes a value 1 if the candidate is a female
Wealth	Election Commission	
Education Index	Election Commission	The index takes a value 0 if the candidates has no formal education, 1 if he has formal education till grade 5; 2 if the education is up to high school; 3 if education is up to undergraduate; and 4 for education level higher than undergraduate, including a technical or professional degree.
Criminal Dummy	Election Commission and Association for Democratic Reforms	The dummy takes a value 1 if the candidate has 2 or more criminal cases against him, and zero otherwise.
Incumbent MP	Various sources on the web	The dummy takes a value 1 if the candidate was a member of the previous Lok Sabha, and zero otherwise.
State Incumbent	Various sources on the web and different issues of Economic and Political Weekly	The dummy takes a value 1 if the candidate belongs to a party which was in power in state government in 2008-09 before the Lok Sabha Elections. The state incumbent parties are: Andhra Pradesh, Indian National Congress (INC), TRS; Bihar: JDU, Bhartiya Janata Party (BJP); Chhattisgarh: BJP; Delhi: INC; Goa: INC, NCP; Gujarat: BJP; Himachal Pradesh: BJP; Haryana: INC; Kerala: CPI (Marxist), CPI; Maharashtra: INC, NCP; Madhya Pradesh: BJP; Orissa: Biju Janata Dal; Punjab: Siromani Akali Dal, BJP; Rajasthan: BJP; Tamil Nadu: Dravida Munnetra Kazhagam, INC; Uttarakhand: BJP; Uttar Pradesh: Bajuhan Samaj Party; West Bengal: CPI (Marxist), RSP; Karnataka: BJP; Jharkhand: JMM, BJP

Appendix A2: Summary Statistics of Variables

Variable	Observations	Average	Minimum	Maximum
Percent of Votes Obtained	7695	6.58	0.02	78.8
Age	7694	45.76	25	88
Gender Dummy	7695	0.07	0	1
Wealth (1000s) log	7196	13.81	0.69	23.71
Education Index	7017	2.58	0	4
Criminal dummy	7676	0.07	0	1
Incumbent MP	7695	0.05	0	1
State incumbency	7695	0.07	0	1
Cognizable Crime in state (per '000 of population)	7695	1.73	0.05	11.62

Table 1: Candidates with Criminal Cases across Party Types

Party Type	Number of Candidates	Number of Candidates with least 2 Criminal Cases	% of Candidates With at least 2 Criminal Cases
	I	II	III: (II/1)*100
National Parties	1353	176	11.5
State Parties	585	108	15.6
Unrecognized Parties	1790	110	6.2
Independent Candidates	3659	124	3.4

Source: Authors' own calculations using the data mentioned in Appendix A1; data refer to the observations on twenty states included in the regressions.

Table 2: Distribution of Contesting and Winning Candidates by the number of Criminal Cases

Number of Criminal Cases	Number of Candidates	Number of Winning Candidates
I	II	III
0	6,551	349
1	607	73
2-4	382	57
5-9	92	16
>10	44	10
Total	7676	506

Source: Authors' own calculations using the data mentioned in Appendix A1; data refer to the observations on twenty states included in the regressions.

Table 3: Distribution of Candidates with Charges across Constituencies

Number of candidates with at least two charges	Number of constituencies
0	206
1	169
2	83
3	25
4	14
5	9

Source: Authors' own calculations using the data mentioned in Appendix A1, data refers to the observations on twenty states included in the regressions.

Table 4: A Comparison of Variables for Candidates with and without Criminal Charges (at least two Criminal Charges)

Criminal Dummy	% votes	Age	Log Assets (in 1000s)	Education Index	Incumbent (percent)
0	5.9	45.7	13.7	2.57	4
1	15.4***	47.2***	15.1***	2.71***	10***
Total	6.59	45.8	13.81	2.58	5

Source: Authors' own calculations using the data mentioned in Appendix A1; data refer to the observations on twenty states included in the regressions. *** indicates that the values are significantly different from those for candidates with one or no charges at 1 percent level of significance.

Table 5: Explaining the Vote Share of Candidates—Benchmark Specification

	I	II		III	IV
	Sample: All Candidates			Sample: No Independent Candidates	
Education	0.10***	0.101***		0.14***	0.142***
	[7.35]	[7.39]		[5.92]	[5.99]
Wealth log*criminal dummy	0.097***	0.095***		0.111***	0.109***
	[4.00]	[3.91]		[3.53]	[3.44]
Incumbent MP	0.842***	0.839***		0.782***	0.779***
	[9.62]	[9.58]		[8.82]	[8.78]
Criminal Dummy	-1.27***	-1.12***		-1.415***	-1.19**
	[3.59]	[3.01]		[2.99]	[2.43]
Candidates with charges (among top 4)*criminal dummy	0.24***	0.23***		0.185**	0.165**
	[4.38]	[4.14]		[2.41]	[2.13]
Relative Wealth	0.003**	0.003**			
	[2.29]	[2.28]			
State Incumbent	1.78***	1.78***		1.755***	1.75***
	[24.35]	[24.30]		[24.36]	[24.31]
Cognizable crime in state*Criminal dummy		-0.072*			-0.104*
		[1.94]			[1.91]
Relative Wealth (no independent candidates)				0.006***	0.001***
				[4.17]	[4.12]
Constituency Fixed Effects	Yes	Yes		Yes	Yes
Party Fixed Effects	Yes	Yes		Yes	Yes
Observations	6,732	6,732		3,621	3,621
Adjusted R-squared	0.781	0.781		0.772	0.772

*,**, *** indicate that the coefficient is significantly different from zero at 10, 5 and 1 percent levels of significance respectively. Robust t statistics are reported in parentheses. The regression equation is in equation 8 and the variables are defined in Appendix A1 and in the text.

Table 6: Explaining the Vote Share of Candidates—With an Alternative Criminal Charge Dummy

	I	II		III	IV
	All Candidates			No Independent Candidates	
Education	0.102***	0.103***		0.144***	0.145***
	[7.50]	[7.53]		[6.08]	[6.11]
Wealth log*criminal dummy (more than 2 cases)	0.105***	0.103***		0.137***	0.136***
	[2.71]	[2.66]		[3.39]	[3.32]
Incumbent MP	0.837***	0.836***		0.781***	0.780***
	[9.58]	[9.55]		[8.81]	[8.79]
Criminal dummy (more than 2 cases)	-1.378**	-1.188*		-1.790***	-1.629**
	[2.33]	[1.90]		[2.81]	[2.45]
Candidates with charges (among top 4)*criminal dummy (more than 2 cases)	0.271***	0.250***		0.183*	0.167*
	[3.68]	[3.30]		[1.88]	[1.68]
Relative Wealth (absolute)	0.003**	0.003**			
	[2.27]	[2.27]			
State Incumbent Party	1.808***	1.805***		1.777***	1.775***
	[24.53]	[24.44]		[24.59]	[24.47]
Cognizable crime*Criminal dummy (more than 2 cases)		-0.092			-0.083
		[1.12]			[0.79]
Relative Wealth (absolute, no independent candidates)				0.001***	0.001***
				[3.86]	[3.86]
Constituency Fixed Effects	Yes	Yes		Yes	Yes
Party Fixed Effects	Yes	Yes		Yes	Yes
Observations	6,732	6,732		3,621	3,621
Adjusted R-squared	0.78	0.78		0.77	0.77

*,**, *** indicate that the coefficient is significantly different from zero at 10, 5 and 1 percent levels of significance respectively. Robust t statistics are reported in parentheses. The regression equation is in equation 8 and the variables are defined in Appendix A1 and in the text.

Table 7: Explaining the Vote Share of Candidates: Robustness Tests
(All Candidates)

	I	II	III	IV
Criminal dummy	-1.49***	-1.19***		-1.239***
	[4.07]	[3.36]		[3.20]
Wealth log*criminal dummy	0.115***	0.092***		0.114***
	[4.54]	[3.80]		[4.78]
Incumbent MP	0.834***	0.819***	0.82***	0.836***
	[9.53]	[9.38]	[9.42]	[9.54]
State Incumbency	1.872***	1.785***	1.81***	1.793***
	[24.2]	[24.4]	[24.84]	[24.51]
Education	0.099***	0.099***	0.10***	0.100***
	[7.29]	[7.24]	[7.44]	[7.29]
Candidates with charges (among top 4)*criminal dummy	0.264***	0.240***		
	[4.77]	[4.33]		
Relative wealth (wealth/average wealth of others)	0.003**	0.003**	0.003**	0.003**
	[2.26]	[2.27]	[2.28]	[2.31]
State Incumbency*criminal dummy	-0.56***			
	[3.47]			
Age		0.005***		
		[4.06]		
Gender		0.048		
		[0.92]		
Number of cases, log			-0.647**	
			[2.42]	
Wealth log*number of cases log			0.067***	
			[3.85]	
Number of cases against other candidates log* cases against the candidate log			0.064*	
			[1.72]	
Candidates with charges (all)*criminal dummy				0.038
				[1.22]
Constituency Fixed Effects	Yes	Yes	Yes	Yes
Party Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	6,732	6,731	6,732	6,732
Adjusted R-squared	0.782	0.782	0.78	0.781

*,**, *** indicate that the coefficient is significantly different from zero at 10, 5 and 1 percent levels of significance respectively. Robust t statistics are reported in parentheses. The regression equation is in equation 8 and the variables are defined in Appendix A1 and in the text.

Table 8: Explaining the Vote Share of Candidates: More Robustness Tests

	I	II	III	IV
Criminal dummy	-1.13***	-1.894*	-1.19***	-1.39***
	[2.99]	[1.95]	[2.86]	[3.88]
Wealth log*criminal dummy	0.112***	0.146**	0.088***	0.092***
	[4.64]	[2.03]	[3.18]	[3.73]
Candidates with charges (at least 3 percent vote share)*criminal dummy	0.079			
	[1.39]			
Relative wealth (candidates with at least 3 percent vote share)	0.011***			
	[3.85]			
Candidates with charges (among top 4)*criminal dummy		0.308*	0.24***	0.25***
		[1.95]	[3.85]	[4.53]
Relative wealth		0	0.003**	0.003**
		[0.17]	[2.41]	[2.28]
Incumbent MP	0.825***	0.66***	0.89***	0.85***
	[9.39]	[3.24]	[8.6]	[9.69]
State incumbency	1.791***	1.72***	1.774***	1.79***
	[24.58]	[11.27]	[20.63]	[24.44]
Education	0.10***	0.146***	0.083***	0.10***
	[7.28]	[5.12]	[4.83]	[7.36]
Serious Crime Dummy				0.262**
				[2.15]
Constituency Fixed Effects	Yes	Yes	Yes	Yes
Party Fixed Effects	Yes	Yes	Yes	Yes
State-Party Fixed Effects	No	No	No	No
Observations	6717	2021	4711	6732
Adjusted R-squared	0.781	0.764	0.786	0.782

*,**, *** indicate that the coefficient is significantly different from zero at 10, 5 and 1 percent levels of significance respectively. Robust t statistics are reported in parentheses.

Table 9: Voter Turnout and the Number of Candidates with Criminal Charges
(Dependent Variable: Voter Turnout)

	I	II	III	IV	V	VI
Number of Candidates with at least two Charges	0.079	0.341	0.227			0.201
	[0.28]	[1.21]	[0.79]			[0.72]
Total Candidates		-0.23***	-0.24***			-0.23***
		[4.14]	[4.30]			[4.18]
Number of Candidates with at least Two Charges from a Large Party				0.86**	1.00**	-2.23***
				[2.08]	[2.38]	[2.85]
Total Candidates from a Large Party					-0.579*	0.514
					[1.86]	[0.41]
Dummy for a Constituency Reserved for SC Candidates			-2.25***	-1.25	-1.39*	
			[2.91]	[1.65]	[1.83]	
Dummy for a Constituency Reserved for ST Candidates			1.137	2.61**	2.74**	
			[0.94]	[2.24]	[2.31]	
Literacy						-0.09**
						[2.09]
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	506	506	506	506	506	506
Adj. R-squared	0.78	0.788	0.792	0.784	0.785	0.794

*,**, *** indicate that the coefficient is significantly different from zero at 10, 5 and 1 percent levels of significance respectively. Robust t statistics are reported in parentheses.

Table 10 : Candidate Wealth and Criminal Dummy
(Dependent Variable: Candidate wealth)

	I	II	III	IV	V
Criminal dummy	0.778***	0.763***	0.638***	0.747***	0.691***
	[7.55]	[7.43]	[6.23]	[5.68]	[5.87]
Dummy for national party	2.58***	2.07***			1.912***
	[42.70]	[31.04]			[22.08]
Dummy for state party	1.88***	1.604***			1.402***
	[19.81]	[17.32]			[12.92]
Education		0.44***	0.358***	0.468***	0.449***
		[17.34]	[13.69]	[8.82]	[12.18]
Incumbent candidate		1.14***	0.876***	1.09***	1.04***
		[11.75]	[8.47]	[10.50]	[10.72]
Constituency fixed effects	Yes	Yes	Yes	Yes	Yes
Party fixed effects	No	No	Yes	No	No
Observations	7,177	6,737	6,733	2,077	3,633
Adj. R-squared	0.251	0.305	0.374	0.213	0.289

*,**, *** indicate that the coefficient is significantly different from zero at 10, 5 and 1 percent levels of significance respectively. Robust t statistics are reported in parentheses