Incumbents and Criminals in the Indian National Legislature

Toke Aidt, Miriam A. Golden and Devesh Tiwari

13 September 2011

CWPE 1157

Incumbents and Criminals in the Indian National Legislature[1]

Toke Aidt

University of Cambridge

toke.aidt@econ.cam.ac.uk

and

Miriam A. Golden

University of California at Los Angeles

golden@ucla.edu

and

Devesh Tiwari

University of California at San Diego

dtiwari@ucsd.edu

September 13, 2011

Version 5.1. Comments welcome.

[1]Portions of the work reported here were presented at the 2009 and the 2010 Annual Meetings of the American Political Science Association, the 2011 meetings of the International Society for New Institutional Economics, the University of California at Riverside, Georgetown University, the World Bank (New Delhi), the Johns Hopkins School of Advanced International Studies, the University of Virginia, Loyola Marymount University, and the University of California at Los Angeles. For comments, we are grateful to Scott Desposato, John Echeverri-Gent, Philip Keefer, Stuti Khemeni, and Kevin Morrison. Funding for the research reported here was provided to Miriam Golden by the Academic Senate, the International Institute, and the Center for International Business Education and Research at the University of California at Los Angeles.

Abstract

Utilizing data on criminal charges lodged against candidates to the Fourteenth and Fifteenth Lok Sabha, India's lower house of representatives, we study the conditions that resulted in approximately a quarter of members of parliament elected in 2004 and in 2009 facing or having previously faced criminal charges. Our results document that Indian political parties are more likely to select alleged criminal candidates when confronting greater electoral uncertainty and in parliamentary constituencies whose populations exhibit lower levels of literacy. We interpret the decisions of political parties to enlist known criminals as candidates as a function of the capacity of these candidates to intimidate voters. To substantiate this, we show that criminal candidates depress electoral turnout. In addition, our results suggest that India's well-known incumbency disadvantage stems from the superior electoral performance of allegedly criminal candidates, who drive parliamentary incumbents from office. Our study raises questions for democratic theory, which claims that electoral competition improves accountability, and for the future of the Indian polity, which is experiencing a growing criminalization of the national political arena.

1 Introduction

Why would a political party in a competitive democratic system recruit a known criminal to run for national public office? There are varied puzzling aspects to this phenomenon. Not only is it puzzling that a party would field a candidate who faces criminal charges since, especially if the charges involve acts of non-familial violence, this ought to prove electorally disadvantageous, it is puzzling that voters, instead of repudiating such a candidate for public office, would elect him. And it is genuinely astonishing that this would occur in fully a quarter of 543 single-member parliamentary constituencies not once but twice in a row in a highly competitive multiparty system of a well-established, long-standing democratic polity known for its vigilant civil society and aggressive free press.

We analyze data on the 2004 and 2009 Indian national elections, the first and second legislative elections in which candidates to the national parliament were required by a 2003 Supreme Court ruling to file sworn affidavits that included, among other pieces of information, declarations of criminal records and then-current indictments.[1] Our dataset contains information drawn from all the affidavits filed by the more than $12,000$ parliamentary candidates in these two elections. Empirically, our goal is to isolate the main electoral, demographic, and other factors associated with the appearance of allegedly criminal candidates on the ballot. Theoretically, we seek to shed light on how electoral democracies generate situations in which accountability appears so seriously compromised.

Our story is simple although troubling. We contend that alleged criminals are selected by political parties because these candidates specialize in the use of violent pre-electoral tactics. We document accordingly that allegedly criminal candidates appear on the ballot when the parliamentary seat is highly contested, where there are more illiterate and therefore more vulnerable voters, and that these candidates depress electoral turnout. We interpret all this as support of our claim that these candidates intimidate opposition voters from going to the polls. In addition, we show that these candidates successfully drive incumbents out of office. We interpret this as evidence that holding a seat in the national legislature in India is highly valued — indeed, so highly valued that individuals regularly use violence to obtain one.

Substantively, who wins an election is of greater interest than who runs for office. If alleged criminals are listed on the ballot but win only a handful of seats, the phenomenon would be of only marginal interest

[1] We discuss below the validity of the information reported. We do not distinguish between persons who report having been convicted of a criminal offense from those who report facing indictment, and we refer interchangeably to persons who report facing indictment on criminal charges as "self-reported," "acknowledged," or "publicly known" criminals.

and of limited political importance. But in fact, the reverse is the case. As the data featured in Table 1 shows, in both 2004 and 2009 Indian candidates to the lower house whose affidavits report criminal charges have a much larger likelihood of winning than other candidates. In 2004, more than a quarter of those facing criminal charges won their seats compared with a success rate of only 8 percent for other candidates. The 2009 elections were much more competitive — the total number of candidates rose 50 percent over the 2004 total — but even so, allegedly criminal candidates won 14 percent of the time compared with a success rate of 6 percent for those who did not report criminality on their affidavits. In both elections, candidates reporting criminal charges were two to three times more likely to win than others. This obvious electoral advantage makes candidates reporting criminal charges unusually attractive to political parties. Although in most circumstances we expect criminality to constitute an electoral hindrance, in contemporary India the reverse appears to be the case. We seek to understand why.

[Table 1 about here]

Our paper is organized as follows. First, we briefly review various strands of literature related to our study. Second, we study a formalization of the party's decision to select a criminal candidate. We then turn to statistical estimations. After discussing our data and explaining how we operationalize our variables, the first set of statistical estimations that we report studies the likelihood that an allegedly criminal candidate appears on the ballot. We then study whether there are theoretically-expected patterns in the numbers of self-reported criminals selected to stand in each constituency. A fifth section employs a regression discontinuity procedure to investigate the electoral success of self-reported criminal candidates and their impact on the electoral success of incumbents. A final set of regressions examines the effects of allegedly criminal candidates on electoral turnout. Finally, a concluding section raises unanswered questions and suggests future research paths.

Our study has two primary theoretically unexpected implications. First, as we discuss in the next section, prior research documents that once information about political malfeasance is released and disseminated to voters, they will use the information to elect honest challengers or, what comes to the same thing, public officials, anticipating voter backlash, will substantially reduce the frequency of corrupt or malfeasant activities. We observe voters successfully using new information to reduce the frequency and scope of political corruption in settings as diverse as Uganda (Reinikka and Svensson 2005), Brazil (Ferraz and Finan 2008), and Italy (Chang, Golden and Hill 2010). On the basis of such studies, we expect that the release of public

information about criminal charges lodged against members of the Indian political elite should reduce the number of alleged criminals elected — if not immediately, then certainly by the second election in which candidates file affidavits reporting their legal profiles. Instead, we observe a nearly uniform persistence in the rate of self-reported criminality among the elected. This suggests that information alone is inadequate in reducing political corruption and criminality, and that prior studies may have omitted important variables from consideration. Other research that we refer to below indicates that ethnic or programmatic attachments to a party may dissuade voters from turning against that party even if its leaders engage in corrupt, criminal, or otherwise malfeasant behavior. Our research refocuses attention on the importance of violence as well. It thereby builds on other studies of the strategic use of violence in electoral contests, especially Wilkinson (2004): where voters are impoverished, illiterate, and easily coerced, organized groups may strategically deploy violence to reelect corrupt or criminal political representatives.

A second implication of our study speaks to the literature on legislative incumbency. Multiple papers document an incumbency disadvantage for legislators elected to the federal or to the state level in India. Our results show that the Indian incumbency disadvantage, which until now has lacked explanation, is linked to the high rate of allegedly criminal candidates. Incumbency disadvantage occurs because self-reported criminals boast an extraordinary political advantage against even experienced politicians. In India, we appear to be witnessing the nearly wholesale take-over of large chunks of an established democratic political system by persons with criminal records. Moreover, this is occurring even as Indian party politics becomes increasingly competitive. This suggests that democratic accountability may encounter unexpected obstacles that have not been previously identified and that are not well understood or analytically appreciated.

2 Related Literature

Our paper builds on four distinct strands of literature: first, prior studies that also use Indian candidate affidavit information; second, studies of electoral responses to political malfeasance; third, studies of incumbency disadvantage in India; and, fourth, studies of political competition and accountability. We briefly review these various classes of studies.

We have identified three other scholarly studies that analyze Indian candidate affidavits, although for somewhat different purposes than ours or at levels of government below the national, which we study. Chemin (2008) studies local outcomes after the 2004 parliamentary elections and finds that bureaucratic

corruption falls but poverty rates increase in constituencies that elect self-reported criminals.[2] These outcomes are consistent with a theoretical framework in which criminal elements enter and capture the polity; such a view expects political control by organized crime to reduce the unorganized criminality of the petty bureaucrat while also encouraging economic inequality. Banerjee and Pande (N.d.) examine the 2004 affidavits of candidates to the legislative assembly in Uttar Pradesh, a state well known for a high level of political corruption and criminality, and find that parliamentarians are more likely to be alleged criminals in constituencies with more low-caste voters. The broad lines of their argument is that corruption and criminality among politicians in Uttar Pradesh have been due to the rise of low-caste and ethnic voting and that ethnic party politics creates incentives for political corruption. These findings are not inconsistent with ours. We find that areas with more illiterate voters (which broadly coincide with those with more low-caste voters) also see more alleged criminals on the ballot.

The study closest to ours, finally, is Vaishnav (2010), who uses a political selection framework to analyze the affidavit information of party-affiliated candidates to state (not federal) elections in India. His main argument is that allegedly criminal candidates are attractive to political parties because such candidates are self-financing, perhaps because of assets amassed thanks to criminal activities. However, in contrast to the results we report, Vaishnav finds no evidence that electoral competitiveness increases the likelihood that an alleged criminal appears on the ballot.

In addition to studies that analyze the affidavits of Indian parliamentary candidates, various studies, including Reinikka and Svensson (2005); Ferraz and Finan (2008); Bobonis and Cámara Fuertes (2009); Brollo (2009); Chang, Golden and Hill (2010); Banerjee et al. (2010), have analyzed how voters respond to revelations of wrongdoing on the part of public officials. They find that revelations of malfeasance by elected officials result in electoral retribution under certain conditions — for instance, when information about malfeasance is disseminated widely to voters. When it is not, elected officials with records of wrongdoing may be repeatedly reelected.

We highlight another reason why, even if they are informed of candidate malfeasance, voters may end up electing such individuals into public office. Information is not fully effective when voters are subject to physical intimidation. This may compromise political accountability by scaring voters into staying home on election day. If potential opponents of reputably malfeasant candidates do not turn out to vote for another

[2] In India, electoral districts for the national legislature are known as parliamentary constituencies. Throughout our paper, we employ this nomenclature since the term "district" is used in the Indian context to refer to administrative units.

candidate, the former may win the seat. Such practices are likely to be all too common in environments characterized by weak rule of law, endemic political corruption, and entrenched organized crime. Research on countries in subsaharian Africa has documented that voter intimidation and pre-electoral violence reduce turnout (Bratton 2008). Our study extends this line of investigation to the Indian context.

A third strand of relevant literature studies the electoral (dis)advantage of incumbent politicians in India. Various studies show that in India, unlike many other political systems, federal and state legislators face an incumbency disadvantage (Linden 2004; Nooruddin and Chhibber 2007; Uppal 2009; Chakrabarti, Gangopadhyay and Krishnan N.d.). This disadvantage is apparently recent and it overlaps in part with the growing turnover of legislative representatives in India and the decline of the Indian National Congress. Nooruddin and Chhibber (2007), for instance, show that electorally more volatile states in India also exhibit larger anti-incumbent swings. Likewise, Wilkinson (2007) contends that party volatility, electoral turnover, and clientelism have risen together in India since the late 1960s. Our work offers an entirely new understanding of these phenomena. We analyze incumbency disadvantage together with allegations of candidate criminality to estimate the effects of allegedly criminal opponents on incumbent reelection probabilities. We evaluate whether the proliferation of alleged criminals on the ballot is causally linked to the growing incumbency disadvantage in India, and we show that it is.

Finally, many studies of political competitiveness and accountability report that accountability is enhanced by electoral competition. The most compelling statement of this view is the observation advanced by Sen (1981) that famines occur only in non-democratic regimes where leaders do not face reelection prospects. In the Indian context, this argument has been extended by Besley and Burgess (2002), who show that public food distribution across Indian states in response to declines in grain production and flood damage is improved where newspaper circulation and political competition are greater. Our results fail to corroborate that political competition enhances accountability and responsiveness. Instead, we show that where partisan competition is more intense and when a party has less assurance about whether it will win the seat, that party is more likely to list an alleged criminal on the ballot. The aggregate outcome of intense partisan competition is thus to diminish political accountability.

We are not the first to observe that competitive electoral pressures may encourage wrongdoing by elected officials (Nyblade and Reed 2008). However, the mechanism that we identify to explain why Indian political parties list allegedly criminal candidates is new. We contend that the self-reported criminal candidates that we study are distinguished not only by their involvement in illegal fund-raising and any resulting financial

advantage but also by expertise in the use of violence. Although we cannot observe this directly, we test some implications of this theory in more than one way using as much data as we have been able to assemble.

3 A Model of Candidate Selection and Criminality

The standard probabilistic voting model (such as Persson and Tabellini (2000)) studies political competition as a function of the choice of platform by political parties. In these models, voters have variable degrees of attachment to their party's program, and as a result, their votes can potentially be "bought" with larger or smaller other (material) benefits and inducements. Given this framework, debate has centered on whether parties provide such material inducements to voters with lesser or greater programmatic attachments — so-called "swing" or "core" voters (Dixit and Londregan 1996; Stokes 2005). Studies of vote buying assess effects on party switching of these inducements, taking the decision to vote as given.

Although vote-buying is an important consideration in many contexts, our focus lies with candidate selection. We think of platform choice as determined centrally by the party organization for all constituencies and so as given for any particular constituency, and we study mechanisms used by political parties to manipulate election turnout. The mechanisms we focus on are violence and intimidation. We might imagine that parties have exhausted their resources to buy votes and so turn their attention to affecting the decision to vote. More realistically, perhaps, given evidence in the Indian context that parties systematically fail to utilize all possible sources of material inducement that are easily available with which to buy votes (Keefer and Khemeni 2009), it may be the case that it is less expensive for them to discourage opposition voters from coming to the polls than it is to attract swing voters to shift their allegiance from another party or even to induce their own core supporters to turn out to vote (on using material benefits to encourage turnout, see Nichter (2008)).

We present a formal model designed to illustrate possible links between electoral competition, voter turnout, and the incentives of political parties to field candidates who publicly acknowledge facing criminal indictment. The model highlights what we believe is the central mechanism that makes known criminal candidates attractive to political parties: their capacity to intimidate opposition voters into staying home on election day. Formalizing these issues allows us to develop testable hypotheses that in turn facilitate interpreting and understanding the empirical results that we report in subsequent sections. In the model, we assume that self-reported criminals are in fact actually engaged in criminal activities; in reality, of course,

not all may be. For the results of the formal model to be empirically meaningful, it is not necessary that every self-reported criminal in fact engage in criminal activities. There need only be a positive correlation between the two.

3.1 Assumptions

We consider two political parties $k \in \{A, B\}$ that compete in a constituency for a seat in parliament. The parties have fixed ideological positions and we take their electoral platforms as given. Each party has a natural support base of voters who for ideological or identity reasons are committed to support it over other political parties. We denote the number of party k supporters by N_k. Since voting is costly, only a fraction of the supporters will turn out to vote on election day. The turnout cost is, for simplicity, the same for all voters and denoted by $c > 0$. It captures the opportunity cost of voting net of any electoral intimidation. We assume that voting is an expressive act and that each voter gets some benefit out of the act of voting irrespective of whether the vote influences the outcome or not.[3]

Potential candidates are of two different types, which we refer to as criminal and non-criminal, indexed by C and N.[4] The two parties simultaneously select the type of candidate they each field before the election. On election day, voters decide whether to go to the polls to vote for their favored party or stay home. We refer to this as the candidate fielding game.

[3] This formulation circumvents the paradox of voting. For a discussion of expressive political behavior see, for instance, Jennings and Hamlin (2011) and Hillman (2010). The benefit of voting is distributed according to a uniform distribution on the interval $[0, \theta_k]$, where θ_k is specific to a party and may differ across parties. The parameter θ_k captures how much value the average voter of each party attaches to voting. A party whose average voter has a high benefit to voting for her party is more likely to vote even in the face of increased costs, including violence and intimidation. The differences in the value that each party's supporters attach to voting could be related to the degree of ideological or ethnic and caste activation by each party.

[4] We focus on the demand side — that is, on why parties select candidates with criminal records — taking the supply of criminal candidates as given. However, there is good reason to believe that public office is highly valued in India. Recent research shows that winning office increases the average candidate's family assets by 28 percent over five years (Bhavnani 2011). Another indication of the value of public office in India is the increasingly extensive use of nepotism on the part of existing politicians in advantaging their children and other family members in entering politics. This has been widely reported (e.g. "On Cluttered Ballots of India, Families Proliferate," *New York Times*, Oct. 11, 2009), and journalist Patrick French has compiled data showing that although more than a quarter of the Lok Sabha members elected in 2009 entered politics through familiy connections (p. 107), 100 percent of the members of parliament under the age of 30 and 65 percent of those in the 31-40 age range had "in effect inherited a seat" (French 2011, p. 110). These considerations suggest that financial incentives contribute to the decisions by criminal elements to enter the political arena in India.

We conceptualize the electoral benefit to a party of fielding a known criminal rather than a non-criminal candidate as the access enjoyed by the former to an organized network of persons who can target opposition voters with threats of or the actual use of violence, thereby intimidating some of them not to show up at the polls.[5] In particular, we assume that a criminal candidate through this mechanism increases the cost of voting for opposition voters. This reduces electoral turnout among opposition voters, potentially swinging the election in favor of the party represented by the criminal. Criminals, however, have limited resources to produce electoral violence and therefore to intimidate. We capture this by assuming that a criminal candidate can at most increase the cost of voting for each opposition voter by $t > 0$. The parameter t can be interpreted as a measure of the productivity of violence. Intuitively, we might imagine that if a thug who is associated with a criminal politician beats up one opposition voter, he scares off another nine from showing up at the polls on election day, but does not scare off all opposition voters or even another 99 of them. Non-criminal candidates do not have access to this technology and cannot scare off any voters.[6] To ensure some turnout even with criminals in the race, we assume that $\theta_k > c + t$ for all k.

Technologies of intimidation will be more effective in some parliamentary constituencies and among some voting populations. In particular, constituencies with higher illiteracy rates allow criminal candidates to depress turnout more effectively. Our reasoning is that literate voters are endowed with a whole panoply of resources, including access to information and access to police protection, that together offers them greater resilience in the face of electoral violence. Citizens with low socio-economic status are less likely to access police services even when confronting circumstances that would warrant using them. Higher status voters,

[5] Other work supports the hypothesis that violence reduces turnout. Research in Africa documents the effects of voter intimidation on turnout; in that context, as perhaps also in India, violence is especially effective in reducing turnout among the rural poor, where even highly selective intimidation penetrates larger communities of voters. See Bratton (2008).

[6] We could assume that voters, in general, do not like known criminal candidates and that a party that selects such a candidate erodes its own support base. This feature is not essential for our results as long as criminal candidates compensate for this loss of votes through their capacity to reduce the turnout of opposition voters. To simplify the analysis, we ignore this effect but it would be straightforward to incorporate it into the model. Empirically, we have no way to distinguish reductions in turnout due to intimidation by a criminal candidate of supporters of another party and abhorrence of criminality on the part of supporters of the party that selects the criminal. However, if none of the effect of including a criminal on the ballot came from the intimidation of opposition voters, a party would never select a criminal, since doing so would reduce net turnout for that party. Therefore, our theory allows us to assume that criminals must engage in violence and intimidation.

by contrast, are more knowledgeable about their legal rights, economically less vulnerable, and generally more capable in their interactions with providers of government services. As a result, we argue, literate voters are less likely to be intimidated by threats of violence and to stay home on election day, which makes t an increasing function of constituency-level illiteracy.[7]

We assume that there is a small (but strictly positive) cost associated with fielding a criminal candidate. We denote the cost by μ and assume, for simplicity, that it is the same for both parties.[8]

In the absence of criminal candidates in the race, the number of voters turning out in support of party k is $N_k(1 - \frac{c}{\theta_k})$ and party A wins if

$$\Delta \equiv (1 - \frac{c}{\theta_A}) - \gamma(1 - \frac{c}{\theta_B}) > 0 \tag{1}$$

where $\gamma = \frac{N_B}{N_A}$ is the relative size of group B.[9] We assume that party A wins the seat with certainty if both parties field non-criminal candidates (i.e., $\Delta > 0$). We refer to party A as the top dog and party B as the underdog in the race. The closeness or contestability of the election is determined by the relative size of the two parties' groups of supporters. We can, using equation (1), express the assumption that party A is the top dog as a condition of relative group size:

$$\gamma < \frac{\theta_B(\theta_A - c)}{\theta_A(\theta_B - c)} \equiv \widehat{\gamma}. \tag{2}$$

As γ is increasing (and approaches $\widehat{\gamma}$),then Δ is close to zero and the race is close; conversely, as γ is decreasing (and approaches zero), then Δ is large and the seat is safe for party A. We can therefore use γ as an index of electoral competition: a low value corresponds to a situation with little competition whereas

[7] Variations in constituency-level income levels are also likely to be systematically related to t, with poorer voters being easier to intimidate. Since constituency-level data on income is unavailable in India, we do not stress this aspect.

[8] The fielding cost represents whatever disadvantages a party suffers irrespective of its electoral success in the particular constituency as a consequence of allowing a criminal on the ballot under its symbol. This, among other things, includes reputation costs for the party nationally as well as the inconvenience for the local party organization of having to associate with criminals. The cost can be arbitrarily small and it plays a role only as a tie breaker.

[9] If $\Delta = 0$, we assume that the election is determined by a toss of a coin.

a high value corresponds to a situation in which the race, in the absence of any criminal candidates, is close and the seat is therefore competitive.

Finally, we assume that the value to a party of wining the seat is $M > 0$. It is reasonable to assume that the benefit of wining the seat for the party is always greater than the cost of fielding a criminal ($M - \mu > 0$). If not, there would be no reason to ever do so.

3.2 Analysis and Results

A criminal candidate can increase the voting cost of opposition voters by t. The effect of this on the number of voters intimidated depends on how attached the targeted voters are to their parties (i.e., on how big θ_k is). The maximum fraction of the supporters of party A that a criminal candidate fielded by party B can intimidate not to vote is $\Delta_B = \frac{t}{\theta_A}$ and similarly $\Delta_A = \frac{t}{\theta_B}$ is the maximum fraction of party B voters that a criminal candidate fielded by party A can intimidate into not voting. Depending on how close the race is to begin with (i.e., on the size of Δ and γ), three different situations can emerge. We say that the seat is *competitive* if party B — the underdog — can swing the election by fielding a criminal, i.e., if $\Delta - \Delta_B < 0$, and that the seat is *defensible* if party A — the top dog — can swing the election back if, in response, it fields a criminal, i.e., if $\Delta - \Delta_B + \gamma\Delta_A > 0$ and *indefensible* if not. If party B cannot swing the election by fielding a criminal, we say that the seat is *safe*. For ease of reference, we refer to these three situations or regimes as safe seat, competitive defensible seat, and competitive indefensible seat. With these preliminary remarks, we are ready to characterize the outcome of the candidate fielding game in the three regimes in which the two parties simultaneously field candidates. We do so intuitively here (the technical details appear in Appendix B).

1. Safe seat: $\gamma < \widehat{\gamma} - \frac{\theta_B t}{\theta_A(\theta_B - c)} \equiv \underline{\gamma}$. Even if party B fields a criminal candidate and party A does not, party A wins the seat. The candidate fielding game thus has a unique pure strategy Nash equilibrium in which both parties field non-criminals. The intuition is that fielding a criminal candidate is a

dominated strategy for party B. Doing so would not swing the election but would impose on party B the (small) fielding cost. Hence, party B does not field a criminal in this regime and neither does party A.

2. Competitive defensible seat: $\underline{\gamma} \leq \gamma < \frac{\theta_B(\theta_A - c - t)}{\theta_A(\theta_B - c - t)} \equiv \overline{\gamma}$. Party B can swing the election if it fields a criminal candidate so long as party A does not. Party A can, however, defend the seat if it, in response, also fields a criminal. The candidate fielding game has a unique Nash equilibrium in mixed strategies where the two parties each field a criminal with positive probability.[10] The probability that party A fields a criminal is $\frac{M-\mu}{M}$ and the probability that party B fields a criminal is $\frac{\mu}{M}$. The equilibrium outcome, then, may be that one, both or neither of the parties fields a criminal candidate. The probability that we observe at least one criminal candidate is $\frac{\mu(M-\mu)}{M^2} > 0$.

3. Competitive indefensible seat: $\theta_B > \theta_A$ and $\widehat{\gamma} > \gamma \geq \frac{\theta_B(\theta_A - c - t)}{\theta_A(\theta_B - c - t)} \equiv \overline{\gamma}$. Party B can swing the election if it fields a criminal candidate and party A does not, and party A is unable to defend the seat even if it offers a ticket to a criminal candidate. As a result, the candidate fielding game has a unique pure strategy Nash equilibrium in which party B fields a criminal candidate whereas party A does not. This is an equilibrium because party B — the underdog — by fielding a criminal wins the election with a net gain of $M - \mu > 0$ irrespective of what party A — the top dog — does. Given this, the best response for party A is to avoid the (small) fielding cost associated with a criminal candidate and to select a non-criminal instead. The condition that $\theta_B > \theta_A$ is required to ensure that $\overline{\gamma} < \widehat{\gamma}$. Intuitively, for the seat to be indefensible, the supporters of party A must be relatively easy to intimidate because they, on average, get relatively little benefit from voting.

[10] The reason the game does not have a pure strategy Nash equilibrium is that party A will field a criminal to defend the seat if party B fields one. Given that, party B does not want to field a criminal. But then party A prefers not to field a criminal either which, in turn, gives party B an incentive to do so, etc.

3.3 Empirical Implications

The model generates four predictions that we investigate empirically. The first is that criminal candidates are more likely to be given tickets in races where the seat is competitive, irrespective of whether it is defensible or indefensible. Hence, when the seat is safe, criminals do not appear on the ballot. Figure 1 illustrates this result. On the horizontal axis, we record the value of γ which controls the degree of electoral competition, with larger values corresponding to a closer race and thus more electoral competitiveness. On the vertical axis, we record Δ, $\Delta - \Delta_B$ and $\Delta - \Delta_B - \gamma\Delta_A$ for the case where $\theta_B > \theta_A$.

[Figure 1 about here]

Consider an electoral district or parliamentary constituency located at point 1 with a value of γ close to zero. From the point of view of party A, whose support base is much larger than that of party B, this seat is completely safe: no matter what party B does, it cannot swing enough votes to win. In constituencies such as this, characterized by low electoral competition, criminal candidates are not fielded by either party. In contrast, a constituency located at point 2 houses a more competitive race. Here, although party A is the favorite, party B can, in principle, swing the election if its criminal candidate does his work ($\Delta - \Delta_B < 0$). However, the position of party A is sufficiently safe to make the seat defensible ($\Delta - \Delta_B - \gamma\Delta_A > 0$). In constituencies like this, parties occasionally field criminals and we may observe situations with two criminals, one criminal, or no criminal on the ballot.

When competition becomes even more intense, the seat may become indefensible ($\Delta - \Delta_B - \gamma\Delta_A < 0$). This is the case for a constituency located at point 3 in the figure. In this locality, party A cannot counter the challenge from party B by fielding a criminal. It thus refrains from doing so. In a constituency such as this, we expect to observe a race with one criminal on the ballot.

The testable implication of this framework is that a political party is more likely to give a ticket to a criminal candidate in a constituency where it faces more competition for a seat. We model this empirically by assessing systematically whether, all else equal, parties are more likely to list self-reported criminals

where they perceive the seat to be competitive.

A second prediction of the formal model is that criminals are more likely to be fielded in constituencies where criminality is more effective in reducing voter turnout, i.e., where t is high. An increase in t reduces the threshold $\underline{\gamma}$ and makes it more likely that a criminal candidate is fielded. Recall that the size of t is empirically related to constituency characteristics such as the degree of literacy, average income, etc. Accordingly, the prediction that we test is that criminals are more likely to be fielded in constituencies with more illiterate voters.[11]

We model the first two predictions using a multiple regression framework which incorporates competitiveness and literacy simultaneously into models that estimate the probability of observing a known criminal on the ballot.

A third prediction of the model concerns the numbers of acknowledged criminals who are likely to be listed in any one constituency. Unlike other models that study when criminal candidates are listed (for instance, Vaishnav (2010)), our's does not predict that a specific pattern to the number of criminal candidates observed in parliamentary constituencies. In particular, we do not expect to see criminal candidates appearing systematically in pairs, as occurs when a party always matches another party's decision to list a self-reported criminal by also selecting an alleged criminal. Instead, our model predicts that we should observe the full range of possible outcomes of candidate selection: constituencies with symmetric outcomes (none or two alleged criminal candidates on the ballot) and those with asymmetric outcomes (only one alleged criminal on the ballot). We assess this empirically by examining the distribution of the numbers of self-reported criminals on the ballot across India's parliamentary constituencies.

Fourth, our model carries implications for longevity in office. In some contexts, it is natural to think about party A's initial electoral advantage as an incumbency advantage. If we do so, our model suggests

[11] An additional prediction of the model, which we do not test, is that criminals will be given tickets when the average voter receives a relatively low benefit from voting; i.e. γ is increasing in θ_A and θ_B. The average benefits of voting could be related to the degree to which other political parties activate ethnic, religious, or, in the Indian context, caste, attachments. In other words, we could hypothesize that criminal candidates are more likely to be fielded in constituencies where caste or religiously based parties compete against opponent parties which are not not caste or religiously distinctive and where voters have a low average θ.

that acknowledged criminal candidates fielded by opposition parties erode whatever incumbency advantage exists. In other words, even if the incumbent party has an ex ante advantage, perhaps because of the size of its natural support base in the constituency, the fact that opposition parties can (and do) field criminals to intimidate the incumbent's voters may turn the ex ante incumbency advantage into an ex post incumbency disadvantage. This effect arises in cases characterized by what we have labeled competitive indefensible seats. To examine this prediction empirically, we utilize a regression discontinuity design to study whether the proliferation of self-reported criminals on the ballot systematically disadvantages incumbent legislators.

In addition to the four predictions derived from the formal model, we also assess an empirical implication of an assumption of our model. Our theory assumes that criminal candidates are distinguished by their use of technologies of violence in their interactions with voters. Although Vaishnav (2010) reports that allegedly criminal candidates in India possess unusually high levels of wealth, wealth is unlikely to distinguish the criminal. Successful business people and Bollywood actors also bring considerable financial resources into the political arena. If wealth were the specific feature that political parties sought in candidates, perhaps because it allowed them to self-finance their electoral campaigns, parties would select wealthy candidates who did not carry with them the obvious disadvantages of the criminal: namely, that some voters find criminality morally distasteful and will therefore avoid voting for the criminal, that criminality carries some negative reputational costs for the party, and that other elected party officials will be forced to do more work to compensate for the lack of political commitment of the criminal politician.[12] To assess our assumption that the use of violence is the specific characteristic distinguishing the criminal candidate, we investigate whether parliamentary constituencies with at least one criminal on the ballot see a reduction in turnout.

To summarize, we study empirically the following:

1. Parties are more likely to field criminal candidates when they are close to winning or losing a seat;

[12] Tiwari (In Progress) documents that members of the Lok Sabha with criminal indictments have worse parliamentary attendance records than other legislators.

2. Parties are more likely to field criminal candidates when there is a higher proportion of illiterate voters in the constituency;

3. Criminal candidates appear on the ballot as singletons or as pairs, and neither configuration dominates;

4. Criminal candidates erode incumbency advantage;

5. Criminal candidates reduce electoral turnout.

Before proceeding, we discuss why we study the listing of candidates rather than the winning of elections and related endogeneity issues. We theorize that the relationship between fielding a candidate who reports a criminal indictment and the candidate's subsequent electoral success is endogenous. Because candidate selection is a calculated decision on the part of the sponsoring political party, only indicted candidates who have the skills necessary to win an election are given tickets to run. They therefore differ from other alleged criminals in that they have higher levels of political skill. Even if voters experience less utility if they are represented politically by an alleged criminal (hence reducing the likelihood they vote for one), parties compensate by only fielding especially skilled candidates who are able to overcome this disadvantage. Given this selection bias, we should not be surprised that alleged criminals are more likely to win elections; they win because they are politically skilled.

Correcting for this bias would require data we do not have and that we believe impossible to obtain: a complete list of all *potential* candidates, their criminal, financial, and political histories, and whether they were given a ticket. Substantively speaking, the correction would allow us to know whether, all else equal, having a criminal record increases or decreases the probability of winning a seat. The "all else equal" condition is important. This correction would allow us to estimate the causal impact of the reported criminal record itself on electoral success. In essence, then, the correction would allow us to estimate the preferences of Indian voters and to ascertain the extent to which they view self-reported criminality as a positive or

negative candidate characteristic.

But even if it were possible to undertake such an analysis, such is not the theoretical objective of this paper. We do not seek to determine whether or not the Indian voting public prefers known criminal candidates, all else equal. The argument that we make here is that the "all else equal" condition does not apply. Self-reported indicted candidates are fundamentally different from other candidates, and they are also different from alleged criminals who do not enter politics. That is why the former are selected by political parties to run and why they win so often. Self-reported criminal candidates possess distinctive combinations of illicit skills and political acumen. Parties field them precisely because of their abilities to parlay these skills into political success. This ability to win elections increases the attractiveness of acknowledged criminality to political parties, and is in itself a challenge to Indian democracy. So while we admit that our data are characterized by selection bias, we cannot nor do we wish to correct for it.

4 Data

We use information drawn from the affidavits filed by candidates to India's Fourteenth and Fifteenth national lower house of parliament (the Lok Sabha) to measure the criminality of Indian legislative candidates. We have coded all candidates by whether they report having been convicted of, or currently face indictment for, criminal activity.[13] Although the affidavits that Indian legislative candidates are legally required to file are self-reported, candidates who perjure themselves place themselves in legal jeopardy and may be disbarred from the race. Their electoral opponents, moreover, have obvious incentives to scrutinize their affidavits, as do the many anti-corruption nongovernmental organizations active on the ground. So although the measure of criminality that we use may underestimate the real extent of criminal indictment or conviction of parliamentary candidates, we believe the underestimation is modest and unlikely to be systematically related to the explanatory variables of interest. Moreover, to the extent that our data underestimates indictments of

[13] Due to the intricacies of the often multiple charges against individual candidates, we do not attempt to differentiate candidates according to the nature of the charges. However, we shortly present evidence that the charges often involve violence.

politicians by including among the self-reported non-criminal candidates some persons who are in fact criminals, this works against our hypotheses. If there is a bias to the underreporting of criminality, it therefore suggests that we are less likely to identify the relationships hypothesized.

We have merged affidavit data on candidates to India's national legislature with data on electoral outcomes in 2004 and 2009, as well as data on electoral turnout, information on the party affiliation of candidates, information on which parties are national or regional in scope, the incumbency status of each candidate, and theoretically relevant characteristics of the voting population in each constituency, namely, literacy rates. We have also merged in data for relevant constituency-level controls. We detail the operational indicators below.

4.1 The Institutional and Political Context

Figure 2 depicts a map of India's 543 parliamentary constituencies as of 2004, visually differentiated according to whether at least one alleged criminal appeared on the ballot or not.[14] In 2004, half of India's parliamentary constituencies witnessed known criminal candidates; in 2009, self-reported criminal candidates appeared on the ballot in nearly three-quarters of Lok Sabha constituencies, attesting to a diffusion of publicly acknowledged criminality in national political life.[15] A chi-squared test on the distribution of candidates with criminal records by state shows that criminality is not randomly distributed. However, a visual inspection of the map also shows that candidates facing criminal charges are widely dispersed across the subcontinent. Note that this contradicts conventional wisdom, according to which political criminality clusters in Bihar and Uttar Pradesh, two states in the north of the country.

[Figure 2 about here]

Although the data depicted in Figure 2 refer to all Lok Sabha candidates, our statistical analysis omits candidates who are unaffiliated with any of India's hundreds of political parties. We are interested in know-

[14] We have not been able to locate the GIS information required to create a similar map for 2009.

[15] These figures include self-nominated known criminals. For reasons we detail shortly, the analysis to follow includes only party-sponsored known criminals.

ing when parties list criminal candidates, and independent candidates by definition self-nominate. In addition to this theoretical justification for excluding them, including independent candidates introduces considerable uninformative noise into the statistical analysis. A third reason for excluding independent candidates is that they are politically irrelevant. In 2004, 43 percent of India's legislative candidates were unaffiliated with any political party and in 2009 independent candidates rose to comprise 47 percent of all candidates. Despite their numbers, unaffiliated candidates have almost no chance of winning seats: a mere five of the 2,385 independent candidates were elected in 2004, and in 2009 nine of 3,831 won their seats. As a result, only about 1 percent of the Lok Sabha's 543 members are unaffiliated with any party. Unaffiliated candidates are numerous but close to politically irrelevant. The median vote share collected by the winner and the first runner-up together is 87 percent, making most parliamentary races effectively two candidate contests.[16] The average vote share of the unaffiliated candidate was less than 1 percent in 2004 as well as in 2009, and no independent candidate won more than 7 percent of the vote in either election.

Examining only candidates who are put on the ballot by a political party and therefore excluding independent candidates, we find the ability of self-reported criminal candidates to gain seats in the legislature even more pronounced than for all candidates. We refer to the data presented earlier in Table 1. For partisan-affiliated candidates in 2004, reporting a criminal charge more than doubles the rate of winning a seat, increasing it from 15 to 36 percent. In 2009, the difference is less marked but even so the success rate for publicly-identified criminals is 20 percent compared with 11 percent for other candidates. For candidates listed as affiliated with one of India's numerous political parties, we find that reporting a criminal charge thus proves especially electorally advantageous.

4.2 Measurement and Definitions

In this section, we introduce the specific variables we use to operationalize the key concepts derived from

[16] Chakrabarti, Gangopadhyay and Krishnan (N.d., p. 4) note that the share of the vote received by candidates not among the top two averages 17 percent, enough to unsettle the final outcome, but this is much less true if we exclude the unaffiliated, who have virtually no chance of winning the seat to begin with.

our theory. These variables allow us to to test the empirical predictions and implications laid out above.

4.2.1 Criminality

Candidates for the two Lok Sabha elections that we analyze were required to file sworn affidavits in which they report criminal histories or pending criminal charges for any offence punishable with imprisonment of two years or more and that was brought prior to six months before filing for nomination.[17] The variable *CRIM* is coded 1 if the affidavit reports any charges against the candidate at any time regardless of the court's outcome and 0 otherwise. We use this as our dependent variable in the tests of the first two hypothesis outlined above, where the unit of analysis is a political party in a parliamentary constituency, and as an independent variable in the investigation of our assumption that criminals reduce turnout.

One potential objection to measuring criminality this way is that perhaps criminal charges encourage political entry because members of parliament (MP's) enjoy immunity from prosecution while in the legislature. However, Indian parliamentarians are not protected from prosecution while holding legislative office.[18]

A second potential objection to using charges drawn from affidavits is that the charges may be politically motivated rather than genuine. Skilled politicians may be charged with crimes by their rivals in efforts to discredit them politically. If this is the case, criminals could naturally be expected to enjoy an electoral advantage. This would not be because of activities involving voter intimidation or attributes specifically associated with the candidate's criminal status but rather because of selection effects. Skilled politicians are more adept at winning elections. If skilled politicians are more likely to be charged with crimes, charged candidates are more likely to win elections.

If this line of argument were true, then it would be reflected in the nature of the charges against can-

[17] We use affidavit information downloaded from the Liberty Institute, an Indian NGO. We initially downloaded copies of the affidavits from the website of the Election Commission of India. We then hired an Indian data input company to input the data recorded on the affidavits electronically for us. However, preliminary analysis revealed that the dataset produced by this company contained too many errors; as a result, we rebuilt our dataset using the information obtained on-line from the Liberty Institute.

[18] This matter was clarified in a Supreme Court judgement rendered on Dec. 6, 2006 specifically regarding corruption cases.

didates, a significant proportion of which would either be plausibly politically motivated or be charges to which politicians are naturally especially vulnerable given the nature of electoral competition. Such charges could involve crimes such as libel and slander (of other candidates, for instance) or activities for which there is little or no direct physical or eyewitness evidence. Charges such as murder, arson, looting, or assault are less likely to be fabricated and are not charges to which politicians are especially or uniquely vulnerable.

Such a minute investigation of the pattern of charges is beyond the scope of this analysis. However, a 2004 press release by the Association for Democratic Reforms (ADR) sheds light on the nature of the criminal charges against Indian parliamentarians elected to the Fourteenth Lok Sabha. Table 2, taken from the ADR report, details some of the charges against legislators. The ADR finds that there were 229 charges for serious violent crimes whereas 87 charges were related to "dishonesty, cheating, fraud, forgery and dealing in stolen property." Inspection of the data reported in the table reveals a relatively large number of criminal charges that are, by almost any definition, extremely serious and not likely to occur in the ordinary course of events during political campaigns; murder and arson, for instance. Although it is possible that some charges were politically motivated, the large proportion of serious violent crimes suggests that *CRIM* is mainly picking up a true criminal element in Indian politics. It is easy to fabricate a charge of libel against a political opponent but difficult to fabricate a charge of murder.

[Table 2 about here]

Ethnographic work substantiates the regular involvement of some Indian politicians in criminal activities and violence (Berenschot 2011). To illustrate the kinds of criminal activities of which prominent politicians are accused, we offer a few examples. Ganesh Singh was elected from the constituency of Satna in the state of Madhya Pradesh on the ticket of the Bharatiya Janata Party (BJP) in both the Fourteenth and Fifteenth Lok Sabha. At the time of each election, Singh stood indicted on multiple counts of cheating and forgery. He remained under indictment while serving as a Member of Parliament, although in January of 2007 the Calcutta newspaper, *The Telegraph*, reported than Singh was under prosecution for events dating back to

1998.[19].

Another example is one about which extensive information is publicly available, — due to its notoriety. Afajal Ansari is the son of a one-time President of the Indian National Congress (INC), India's dominant political party. Elected with 48 percent of the vote on the ticket of the Samajwadi Party (SP) to the 2004 Lok Sabha from the constituency of Ghazipur in the state of Uttar Pradesh (UP), the affidavit that Ansari filed with his candidacy papers reported three criminal charges, including rioting (India Penal Code Section 147), rioting and armed with a deadly weapon (India Penal Code Section 148), and criminal force to deter a public servant from the discharge of his duty (India Penal Code Section 353). We display a page of Ansari's affidavit reporting these charges in Figure 3.[20] (We direct the reader's attention to the statute numbers with an arrow.) In 2009, Ansari switched his party affiliation to the Bahujan Samaj Party (BSP) and again stood in the (redistricted) constituency of Ghazipur, where he again won the seat, this time with slightly more than 40 percent of the votes polled. Perhaps the decline in his vote share was related to the fact that Ansari filed his 2009 candidacy papers from jail, where he was being held in connection with the 2005 murder of Krishnanand Rai, a legislator from the BJP.[21]

These examples suggest that self-reported criminality is an imprecise but not invalid proxy for genuine criminality. There is little here to make us worry that India's acknowledged criminal politicians are simply more skilled at political contests and therefore more vulnerable to fabricated or politically generated charges.

[Figure 3 about here]

Finally, how the information about the alleged criminality of legislative candidates is used within India is also relevant to our evaluation of the validity of the self-reported charges. The Indian press and multiple

[19] See http://www.telegraphindia.com/1070120/asp/nation/story_7287189.asp.

[20] The sample is in Hindi, one of the eight possible languages in which affidavits are filed. It illustrates some of the difficulties in working with the original data used in this analysis. Many affidavits are handwritten, and therefore even more difficult to decipher than the sample page featured in Figure 3.

[21] In 2010, Ansari was expelled from the BSP, along with his brother, Mukhtar, who had also run for the Lok Sabha in 2009 under the BSP symbol and who at the time of his expulsion faced 30 criminal charges, including murder and kidnapping.

non-governmental organizations use the information from the affidavits to publicize the criminal element in Indian politics. An implication of this is that being charged with criminal activity is considered a potential electoral liability. Even if a candidate is unfairly charged, the general public is unlikely to know this. The candidate would have to use limited campaign resources convincing voters that the charges were false. For a party to deliberately select a self-reported criminal as its candidate thus carries with it fielding costs that are higher than those associated with other candidates. Given the added costs of running these candidates, the presence on the ballot of such a large number poses a puzzle. These various considerations give us confidence in the relative validity and reliability of the measure used to operationalize criminality.

4.2.2 Electoral Competitiveness

We operationalize electoral competitiveness by measuring how close each party expects to be to winning the seat in the constituency. In the multiparty setting in which we work, this generates a separate measure for each of the many political parties (and thus, for each partisan-affiliated candidate) in each constituency in each of the two legislative elections for which criminal records are disclosed by candidates. The reason that we construct separate measures of competitiveness for each candidate-party rather than a single constituency-level measure, as is common in single-member districts, lies with the large number of parties that operate in the typical Indian parliamentary constituency. The party of each candidate will gauge the competitiveness of the race in terms of how close it is to winning, not by how close the first runner up is likely to be. In the decision to field a self-reported criminal, we assume that each party's judgement depends on its own competitive position in the constituency.

Our main measure of competitiveness, *COMP*, is the (absolute value of the) percentage point difference between the share of the votes obtained by the winning candidate and the share of votes obtained by the party of the candidate in question in the constituency in the same election. This measure thus varies by party, by constituency, and by election.

Redistricting between elections, which prevents us from matching 2009 districts back to those that existed in 2004, requires that we use the actual constituency-level electoral results as proxies for the expected competitiveness of the race. One justification for this, in addition to the sheer necessity of doing so, is that we believe that candidates are selected by parties on the basis of relatively current and accurate information about how competitive the race is likely to be. Especially given the costs of fielding a known criminal, parties are likely to deploy various instruments to collect information about the sentiments of the electorate in the period immediately leading up to the election. This is especially likely to be true in settings with short parliamentary campaigns, as is the case in India, where campaigns last only a few weeks. Therefore, if the decision to list a self-reported criminal on the ballot is in part a function of electoral competitiveness, assessing the degree of competitiveness in the same election simply means that we believe that parties have a relatively good estimate of how close the election is likely to be in the weeks prior to the election when they must make final decisions about which candidates to list. An analysis of the 2004 election, not reported here, allowed us to estimate the expected competitiveness of the race using the results from the prior (1999) elections. A comparison of the results with results using the measure of competitiveness created with only 2004 data shows that it makes little difference which of the two possible measures is used.[22]

According to our first hypothesis, competitiveness should affect the probability that a party fields a self-reported criminal negatively; recall that smaller values of *COMP* indicate greater electoral competitiveness, which should in turn be associated with a greater likelihood of observing self-reported criminal candidates.

4.2.3 Vulnerability and Resilience to Intimidation and Violence

We proxy the vulnerability of the population to possible political intimidation and violence with a measure of literacy. *LIT* is an estimate of the literacy rate in each parliamentary constituency. As the literacy rate increases, we expect (on the basis of our second hypothesis) that the likelihood that parties list candidates

[22]Results available from the authors on request.

facing criminal indictments will decrease. We therefore expect to see more criminal candidates on the ballot where illiteracy is higher.

4.2.4 Political Control Variables

Even though India uses a simple plurality system (first-past-the-post), it nonetheless has more than two political parties running in each political constituency. We control for the number of parties (*NUMBER*) in each constituency. Where there are more candidates in the race, criminal candidates should find it easier to camouflage their status.

Our empirical specifications also include measures of three other features that we believe affect the nature of electoral competition in a constituency or for a candidate. These are whether the candidate is an incumbent (*INCUM*),[23] whether the seat is reserved for the representative of a scheduled caste or tribe (*RESERVED*), and whether the candidate is affiliated with a political party that is nationally organized or instead is local or regional in scope (*NATIONAL*).

INCUMB is coded 1 if a candidate is an incumbent and 0 otherwise. The standard argument is that incumbents enjoy electoral advantages because their names are more likely to be familiar to voters and because they have had the opportunity to use government resources to consolidate their electoral hold. However, research finds that in India legislators experience an incumbency disadvantage (Linden (2004); Uppal (2009)). This, however, is apparent only using a regression discontinuity design. Raw electoral returns show that incumbents do well in Indian elections. In 2004, 40 percent of those elected to the Lok Sabha were incumbents and in 2009 the equivalent figure rose to 53 percent. Put another way, in 2009, 51 percent of incumbents who ran again won their seat; in 2004, which featured many fewer candidates overall, 53 percent of incumbents who ran were reelected.

RESERVED is coded 1 if a seat is reserved for a member of a scheduled caste or tribe and 0 otherwise. A

[23] For candidates in 2004, we code a candidate as an incumbent if she represented the same parliamentary constituency as of the previous election. Due to redistricting, we cannot use this definition in 2009. Thus, for 2009 we define as an incumbent any candidate who won any seat in the previous election, regardless of where the seat is located in the country.

seat is designated reserved by each delimitation order setting electoral boundaries. The electoral boundaries used in the Lok Sabha elections of 2004 had been set by the Delimitation Order of 1976. In 2009, a new Delimitation Order was in effect that redrew electoral boundaries. Preliminary research shows that these boundaries reflect population changes and not political bias (Iyer and Shivakumar 2009). A seat is reserved on the basis of the share of population that is comprised of members of scheduled castes or tribes. There were 120 seats reserved in the Fourteenth Lok Sabha and 131 in the Fifteenth out of a total of 543 seats. Thus, nearly a quarter of parliamentary seats are now reserved.

Because caste features prominently in Indian politics, it seems reasonable to incorporate its potential impact on criminality in elections. We cannot measure caste directly since information on the distribution of various castes in the population is not included in the current Indian census. But seeing if the selection of known criminals as legislative candidates differs between reserved and non-reserved parliamentary seats is an indirect way to assess the effects of caste.

Reserved seats may differ from other constituencies in three important ways affecting their propensity to see known criminals given parliamentary tickets. First are the demographics of the population resident in the jurisdictions which are reserved. Reserved seats, which by definition have a higher scheduled caste and tribe population, also have more illiterate voters.[24] Even though we control for literacy in our estimations, reserved constituencies may have other unmeasured characteristics that make political parties there more likely to nominate criminals, justifying our decision to code them separately.

A second difference relates to barriers to entry in the political market. Although all voters are eligible to vote for reserved seats, only identified members of scheduled castes or tribes are eligible to stand as candidates. This may constitute a barrier to entry and make reserved seats less competitive.

A third possible difference between reserved and other seats regards aspects of political competition. Politics in reserved constituencies may be "ethnified." Banerjee and Pande (N.d.) argue that candidate qual-

[24]Literacy rates are 52 percent in reserved seats compared to 56 percent generally.

ity is a function of the proportion of a constituency's population that shares the same caste as the political party representing it. Their model predicts that candidate quality deteriorates as there is a closer match between constituency population characteristics and the representative's characteristics. Since reserved seats, by definition, have high shares of scheduled caste and tribe voters, the Banerjee-Pande model predicts that in constituencies with reserved seats, low caste parties will have lower quality candidates. Therefore, the model predicts that more known criminals, which we assume is a low-quality candidate characteristic, will be nominated for reserved seats.

On the basis of these three considerations, we expect that *RESERVED* will have a positive impact on the likelihood of criminals being nominated.

We also control for whether the party with which the candidate is affiliated qualifies as a "national" party (*NATIONAL*). National parties are defined as those parties that run candidates in a certain number of constituencies and across at least four states. The Election Commission of India classifies six parties as national in 2004: the BJP, the BSP, the Communist Party of India (*CPI*), the Communist Party of India (Marxist) (*CPM*), the Indian National Congress, and the Nationalist Congress Party (*NCP*). In 2009, national parties are defined as these same six parties plus the Samajwadi Party, Samata Party, and the Shiv Sena.

We include this variable because the decision-making calculus of national parties about recruiting self-reported criminals onto the ballot is likely to be different from that of more geographically restricted parties. National parties make decisions across parliamentary constituencies and are concerned with the extra-constituency ramifications of candidate selection. Parties that are organized only locally or in a state or two are more parochial in their decision making calculus. This may affect the likelihood of selecting known criminals to run. In particular, we expect that national parties may be less willing to do so, since they should be more sensitive to the costs of fielding criminal candidates.

4.2.5 Demographic Control Variables

In addition to these political control variables, our empirical models include two demographic control variables for which we can estimate data at the level of the parliamentary constituency. *POP* is the population of each constituency. Although India's electoral system is single member, parliamentary constituencies are not equally sized. We hypothesize that, because the technology of intimidation is inherently limited in scope, the ability of criminals to intimidate voters will decrease with size. Therefore, we expect that criminal candidates are less likely to be fielded in large constituencies.

URB is the percent of the population located in urban as opposed to rural areas in the constituency. We include this variable because we believe that urban voters, regardless of their level of literacy, have greater access to government protection from potential intimidation. In more urban areas, we therefore expect the likelihood of a self-reported criminal on the ballot to fall.

Summary statistics for the independent variables appear in Table 3, where we have also indicated the sign that we expect for each. Electoral results for 2004 and 2009 are presented in Table 4 to give readers an overall idea of the extent of Lok Sabha competition. Finally, in Appendix D (Table C-1), we present the data on alleged criminal candidates by state. States are in many contexts the natural units of analysis for Indian politics and knowledgeable observers will be especially interested in seeing the data presented at this level.

[Tables 3 and 4 about here]

5 Analysis of the Probability a Party Selects a Self-Reported Criminal Candidate

We employ a logistic analysis to test hypothesis 1 and 2 regarding the effects of political competitiveness and the literacy rate on the likelihood that a candidate facing criminal indictment appears on the ballot. The unit of analysis is the candidate-constituency. Let $q_{itps} = Prob(CRIM_{itps} = 1)$ be the probability that party p in election t fields a criminal candidate in district i in state s. The basic model that we estimate can then be written as:

$$
\begin{aligned}
logit(q_{itps}) \quad &= \quad \beta_s + \beta_1 COMP_{itps} + \beta_2 LIT_{its} \\
&\quad + \beta_3 CONTROLS1_{itps} + \beta_4 CONTROLS2_{its} + \epsilon_{its}
\end{aligned}
$$

where *COMP* is the measure of how close the candidate is to winning the seat; *LIT* is the proportion of the population in the constituency that is literate; *CONTROLS1* comprise control variables that vary by party (*INCUM* and *NATIONAL*), while *CONTROLS2* comprise the controls that are the same for all parties within a constituency (*RESERVED*, *NUMBER*, *POP*, and *URB*). We allow the intercept β_s to be state-specific in some specifications. Standard errors are clustered at the constituency level.

Because of changes in constituency boundaries that took place just prior to the 2009 elections, we are not able to match constituencies across the two elections we analyze. We therefore present results for each of the two elections separately, as well as for specifications in which we pool the data. Our strategy for pooling the data from the two elections is to attach to each candidate the characteristics relevant to the electoral boundaries in which he operates; *POP*, for instance, is the estimate of total population in the constituency for either 2004 or 2009, as appropriate. Because of these boundary changes, we are unable to include constituency fixed effects in the estimations.

The results of our first set of estimations appear in Table 5. For ease of interpretation, we report odds ratios.

[Table 5 about here]

Our baseline model appears as Model 1. It tests the likelihood a party fields a criminal candidate as a function of only *COMP* and *LIT*, our main theoretically relevant variables. Model 2 adds the control variables *INCUM*, *RESERVED*, *NATIONAL*, *NUMBER*, *POP*, and *URB*. Model 3, finally, adds state fixed effects to control for possible unobserved state-level heterogeneity.[25] For each model, we run three separate

[25]The coefficients for state effects are not included in the tables reporting results. About half the state-level coefficients are

estimations. The first is for candidates who ran in 2004, the second is for those who ran in 2009, and the third is a pooled analysis that includes both sets of observations. Interpretation of the odds ratios that we report follows standard guidelines. Results greater than 1 imply that the covariate increases the likelihood that a party fields a criminal candidate and results less than 1 imply the opposite.

In our baseline model, our principal independent variables, *COMP* and *LIT*, are each statistically significant at the 0.01 level. They retain significance even with the inclusion of statistical controls, although the inclusion of the state level fixed effects (Model 3) diminishes the statistical significance for *LIT* to the 0.10 level in 2004 and the variable loses statistical significance entirely for 2009 and in the pooled sample. The signs on each variable are in the expected direction: greater electoral competitiveness increases the likelihood of observing a self-reported criminal candidate (recall that smaller values of *COMP* imply greater competitiveness, since the difference in the share of the vote received compared to the share received by the winning candidate shrinks) and self-reported criminals are less likely to be fielded in parliamentary constituencies with larger fractions of literate voters.

Some control variables are statistically significant and deserve brief comment. The results reported in Table 5 show that parties are less likely to field known criminal candidates in constituencies with reserved seats and that national parties are more likely to do the same. Both results are contrary to expectations. Other controls, including incumbency status, the number of parties, the size, and the urbanness of the parliamentary constituency, are, typically, not statistically significant.

Given well-known difficulties in comparing logit coefficients across different models, we also report marginal effects for the baseline model (Model 1) and for the model that includes control variables (Model 2). We vary *COMP* and *LIT* one standard deviation below and above their mean values. Thus, *LIT* varies between 42 and 66 percent and *COMP* varies between 14 and 44 percent. All other variables are set to

statistically significant. When state dummies are included, states without variation on the dependent variable drop out of the analysis. This implies that if all the candidates in all the constituencies in a state report no criminal charges, the state drops. In 2004, 14 states or union territories drop out of the analysis reported in Model 3; in 2009, when more self-reported criminals are scattered across India, four states or union territories drop out. The dropped states and union territories are all quite small.

their mean values or, in the case of dummy variables, to zero. The results reported in Table 6 show that the substantive impact of our theoretically relevant independent variables is large. For example, the probability that a party fields a known criminal candidate is reduced by nearly 40 percent in the baseline 2004 model, 30 percent in 2009, and 34 percent in the pooled (2004 and 2009) analysis when the literacy rate shifts from a low of 42 percent to a high of 66 percent. The relationship between *LIT* and the probability of fielding a candidate under indictment is even stronger when control variables are added to the 2004 model (52 percent), although weaker in 2009 (17 percent). *COMP* also has a substantively important impact. For the baseline 2004 model, a party that is 14 percent away from winning the seat is 40 percent more likely to field a self-reported criminal candidate than a party that is 44 percent away from winning. The marginal effect increases to 50 percent in 2009.

[Table 6 about here]

6 Patterns in the Number of Self-Reported Criminal Candidates in Each Constituency

Our theory of how electoral competitiveness affects the decisions by political parties to select criminal candidates does not predict any particular pattern in the numbers of such candidates fielded across parliamentary constituencies. In the mixed strategy equilibrium of the model, we may observe constituencies with no self-reported criminal candidates, with one such candidate, or with two, and in districts with indefensible seats, we should observe only one criminal candidate being fielded. Alternate theories, such as Vaishnav (2010), predict instead that criminal candidates should appear in pairs; where one party selects a self-reported criminal, a close competitor should do so as well. To evaluate this claim empirically, in Table 7 we report the frequency with which we observe different numbers of candidates facing criminal indictment in India's parliamentary constituencies in the elections of 2004 and 2009.

[Table 7 about here]

The data reported in the table offer strong support for our theory. There is considerable dispersion in the

frequency with which we observe no self-reported criminal, one criminal, or two on the ballot in the same parliamentary constituency, just as our theory predicts. Indeed, the number of self-acknowledged criminal candidates in each constituency corroborates our theory in 94 percent of parliamentary constituencies in 2004 and in 86 percent of them in 2009.[26]

7 The Impact of Self-Reported Criminal Candidates on Incumbency Advantage

We now turn our attention to the impact acknowledged criminal candidates play on the electoral fates of incumbent parliamentarians. We examine whether incumbents who face self-reported criminal opponents are less likely to win reelection than incumbents who do not. That is, are indicted criminals eroding whatever incumbency advantage exists in India?

We employ a regression discontinuity design to estimate the incumbency (dis)advantage in the 2009 Lok Sabha elections.[27] Incumbency effects are estimated for all incumbents (baseline), and separately for incumbents who face a self-reported criminal opponent those who do not.

The regression discontinuity design is commonly used to assess the existence and extent of incumbency advantages in elections. Lee (2008) estimates the level of incumbency advantage for incumbent parties in the U.S. House of Representatives and finds that incumbent parties are 40 to 45 percent more likely to win an election. Using similar techniques, an incumbency disadvantage has been identified in India. Linden (2004) estimates that since 1991, incumbents in the Lok Sabha are 14 percent less likely to win than non-incumbents, and Uppal (2009) estimates that since 1991 incumbent members of state assemblies are 25

[26] We note as well that in 2004, 116 unaffiliated candidates report being under indictment, or less than five percent of all self-nominated candidates, compared with 12 percent of candidates who are selected to run by political parties. In 2009, 6.5 percent of independent candidates report being under indictment compared with 15 percent of their party-affiliated peers. Our theory predicts that self-acknowledged criminals should be selected by political parties as candidates; we have no theory of why such persons would self-nominate, since the latter candidates have almost no probability of winning a seat.

[27] Because the design requires data on the electoral results of all candidates from the previous election, we are able to estimate incumbency effects only for 2009 and not for 2004 or any earlier election.

percentage points less likely to win reelection than non-incumbents.

A regression discontinuity analysis compares the probability of winning an election across two groups: candidates who barely won the previous election and candidates who barely lost. The underlying logic is that whether a candidate barely wins or barely loses an election is essentially random and as such, candidates who are very close to winning or losing are otherwise identical. This approach controls for any observed or unobserved characteristics of incumbents that may bias the impact of incumbency on election rates when they are estimated using a regression-based empirical strategy.

There are three main steps involved in performing a regression discontinuity analysis. The first is to estimate the probability that a candidate wins an election conditional on his electoral performance in the prior election. We do this by using a logistic regression that estimates the probability that a candidate wins the 2009 election as a function of his 2004 margin of victory; this is done for incumbents and non-incumbents (i.e. candidates who ran in 2004 but lost) separately. The margin of victory of candidates who won in 2004 is calculated by subtracting the vote share for the second place candidate; i.e. we use the same formula as for COMP above. The margin for those who lost in 2004 is calculated by subtracting the vote share of the winning candidate from the share of each near-loser.[28]

Second, we restrict the sample space to ensure that the two groups of candidates (incumbents and non-incumbents) are as equivalent as possible with respect to observed characteristics. There is a trade-off between sample size and group comparability. If the margin of victory cutoff is set too narrowly, the likelihood of having comparable groups is high but at the cost of a small sample size. Conversely, if the margin of victory cutoff is set too widely, the likelihood of having comparable groups is lower, potentially biasing the results, but the sample size larger. Since, due to the absence of earlier years of data, our analysis only covers one election period, it is particularly sensitive to sample size problems. Though this introduces bias, we show later that the direction of the bias is towards overestimating an incumbency advantage.

[28]In contrast, Linden (2004) uses a non-parametric design and Uppal (2009) utilizes logistic regression with a fourth order polynomial for margin of victory as well as interactions and fixed effects.

The third step in the regression discontinuity analysis is to estimate the difference in election probabilities for incumbents and non-incumbents at the discontinuity (that is, where the electoral distance is zero). This is done by subtracting the probability of being elected in 2009 for non-incumbents from the probability of election for incumbents. It is especially effective to display this graphically by plotting the election probability in 2009 on the y-axis as a function of electoral distance in 2004, which is plotted on the x-axis, with the gap in election probabilities at the discontinuity representing the incumbency effect. If there is no incumbency effect, there is no gap between the two groups. If there is a negative effect, then the line representing the probability of being elected in 2009 is "lower" for incumbents than non-incumbents.

Let $p_{2009i} = Prob(GOTSEAT_{2009i} = 1)$ be the probability that candidate i wins a seat in 2009. Our baseline estimate for all incumbents is thus calculated by estimating the following equation:

$$logit(p_{2009i} = \alpha_0 + \alpha_1 * COMP_{2004i} + \epsilon_i$$

where *GOTSEAT* is a dummy variable that takes a value of 1 if the candidate won a seat in 2009 and 0 otherwise, *COMP* is the electoral distance a candidate was from winning or losing in 2004, and i are the parliamentary constituencies.

In order to estimate the impact on incumbency of facing a self-reported criminal opponent, the basic equation just outlined is expanded to include a dummy variable indicating whether a candidate faced a known criminal opponent. We also include a variable measuring the interaction of a self-reported criminal opponent and electoral marginality (COMP). Our theory of criminal intimidation implies that the electoral environment of constituencies with self-reported criminal candidates differs fundamentally from that of constituencies without them. For instance, we believe that candidates who report criminal indictments discourage opposition turnout, thereby altering electoral results. As a result, we hypothesize that constituencies with known criminals on the ballot may have different "slopes" with respect to the relationship between

electoral performance in 2004 and the probability of winning in 2009. We capture this expected difference with the interaction term. Thus we estimate:

$$logit(p_{2009i}) = \delta_0 + \delta_1 * COMP_{2004i} + \delta_2 * CRIMOPP_{2009i}$$
$$+\delta_3 * COMP_{2004i} * CRIMOPP_{2009i} i + \epsilon_i$$

where $CRIMOPP_{2009}$ is a dummy variable which takes the value of 1 if a candidate faces an opponent with a self-reported criminal record and 0 otherwise; and $COMP * CRIMOPP$ is the interaction term between $COMP_{2004}$ and $CRIMOPP_{2009}$.

Predicted probabilities were simulated using Clarify (Tomz, Wittenberg and King 2003) and the charts were prepared using Zelig (Imai, King and Lau 2009).

7.1 Data

Of the approximately $8,000$ candidates who ran for the Lok Sabha in 2009, 572 were on the ballot in 2004. Of these 572, we restrict attention to candidates within 25 percent of winning or losing in 2004, thereby reducing our candidate pool to 398.[29] Bias-free results require that both incumbents and non-incumbents be balanced on observed as well as unobserved traits. Since in order to obtain a large enough sample size, we set a wide margin of victory (25 percent) as our criterion of inclusion, our groups are not comparable along certain important dimensions. Table 8 compares incumbents and non-incumbents along a variety of dimensions: the probability of winning in 2009; vote shares received in 2004 and 2009; proportions who report facing criminal indictment; proportions with an opponent facing criminal indictment in 2009; and

[29] We also ran an analysis with a cutoff of 10 percent. In that analysis, the incumbency disadvantage for all incumbents disappears but our main result remains; incumbents who face a known criminal opponent are at a disadvantage whereas incumbents who face other opponents are not.

proportions affiliated with the Congress or the BJP.

[Table 8 about here]

As the data presented in the table document, incumbents on average received higher vote shares in 2004 than non-incumbents. This implies that incumbents as a whole may have better political skills than their non-incumbent rivals. The main implication for the analysis that follows is that results may be skewed in favor of incumbents. Thus, results that show an incumbency advantage would need to be viewed with particular caution whereas results that show an incumbency disadvantage may understate the true level of this disadvantage. The bias therefore works against the argument that Indian incumbents are disadvantaged electorally.

7.2 Results

Table 9 summarizes the results of the regression discontinuity analyses. Results are graphically presented in Figures 4 and 5. Results of the baseline analysis are consistent with those reported in previous studies. We find that legislative incumbents face an electoral disadvantage in India. They confront a 6 percentage point disadvantage in winning office in 2009, which translates into being 15 percent less likely to win.[30]

[Table 9 and Figures 4 and 5 about here]

Figures 4 and 5 present the main results of the regression discontinuity analysis. Figure 4 shows that in 2009, incumbents and near-losers from 2004 enjoyed similar probabilities of election. There is a noticeable but small incumbency disadvantage, represented by the break in the line at the discontinuity and the slight drop in the line to the right of 0, representing near winners in 2004. The source of this incumbency disadvantage is clarified in Figure 5. The left-hand panel shows that candidates facing a self-reported criminal opponent have an incumbency disadvantage (8 percentage points) whereas those facing an opponent with no

[30] Our results do not exactly match those reported by Linden (2004) or Uppal (2009), but our analysis differs in three important respects. First, unlike Linden (2004), we use a parametric approach and estimate a logit function. Second, unlike Uppal (2009), we do not include control variables in our estimations. Finally, by using a wider margin of victory cutoff than either study, or than Clots-Figueras (2010), who also performs a regression discontinuity analysis using Indian data, our results understate the incumbency disadvantage.

known criminal record enjoy an incumbency advantage (3 percentage points). The two figures clearly show that the entire incumbency disadvantage is systematically linked to whether an incumbent faces off against a known criminal or not.

These findings document that the well known disadvantage of the Indian legislative incumbent is a product of whether he faces a known criminal among his opponents. That two-thirds of India's electoral districts saw at least one acknowledged party-sponsored criminal on the ballot in 2009, representing a massive diffusion of candidate criminality over 2004, suggests that self-reported criminals are driving out non-criminals at the national level of the Indian polity.

8 The Impact of Self-Reported Criminal Candidates on Election Turnout

The evidence we have presented above is largely consistent with the predictions of our theoretical model. We now present evidence that the underlying mechanism that produces these results is, as we claim, that criminals employ technologies of violence to intimidate voters into not voting. We test this basic claim by studying whether self-reported criminal candidates depress electoral turnout.

If our theory of political criminality in India is correct, acknowledged criminal are selected by political parties as legislative candidates because of their skills in utilizing technologies of violence to intimidate voters. If this is the case, then we should observe lower electoral turnout where such candidates appear on the ballot. In 2004, turnout varied across Indian parliamentary constituencies from a low of 15 percent of the electorate to a high of 91 percent; in 2009, the variation was slightly less. To assess if some of this variation is driven by the presence of self-reported criminals on the ballot, we estimate a series of constituency-level linear regressions of the following sort for each election year:

$$TURNOUT_{ct} = \lambda_0 + \lambda_1 * CRIM_{ct} + \lambda_2 * COMP_{ct} + \lambda_3 * LIT_{ct} + \lambda_4 * CONTROLS_{ct} + \epsilon_{ct}.$$

where c indexes the constituency; t indexes the election; and ϵ is an error term. Because the unit of analysis is the parliamentary constituency, rather than single parties or individual candidates in a constituency, all the variables are measured at the constituency level. *LIT* and most control variables, including *ENP*, *RESERVED*, *POP*, and *UBR*, are measured at the level of the constituency in the work reported above or convert to that level in a natural way. (Note that whereas we used the number of political parties running candidates in each constituency for our candidate-level models, we used the effective number of parties, or *ENP*, for the constituency-level models, as is commonplace. (*INCUM* is now defined as whether any incumbent candidate runs in the constituency and *CRIM* is now defined as whether any candidate in a constituency had a criminal record. The main difficulty is to operationalize electoral competitiveness at the constituency level. We measure *COMP* as the difference in the share of the vote received by the winner and the first runner-up, but this is clearly a rough gauge that does not capture the strategic incentives of individual political parties. However, we have no way to incorporate such information into constituency-level analyses. We utilize an OLS estimator and report results in Table 10.[31]

[Table 10 about here]

As the results reported in Table 10 document, turnout is significantly reduced by the presence of at least one self-reported criminal candidate on the ballot in all but two of the models studied. In a simple t-test of the 2004 data, turnout is about 3 percent higher in constituencies with no acknowledged criminal on the ballot than in those with a single such individual ($p = .0048$) and in 2009, the difference grows to 5 percent ($p = .0002$). The other variables in the models largely perform as reported in prior (state-level) studies of turnout in India (Diwakar 2008) as well as in cross-national studies of turnout, including those of Jackman (1987) and Blais (2006). Our results show that turnout is significantly higher when literacy

[31]The turnout measure is censored from below (no observations are less than zero) and from above (no observations are greater than 100). We performed both a tobit analysis and analyses using a fractional response estimator in addition to the standard OLS that we report below; the results were consistent regardless of which type of estimator was used.

increases. It is also lower as the size of the population rises and with more urbanness. Similar results have been observed before. As studies have consistently shown for other countries (Jackman 1987), we also find that the effective number of parties, *ENP*, is significantly and negatively associated with turnout: as the number of parties running candidates rises, turnout falls.

The only variable that does not behave as expected by our theory or as reported in prior studies (Diwakar 2008) is our proxy for electoral competitiveness. Results for *COMP*, which measures the margin of victory of the winner over the runner-up, are inconsistently signed and in half of the regressions do not meet conventionally defined significance thresholds. These inconsistent results may be due to the noise with which electoral competitiveness is measured at the constituency level in a multi-party setting; with so many parties contesting elections in India, the strategic use of electoral violence will vary by party even within the same constituency and this measure is not sensitive to that. Despite the unexpected results on the noisy measures of competitiveness that we are forced to use in these constituency-level regression models, the chief variable of theoretical interest preforms largely as expected: *CRIM* reduces turnout and the results are generally significant.

9 Conclusions

Self-reported criminal candidates are more likely to appear on the ballot in India in constituencies with more illiterate voters and when a party faces greater electoral competition. Although many parliamentary constituencies observe no known criminals running for the national parliament, when such individuals appear on the ballot, they do so as singletons or, less often, in pairs, reflecting the mixed strategy equilibrium that we theorize underlies their selection as candidates. These candidates reduce voter turnout. We interpret this as corroboration of our theory that acknowledged criminals utilize violent tactics to intimidate opposition supporters and keep them at home on election day. Finally, self-reported criminals appear to be successfully driving Indian incumbent legislators out of office, which may explain their proliferation in 2009.

There is much that we still do not know about the use of violence as an electoral strategy. We have not studied which Indian political parties are more likely to select known criminals and whether established or insurgent party organizations are more likely to do so. We have yet to explore the trade-offs that political parties make between vote buying among possible supporters and intimidation, of opponents, or the relative costs of each. We have not investigated the behavior of criminal individuals on the ground; how they use violence, when in the electoral cycle they do so, how often violence is used, the types employed, and the characteristics of victims. We have not examined how government agents respond, and when they try to limit pre-electoral violence or how effective they are in doing so. We have not looked at how voters react or the impact of the growing use of violent electoral tactics on public opinion and partisanship. Finally, we have not investigated how individuals with criminal backgrounds perform when elected to parliament, and whether they remain a distinct set of legislators whose performance differs from their unindicted peers. Many avenues of future research remain.

References

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

Banerjee, Abhijit V., Selvan Kumar, Rohini Pande and Felix Su. 2010. "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India." Unpublished paper.

Berenschot, Ward. 2011. "On the Usefulness of Goonda's in Indian Politics: 'Moneypower' and 'Musclepower' in a Gujarati Locality." *South Asia: Journal of South Asian Studies* 34(2):255–275.

Besley, Timothy and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *Quarterly Journal of Economics* 117(4):1415–1451.

Bhavnani, Rikhil R. 2011. "Corruption Among India's Politicians: Evidence from Unusual Data." Unpublished paper.

Blais, André. 2006. "What Affects Voter Turnout?" *Annual Review of Political Science* 9:111–25.

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

Bratton, Michael. 2008. "Vote Buying and Violence in Nigerian Election Campaigns." *Electoral Studies* 27(4):621–32.

Brollo, Fernada. 2009. "Who is Punishing Corrupt Politicians — Voters or the Central Government? Evidence from the Brazilian Anti-Corruption Program." Unpublished paper.

Chakrabarti, Rajesh, Shubhashis Gangopadhyay and Shagun Krishnan. N.d. "Incumbency Effects in Indian National Elections — A Preliminary Exploration." Unpublished paper.

Chang, Eric C.C., Miriam A. Golden and Seth J. Hill. 2010. "Legislative Malfeasance and Political Accountability." *World Politics* 62(2):177–220.

Chemin, Matthieu. 2008. Do Criminal Politicians Reduce Corruption? Evidence from India. Cahier de recherche/Working Paper 08-25 Centre Interuniversitaire sur le Risque, les Politiques Economiques et l'Emploi Montreal.: .

Clots-Figueras, Irma. 2010. "Women in Politics. Evidence from the Indian States." *Journal of Public Economics* 95(7-8):664–90.

Diwakar, Rekha. 2008. "Voter Turnout in the Indian States: An Empirical Analysis." *Journal of Elections, Public Opinion & Parties* 18(1):75–100.

Dixit, Avinash and John Londregan. 1996. "The Determinants of Success of Special Interests in Redistributive Politics." *Journal of Politics* 58(4):1132–55.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2):703–45.

French, Patrick. 2011. *India: A Portrait*. New York: Knopf.

Hillman, Ayre L. 2010. "Expressive Behavior in Economics and Politics." *European Journal of Political Economy* 26(4):403–18.

Imai, Kosuke, Gary King and Olivia Lau. 2009. "Zelig: Everyone's Statistical Software." *R package version* pp. 3–4.

Iyer, Lakshmi and Maya Shivakumar. 2009. "Electoral Redistricting in the World's Largest Democracy: An Evaluation." Unpublished paper.

Jackman, Robert W. 1987. "Political Institutions and Voter Turnout in the Industrial Democracies." *American Political Science Review* 81(2):405–24.

Jennings, Colin and Alan Hamlin. 2011. "Expressive Political Behaviour: Foundations, Scope and Implications." *British Journal of Political Science* 41(3):645–670.

Keefer, Philip and Stuti Khemeni. 2009. "When Do Legislators Pass on Pork? The Role of Political Parties in Determining Legislator Effort." *American Political Science Review* 103(1):99–112.

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics* 142(2):675–97.

Linden, Leigh L. 2004. "Are Incumbents Really Advantaged? The Preference for Non-Incumbents in Indian National Elections." Unpublished paper.

Nichter, Simeon. 2008. "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Review* 21026(1):19–31.

Nooruddin, Irfan and Pradeep Chhibber. 2007. "Unstable Politics: Fiscal Space and Electoral Volatility in the Indian States." *Comparative Political Studies* 41(8):1069–91.

Nyblade, Benjamin and Steven R. Reed. 2008. "Who Cheats? Who Loots? Political Competition and Corruption in Japan, 1947–1993." *American Journal of Political Science* 54(4):926–41.

Persson, Torsten and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. MIT Press.

Reinikka, Ritva and Jakob Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3(2-3):259–267.

Sen, Amartya. 1981. *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford: Oxford University Press.

Stokes, Susan C. 2005. "Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina." *American Political Science Review* 99(3):315–25.

Tiwari, Devesh. In Progress. Criminal Candidates and Party Politics in India. PhD thesis University of California at San Diego. La Jolla, CA: .

Tomz, Michael, Jason Wittenberg and Gary King. 2003. "CLARIFY: Software for Interpreting and Presenting Statistical Results." *Journal of Statistical Software* 8(1):1–30.

Uppal, Yogesh. 2009. "The Disadvantaged Incumbents: Estimating Incumbency Effects in Indian State Legislatures." *Public Choice* 138(1):9–27.

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

Wilkinson, Steven I. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. New York: Cambridge University Press.

Wilkinson, Steven I. 2007. Explaining Changing Patterns of Party-Voter Linkages in India. In *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political Competition*, ed. Herbert Kitschelt and Steve I. Wilkinson. Cambridge: Cambridge.

Appendix A: Data and Data Sources

Data units are of four types:

1. National elections occur in India's 543 parliamentary constituencies.
2. Demographic information from the 2001 Indian Census is collected in 593 administrative districts. The administrative districts do not coincide neatly with the parliamentary constituencies.
3. India's 35 states and territories hold elections across 4,140 assembly constituencies.
4. The assembly constituencies aggregate perfectly to the parliamentary constituencies.

To use demographic information at the level of the parliamentary constituencies, we estimate values based on parliamentary voter-weighted values aggregated from administrative districts. For additional details, see below, under *population*.

Caste information is provided by the 2001 Indian census. The information available refers to numbers of persons who are members of scheduled castes and tribes. Data are not available for Manipur and Nagaland. Caste information downloaded from `http:/www.indiastat.com` and `http:/censusindia.gov.in` in April 2008.

Indian Census 2001. Many demographic variables below are taken from the 2001 Indian census, available as CensusInfo [electronic resource], India 2001 (Office of the Registrar General, New Dehli, India), Version 1.0.

Criminal charges are based on sworn affidavits provided by all parliamentary candidates to the Lok Sabha in the 2004 and 2009 elections. The affidavits contain the candidate name and party affiliation organized by state and parliamentary constituency. Downloaded from the Liberty Institute, `http:/www.empoweringindia.org/new/home.aspx`.

Electoral returns from 2004 and 2009 taken from the Electoral Commission of India. Information includes the number of votes received by each candidate where candidates are organized by state, parliamentary constituency, and party. We matched the returns for each candidate with the Liberty Institute's coding on whether the candidate's affidavit reported pending criminal charges. Election results downloaded from `http:/eci.nic.in/StatisticalReports/ElectionStatistics.asp`.

Effective number of parties are calculated at the district level for 2004 and 2009. The formula we use is: $ENP_{it} = \frac{1}{\sum_{it} s_i^2}$, where s_i is the vote share for party i.

Incumbents are coded 1 and non-incumbent candidates are coded 0. Incumbency status as of 2004 coded by matching the names of the winners in the 1999 Lok Sabha elections and winners in any subsequent bye-elections with the names of candidates in 2004 in the same political constituency. For 2009, any candidate that won in 2004 was coded as an incumbent. Prior election results from the website of the Election Commission of India.

Independent vote shares refers to the share of the vote in each constituency won by candidates with no partisan affiliation. Data from the Election Commission of India.

Literacy rates are calculated using 2001 Indian Census (see above); we divide the number of illiterate persons by the total population.

National parties are designated by the Election Commission of India. A national party must be a state party in four or more states. In 2004, the six national parties were the BJP, the BSP, the CPI, and CMP, INC, and NCP. In 2009, the Nationalist Congress Party lost its designation as a national party and the Samajwadi Party, Samata Party, and the Shiv Sena were all designated national parties in addition to the other five parties that had enjoyed national status in 2004.

Reserved seats are coded 1 and regular parliamentary seats coded 0. In the Fourteenth Lok Sabha, there are

79 seats reserved for scheduled castes and 41 for scheduled tribes. In the Fifteenth Lok Sabha, 84 seats are reserved for scheduled castes and 47 for scheduled tribes. Seats are designated for reservation by delimitation orders based on the population of scheduled caste and tribes. The 2004 elections were conducted under the Delimitation Order of 1976 and the 2009 elections under the Delimitation Commission of 2002; the latter was based on 2001 Indian census data.

Population in each parliamentary constituency calculated using data from the 2001 Indian census. Population data is available at the level of administrative districts. We estimate population for parliamentary constituencies in two steps. First we estimate population totals for the state assembly constituencies on the basis of the fraction of votes cast in each state assembly constituency out of the total number of votes cast in the corresponding administrative district. (Parliamentary vote totals are available from the Election Commission of India at the level of administrative districts.) We then aggregate the estimated population totals from assembly constituencies to parliamentary constituencies. Similar manipulations are performed for all variables drawn from the 2001 Indian census.

Urban is the percent of the population that lives in urban areas. Calculated by taking the urban population in each administrative district and dividing by the district's total population. Data available from the 2001 Indian census.

Votes cast (total) in the 2004 and 2009 Lok Sabha elections for each parliamentary constituency from the Election Commission of India.

Vote shares for each candidate in 2004 and 2009 are calculated by dividing the number of votes each candidate receives by the vote totals received by all candidates in each parliamentary constituency and multiplying by 100. We calculate the total votes cast in each parliamentary constituency by adding up the votes for all candidates listed by the Election Commission of India in each parliamentary constituency.

Procedure used for matching 2004 and 2009 parliamentary constituencies with 2001 census districts:

For 2004, the Delimitation Order of 1976 was used to match administrative districts to state assembly districts. Second, each assembly district's population figures from the census is estimated as a fraction of the population in the corresponding administrative district. Each assembly district's fraction was determined by its fraction of votes cast in the 2004 election of the entire administrative district. Third, since state assembly districts are perfect subsets of parliamentary constituencies, population estimates are aggregated up to the parliamentary constituency level.

Some administrative districts were partitioned after 1976. The following parliamentary constituencies were reaggregated to their 1976 boundaries: in the state of Karnataka, Bagalkot was reaggregated to Bijapur, Koppal was reaggregated to Raichur, Gadag and Haveri were reaggregated to Dharwad, Davangere was reaggregated to Chitradurga, Bangalore Rural was reaggregated to Bandalore, Udupi was reaggregated to Dakshina Kannada, and Chamarajnagar was reaggregated to Mysore. In the state of Rajasthan, Bharapur was reaggregated to Dhaulpur, Hanamungarh was reaggregared to Ganganagar, Dausa was reaggregated to Jaipur, Baran was reaggregated to Kota, Kkarauli was reaggregated to Sawai Madhopur, and Rajsamand was reaggregated to Udaipur. For the state of Tamil Nadu, Ariyalur was reaggregared to Perambular. In the union territories, Andaman and Nicobar were combined as were Daman and Diu.

For the 2009 elections, we replicate the above procedures using the new Delimitation Order. However, we were unable to match census data to electoral data for 25 constituencies. The delimitation order did not have information for 12 constituencies in the state of Assam. In addition, several new administrative districts were created after 2004. In Bihar, the administrative district of Arwal was split from Jehenabad after 2001. In Karnataka, Anuppur, Ashok Nagar and Burhanpur are new districts created after 2001. In Manipur, all nine administrative districts either underwent name changes or had new boundaries drawn. In Tamil Nadu, Krishnagiri was split from Dharmapuri and Ariyalur was

split from Perambalur. Finally, in West Bengal, the administrative districts of Paschim Medinipur and Purbo Medinipur were combined into Medinipur. In addition, the 20xx Delimitation Order did not contain the electoral boundaries for any of the seven union territories. As a result, we were unable to merge census data for 25 electoral constituencies for 2009, all of which were dropped from analyses employing demographic data.

Appendix B: Specification of the Formal Model

This appendix provides a full characterization of the equilibrium of the candidate fielding game. The seat is safe if

$$\Delta - \Delta_B > 0 \;\Leftrightarrow \gamma < \widehat{\gamma} - \frac{\theta_B t}{\theta_A(\theta_B - c)} \equiv \underline{\gamma} \tag{B-1}$$

and competitive otherwise. When $\gamma < \underline{\gamma}$, party B cannot swing the election and it is a dominant strategy for that party not field a criminal. When $\gamma \geq \underline{\gamma}$, the seat is competitive and the regime depends on whether party A can defend the seat or not. The seat can be defended if

$$\Delta - \Delta_B + \gamma\Delta_A > 0 \Leftrightarrow \gamma < \frac{\theta_B(\theta_A - c - t)}{\theta_A(\theta_B - c - t)} \equiv \overline{\gamma} \tag{B-2}$$

and is indefensible otherwise.

The normal form representation of the candidate fielding game is

A/B	Criminal	Non-criminal
Criminal	$M - \mu$ $-\mu$	0 $M - \mu$
Non-criminal	$M - \mu$ 0	M 0

where the column player is party A and the row player is party B. We can rule out each of the four potential

pure strategy Nash equilibrium by showing that at least one player has a profitable deviation. First, suppose that both parties field a criminal. Given that party A fields a criminal, party B is better off not doing so, as $-\mu < 0$. Second, suppose that both parties field a non-criminal. Given that party A fields a non-criminal, party B will deviate and field a criminal because $M-\mu > 0$. Third, suppose that party A fields a criminal but party B does not. Given that party B does not field a criminal, party A will deviate and field a non-criminal because $M > M-\mu$. Fourth, suppose that party B fields a criminal but party A does not. Given that party B fields a criminal, party A will deviate and also field a criminal because $M-\mu > 0$. To find the mixed strategy equilibrium, let λ_k be the probability that party k fields a criminal. If a party is using a mixed strategy at equilibrium, then it should have the same expected payoff from the pure strategies over which it is mixing. For party A, this means that the expected payoff of fielding a criminal, $\lambda_B(M-\mu) + (1-\lambda_B)(M-\mu)$, must be equal to the expected payoff of fielding a non-criminal, $\lambda_B 0 + (1-\lambda_B) M$. This implies that

$$\lambda_B = \frac{\mu}{M}. \tag{B-3}$$

For party B, the expected payoff of fielding a criminal, $\lambda_A(-\mu) + (1-\lambda_A)(M-\mu)$, must equal the expected payoff of fielding a non-criminal, $\lambda_A 0 + (1-\lambda_A) 0$. This implies that

$$\lambda_A = \frac{M-\mu}{M}. \tag{B-4}$$

The seat cannot be defended if

$$\Delta - \Delta_B - \gamma\Delta_A < 0 \Leftrightarrow \gamma > \overline{\gamma}. \tag{B-5}$$

To make sure that $\overline{\gamma} < \widehat{\gamma}$, we require that $\theta_B > \theta_A$. If this fails, then the seat is always defensible. The normal form of the candidate fielding game is as follows:

A/B	Criminal	Non-criminal
Criminal	$-\mu$ $M-\mu$	0 $M-\mu$
Non-criminal	$M-\mu$ 0	M 0

We observe that fielding a criminal is a dominant strategy for party B. Given this, the best response for party A is not to field a criminal and the unique pure strategy Nash equilibrium is that party B fields a criminal and party A does not. Since dominated strategies are never used in mixed Nash equilibria, we can rule out mixed strategy equilibrium in this regime.

Appendix C: State-Level Information on Candidates

Table C-1: Criminal Candidates and Electoral Outcomes by State, Fourteenth and Fifteenth Lok Sabha

	Percent Listed 2004	**Percent Elected 2004**	**Percent Listed 2009**	**Percent Elected 2009**	**Number of seats**
Andaman Nicoba	0	0	45	100	1
Andhra Pradesh	5	14	7	21	42
Arunachal Prades	0	0	12	0	2
Assam	0	0	8	14	14
Bihar	18	38	22	45	40
Chandigarh	0	0	0	0	1
Chattisgarh	2	18	4	18	11
Dadra Nagar Ha	20	0	60	100	1
Daman Diu	33	100	29	0	1
Goa	25	50	6	0	2
Gujarat	14	27	15	27	26
Haryana	1	10	10	10	10
Himachal Pradesh	0	0	0	0	4
Jammu Kashmir	0	0	9	17	6
Jharkhand	21	50	25	50	14
Karnataka	3	21	7	21	28
Kerala	16	35	13	35	20
Lakshadweep	0	0	25	0	1
Madhya Pradesh	2	21	9	17	29
Maharashtra	5	42	9	25	48
Manipur	0	0	0	0	2
Meghalaya	0	0	9	0	2
Mizoram	0	0	0	0	1
NCT of Delhi	7	29	8	0	7
Nagaland	0	0	0	0	1
Orissa	3	14	19	24	21
Puducherry	14	0	14	0	1
Punjab	8	38	8	15	13
Rajasthan	10	12	5	4	25
Sikkim	0	0	14	0	1
Tamil Nadu	7	21	6	23	39
Tripura	0	0	11	0	2
Uttar Pradesh	11	30	14	34	80
Uttarakhand	0	0	9	20	5
West Bengal	9	10	11	12	42
Total	9	24	11	24	543

Notes: Criminal candidates defined as those who report having been convicted of or currently facing criminal indictment in their affidavit filed with candidacy papers.

Tables and Figures

Table 1: Criminal Candidates and Electoral Outcomes, Fourteenth and Fifteenth Lok Sabha

	Number of Candidates	Number Elected	Percent Elected
All candidates, 2004			
Not criminals	4,960	415	8.37
Criminals	475	128	26.95
Total	5,435	543	9.99
All candidates, 2009			
Not criminals	7,177	414	5.77
Criminals	893	129	14.45
Total	8,070	543	6.73
Party-affiliated candidates, 2004			
Not criminals	2,691	410	15.24
Criminals	359	128	35.65
Total	3,050	538	17.64
Party-affiliated candidates, 2009			
Not criminals	3,596	406	11.29
Criminals	643	128	19.91
Total	4,239	534	12.60
Independent candidates, 2004			
Not criminals	2,269	5	<1
Criminals	116	0	0
Total	2,385	5	<1
Independent candidates, 2009			
Not criminals	3,581	8	<1
Criminals	250	1	<1
Total	3,831	9	<1

Notes: Criminal candidates defined as those who report having been convicted of or currently facing criminal charges in their affidavit filed with candidacy papers.

Table 2: Serious Criminal Charges Against MPs by Party, Fourteenth Lok Sabha

Party	**BJP**	**INC**	**CPM**	**CPI**	**BSP**	**NCP**	**Other**	**Total**
Murder, attempted murder, etc.	7	4	2	1	17	0	56	84
Robbery	0	4	0	0	8	0	5	17
Kidnapping	0	1	0	0	2	0	9	11
Theft and extortion	1	0	0	0	3	0	24	28
Rape	0	0	0	0	0	0	0	1
Other violent crimes	9	7	3	0	13	2	54	88
Total violent crimes	17	16	5	1	43	2	149	229
Dishonesty, cheating, fraud, forgery, dealing in stolen property	5	17	6	0	23	0	36	87
False oaths	5	4	0	0	0	0	7	16
Defiling place of worship	1	0	0	0	0	0	0	1
Total other serious crimes	11	21	6	0	23	0	43	104
Total all crimes	28	37	11	1	66	2	192	333

Notes: Adapted from "Lok Sabha Elections: Press Release July 21, 2008," issued by the Association for Democratic Reforms (ADR) and partner NGOs from All India Election Watch Network; downloaded from `www.adrindia.org/downloads/LokSabha_High_Level_Analysis.doc`. Data refer to the number of crimes committed, not number of MPs charged. BJP: Bharatiya Janata Party. INC: Indian National Congress. CPM: Communist Party of India (Marxist). CPI: Communist Party of India. BSP: Bahajan Samaj Party. NCP: Nationalist Congress Party.

Table 3: Summary Statistics and Expected Signs of Independent Variables on the Selection of Self-Reported Criminals as Candidates

	Mean	SD	Min.	Max.	Exp. Sign
Electoral year 2004					
COMP	31.23	18.28	.06	72.58	-
LIT	54.10	11.97	25.86	85.42	-
INCUM	.13	.34	0	1	-
RESERVED	.21	.41	0	1	+
NATIONAL	.44	.50	0	1	-
NUMBER	5.52	2.52	0	15	+
POP	1,957,936	538,101.6	60,595	5,410,783	-
URB	27.57	20.86	3.48	100	-
Electoral year 2009					
COMP	31.93	16.72	.04	78.24	-
LIT	53.99	12.00	25.42	85.29	-
INCUM	.07	.25	0	1	-
RESERVED	.33	.47	0	1	+
NATIONAL	.38	.49	0	1	-
NUMBER	7.84	3.04	1	20	+
POP	1,930,203	358,941.7	186,189	4,013,609	-
URB	27.89	20.72	3.48	100	-

Notes: COMP is the absolute value of the difference between the share of vote won by the candidate and the winner's share; LIT is the percent of population that is literate; INCUM is whether the candidate is an incumbent; RESERVED is whether seat is reserved for a scheduled caste or tribe representative; NATIONAL is affiliation with national party; NUMBER is number of other partisan-affiliated candidates; POP is total population; URB is percent of total population in urban areas. LIT, RESERVED, NUMBER, POP, and URB are measured at the level of the parliamentary constituency. COMP, INCUM, and NATIONAL are measured at the level of the individual candidate.

Table 4: Fourteenth and Fifteenth Lok Sabha Electoral Results by Party

Party	**No. of Candidates**	**No. of Seats Won**	**Percent of Cands. Winning**	**Percent of Cands. Winning**
2004 Electoral Results				
BJP	364	138	25.4	37.9
BSP	435	19	3.5	4.4
CPI	34	10	1.8	29.4
CPM	69	43	7.9	62.3
INC	417	145	26.7	34.8
NCP	32	9	1.7	28.1
Other	1,699	174	32.0	10.2
Total	3,050	538	99.1	17.6
2009 Electoral Results				
BJP	433	116	21.36	26.79
BSP	500	21	3.87	4.20
CPI	56	4	0.74	7.14
CPM	82	16	2.95	19.51
INC	440	206	37.94	46.82
NCP	68	9	1.66	13.24
RJD	44	4	0.74	9.09
AITC	27	19	3.50	70.37
DMK	22	18	3.31	81.82
JD(U)	27	20	3.68	74.07
SHS	22	11	2.03	50.00
SP	95	23	4.24	24.21
Other	2,423	67	12.34	2.77
Total	4,239	534	98.34	12.6

Notes: BJP: Bharatiya Janata Party. BSP: Bahujan Samaj Party. CPI: Communist Party of India. CPM: Communist Party of India (Marxist). INC: Indian National Congress. NCP: National Congress Party. RJD: Rashtriya Janata Dal. AITC: All India Trinamool Congress. DMK: Dravida Munnetra Kazhagam. JD(U): Janata Dal (United). SHS: Shivsena. SP: Samajwadi Party. Figures excludes independent candidates.

Table 5: Logit Estimation Results of the Determinants of Self-Reported Criminal Candidates in Parliamentary Constituencies, Fourteenth and Fifteenth Lok Sabha Elections (Odds Ratios)

		Model 1			Model 2			Model 3	
Election	2004	2009	Pooled	2004	2009	Pooled	2004	2009	Pooled
	Model 1a	**Model 1b**	**Model 1c**	**Model 2a**	**Model 2b**	**Model 2c**	**Model 3a**	**Model 3b**	**Model 3c**
COMP	0.981***	0.973***	0.977***	0.982***	0.976***	0.979***	0.978***	0.977***	0.979***
	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)
LIT	0.977***	0.983***	0.981***	0.967***	0.986**	0.979***	0.981	0.997	0.990
	(0.006)	(0.005)	(0.004)	(0.009)	(0.007)	(0.005)	(0.012)	(0.008)	(0.007)
INCUM				1.295	1.194	1.176	1.256	1.265	1.207
				(0.208)	(0.194)	(0.132)	(0.219)	(0.212)	(0.141)
RESERVED				0.641**	0.594***	0.652***	0.625***	0.636***	0.680***
				(0.114)	(0.063)	(0.058)	(0.110)	(0.068)	(0.062)
NATIONAL				0.998	1.356***	1.221**	1.048	1.364***	1.262***
				(0.125)	(0.138)	(0.097)	(0.139)	(0.145)	(0.103)
NUMBER				1.022	1.011	1.041***	1.000	0.970	1.018
				(0.036)	(0.021)	(0.016)	(0.053)	(0.023)	(0.017)
POP				1.000	1.000	1.000	1.000	1.000	1.000
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
URB				1.006	0.998	1.001	1.001	1.000	1.001
				(0.005)	(0.004)	(0.003)	(0.007)	(0.004)	(0.003)
Constant	0.812	0.983	0.890	1.460	0.572	0.767	0.323	0.209***	0.270***
	(0.213)	(0.204)	(0.145)	(0.680)	(0.230)	(0.225)	(0.225)	(0.111)	(0.105)
State fixed effects	NO	NO	NO	NO	NO	NO	YES	YES	YES
Observations	3,050	4,052	7,102	3,050	4,052	7,102	2,828	4,018	7,014

Notes: For variable definitions, see the notes for Table 3. Standard errors in parenthesis. State fixed effects not reported.
*** p<0.01, ** p<0.05, * p<0.1.

Table 6: Marginal Effects of Electoral Competitiveness and Literacy Rates on the Probability a Party Selects a Self-Reported Criminal Candidate

Model	At baseline	LIT = 42%	LIT = 66%	Marginal effect	COMP = 14%	COMP = 44%	Marginal effect
Model 1, 2004	10.9	14.1	8.5	-39.7	14.7	8.8	-40.2
Model 1, 2009	14.1	16.8	11.8	-29.6	21.0	10.4	-50.2
Model 1, 2004 and 2009	12.8	15.6	10.4	-33.6	18.1	9.8	-45.9
Model 2, 2004	11.4	16.0	8.0	-49.9	15.0	9.2	-38.5
Model 2, 2009	14.3	16.8	12.4	-26.3	20.3	11.0	-45.7
Model 2, 2004 and 2009	12.8	15.9	10.3	-35.2	17.6	10.1	-42.7

Notes: Each cell represents the probability, in percent, that a party fields a self-reported criminal candidate. Covariate levels are set to their mean or zero for dummy variables unless noted otherwise. For the regressors included in Models 1 and 2, see Table 5.

Table 7: Frequency Distribution of Self-Reported Criminal Candidates Across Parliamentary Constituencies, Fourteenth and Fifteenth Lok Sabha Elections

Number	**Freq. 2004**	**Perc. 2004**	**Freq. 2009**	**Perc. 2009**
0	307	56	200	37
1	159	29	174	32
2	46	8	94	17
3	22	4	40	7
4	4	< 1	20	4
5	4	< 1	10	2
6	1	< 1	4	< 1
7	0	0	1	<1

Table 8: Comparison of Characteristics of Candidates within 25 Percent of Winning or Losing a Seat in 2004 who Ran Again in 2009

	Won in 2004	**Lost in 2004**	**Difference**
Proportion winning in 2009	0.45	0.33	0.12**
	(0.03)	(0.04)	(0.05)
Vote share, 2009	0.37	0.29	0.07***
	(0.01)	(0.01)	(0.01)
Vote share, 2004	0.46	0.33	0.13***
	(0.00)	(0.01)	(0.01)
Self-reported criminal indictment	0.23	0.16	0.07
	(0.03)	(0.03)	(0.04)
Criminal opponent, 2009	0.68	0.73	-0.05
	(0.03)	(0.04)	(0.05)
Proportion of INC candidates	0.34	0.34	0.00
	(0.03)	(0.04)	(0.05)
Proportion of BJP candidates	0.24	0.28	-0.04
	(0.03)	(0.04)	(0.05)
Number of obs.	253	145	

Notes: Standard errors in parenthesis. INC: Indian National Congress. BJP: Bharatiya Janata Party. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Summary of Regression Discontinuity Results

	2004 incumbents	**2004 non-incumbents**	**Difference**	**Perc. difference**
All observations	0.31	0.37	-0.06	-19
Candidate's opponent does not report indictment	0.36	0.33	0.03	12
Candidate's opponent reports indictment	0.30	0.38	-0.08	-22

Table 10: OLS Estimates of the Determinants of Electoral Turnout in Parliamentary Constituencies, Fourteenth and Fifteenth Lok Sabha Elections

	2004				2009			
	Model 1a	**Model 1b**	**Model 1c**	**Model 1d**	**Model 2a**	**Model 2b**	**Model 2c**	**Model 2d**
CRIM	-3.563***	-2.778***	-1.818**	-0.259	-3.911***	-3.534***	-2.183**	0.261
	(1.016)	(1.005)	(0.920)	(0.600)	(1.228)	(1.193)	(1.095)	(0.589)
COMP		0.108**	0.017	-0.065**		-0.020	-0.076	-0.075**
		(0.048)	(0.047)	(0.033)		(0.063)	(0.060)	(0.033)
LIT		0.189***	0.364***	0.182***		0.335***	0.447***	0.075
		(0.042)	(0.050)	(0.046)		(0.048)	(0.061)	(0.046)
INCUM			-0.389	-0.694			2.031*	0.537
			(1.061)	(0.643)			(1.035)	(0.535)
RESERVED			-1.245	-1.061			0.959	-0.553
			(1.103)	(0.653)			(1.079)	(0.564)
NUMBER			-3.829***	-2.374***			-1.903**	-1.731***
			(0.797)	(0.685)			(0.838)	(0.541)
POP			-0.000**	-0.000			-0.000***	-0.000***
			(0.000)	(0.000)			(0.000)	(0.000)
URB			-0.251***	-0.195***			-0.216***	-0.172***
			(0.029)	(0.021)			(0.034)	(0.020)
Constant	60.525***	48.455***	61.403***	73.592***	61.875***	43.101***	65.224***	86.231***
	(0.670)	(2.462)	(4.375)	(3.187)	(0.976)	(2.954)	(5.620)	(3.560)
State fixed effects	NO	NO	NO	YES	NO	NO	NO	YES
Observations	543	543	543	543	553	519	519	519
R-squared	0.022	0.071	0.247	0.773	0.018	0.112	0.299	0.832

Notes: For variable definitions, see the notes for Table 3. All variables measured at the level of the parliamentary constituency. Standard errors in parentheses. State fixed effects not reported.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1: Degree of Electoral Competition and Equilibrium Candidate Choice

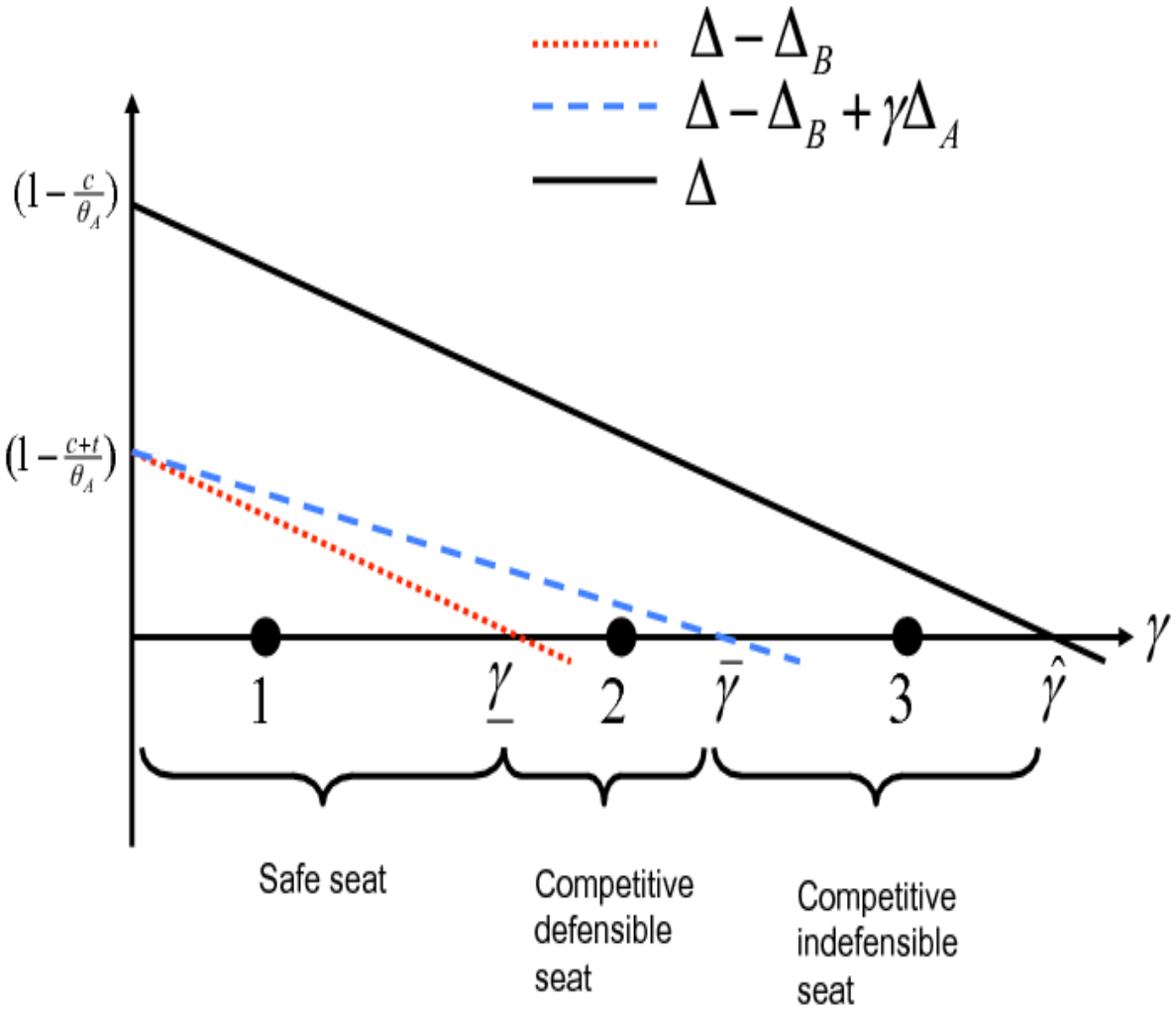
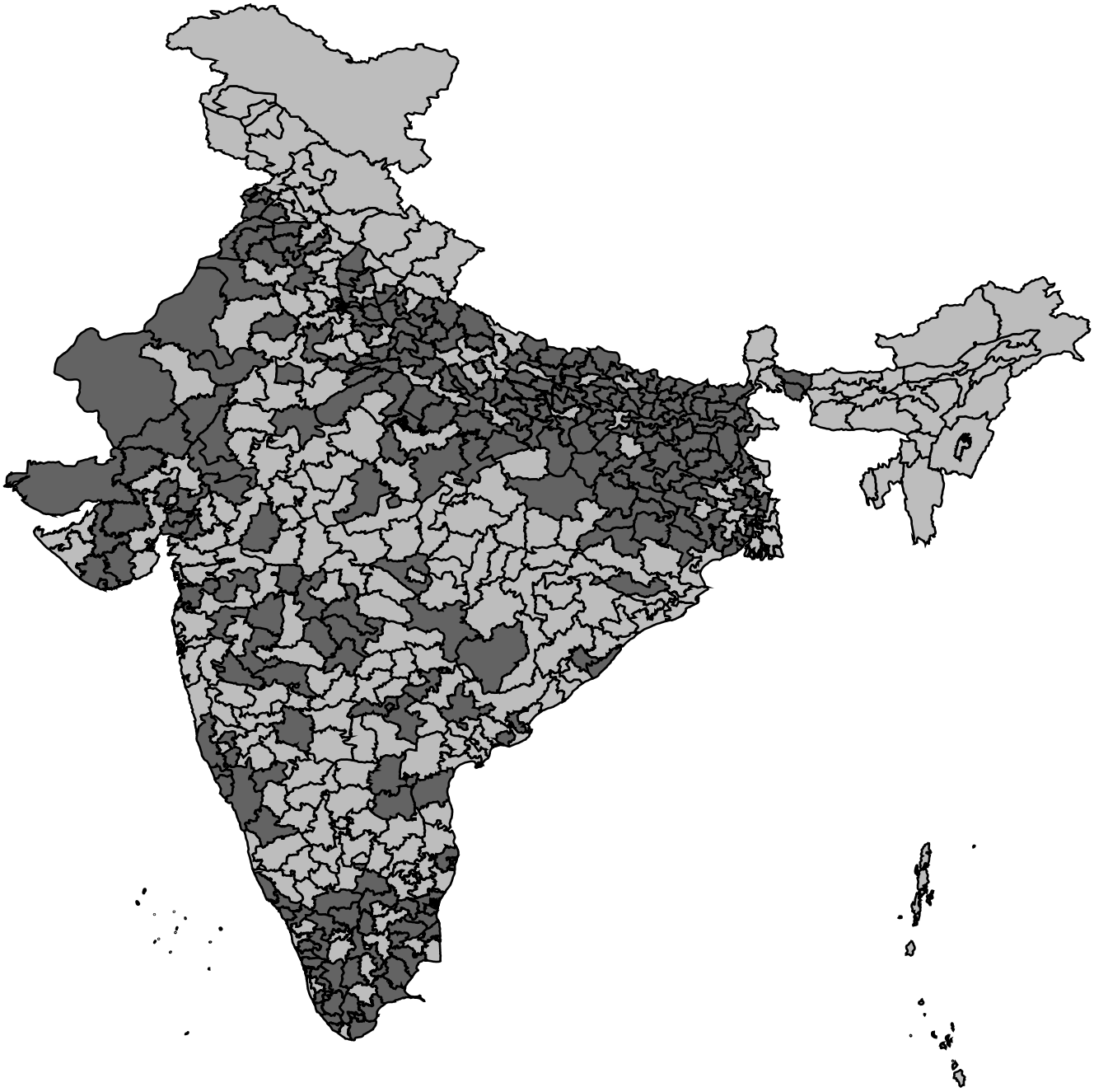
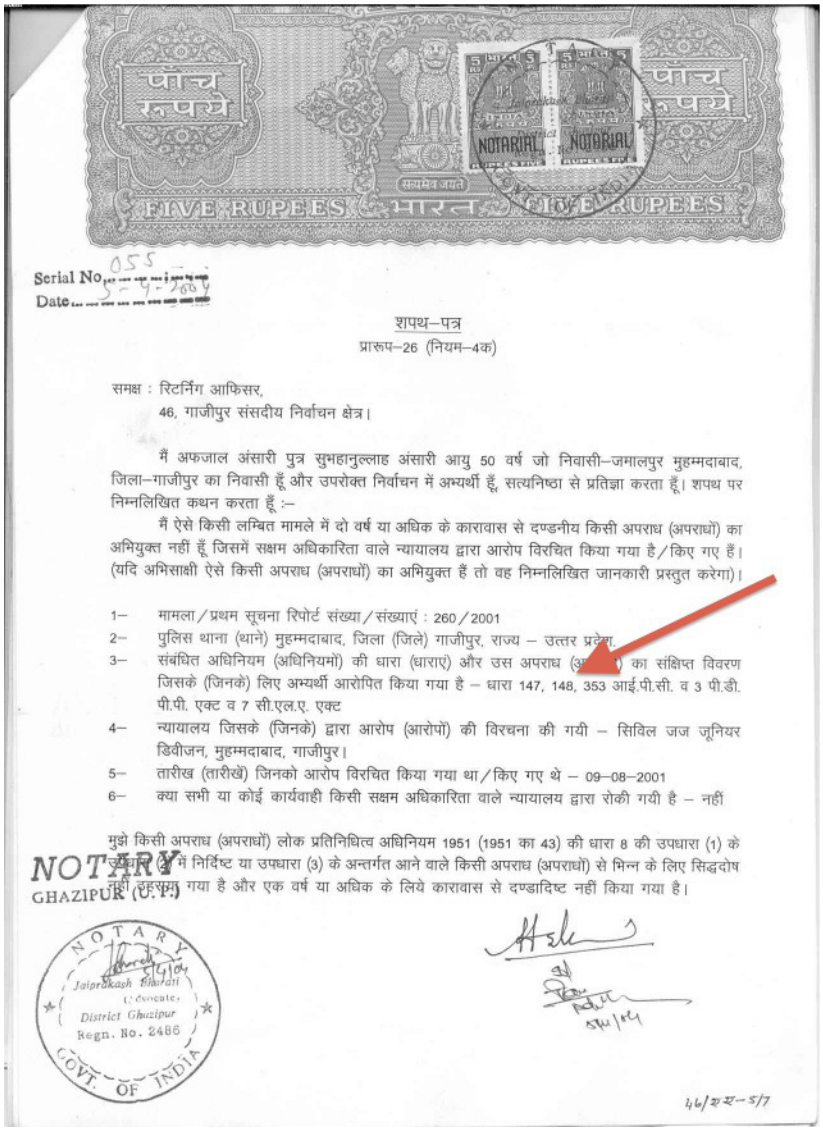


Figure 2: Map of Self-Reported Criminal Candidates in Parliamentary Constituencies for Elections to the Fourteenth Lok Sabha (2004)



Notes: Darker constituencies represent those with at least one self-reported criminal.

Figure 3: 2004 Affidavit of Afajal Ansari, Candidate to the Lok Sabha from Ghazipur (UP), p. 3



पाँच रुपये FIVE RUPEES सत्यमेव जयते भारत NOTARIAL

Serial No. 055
Date 5-4-2004

शपथ–पत्र
प्रारूप–26 (नियम–4क)

समक्ष : रिटर्निंग आफिसर,
46, गाजीपुर संसदीय निर्वाचन क्षेत्र।

मैं अफजाल अंसारी पुत्र सुभहानुल्लाह अंसारी आयु 50 वर्ष जो निवासी–जमालपुर मुहम्मदाबाद, जिला–गाजीपुर का निवासी हूँ और उपरोक्त निर्वाचन में अभ्यर्थी हूँ, सत्यनिष्ठा से प्रतिज्ञा करता हूँ। शपथ पर निम्नलिखित कथन करता हूँ :–

मैं ऐसे किसी लम्बित मामले में दो वर्ष या अधिक के कारावास से दण्डनीय किसी अपराध (अपराधों) का अभियुक्त नहीं हूँ जिसमें सक्षम अधिकारिता वाले न्यायालय द्वारा आरोप विरचित किया गया है/किए गए हैं। (यदि अभिसाक्षी ऐसे किसी अपराध (अपराधों) का अभियुक्त हैं तो वह निम्नलिखित जानकारी प्रस्तुत करेगा)।

1– मामला/प्रथम सूचना रिपोर्ट संख्या/संख्याएं : 260/2001
2– पुलिस थाना (थाने) मुहम्मदाबाद, जिला (जिले) गाजीपुर, राज्य – उत्तर प्रदेश
3– संबंधित अधिनियम (अधिनियमों) की धारा (धाराएं) और उस अपराध (अ[illegible]) का संक्षिप्त विवरण जिसके (जिनके) लिए अभ्यर्थी आरोपित किया गया है – धारा 147, 148, 353 आई.पी.सी. व 3 पी.डी. पी.पी. एक्ट व 7 सी.एल.ए. एक्ट
4– न्यायालय जिसके (जिनके) द्वारा आरोप (आरोपों) की विरचना की गयी – सिविल जज जूनियर डिवीजन, मुहम्मदाबाद, गाजीपुर।
5– तारीख (तारीखें) जिनको आरोप विरचित किया गया था/किए गए थे – 09–08–2001
6– क्या सभी या कोई कार्यवाही किसी सक्षम अधिकारिता वाले न्यायालय द्वारा रोकी गयी है – नहीं

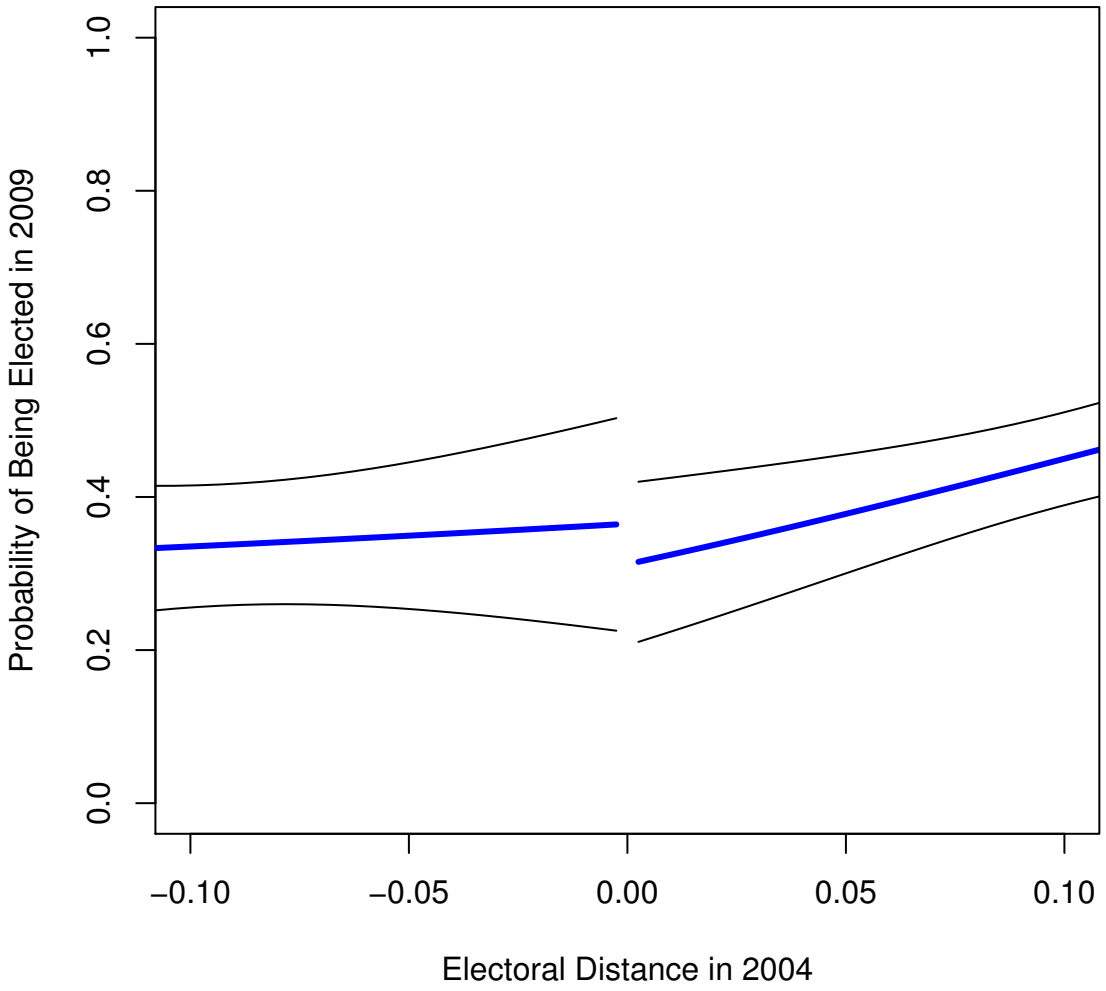
मुझे किसी अपराध (अपराधों) लोक प्रतिनिधित्व अधिनियम 1951 (1951 का 43) की धारा 8 की उपधारा (1) के [illegible] में निर्दिष्ट या उपधारा (3) के अन्तर्गत आने वाले किसी अपराध (अपराधों) से भिन्न के लिए सिद्धदोष [illegible] गया है और एक वर्ष या अधिक के लिये कारावास से दण्डादिष्ट नहीं किया गया है।

NOTARY
GHAZIPUR (U.P.)

NOTARY
Jaiprakash Bharati
Advocate
District Ghazipur
Regn. No. 2486
GOVT. OF INDIA

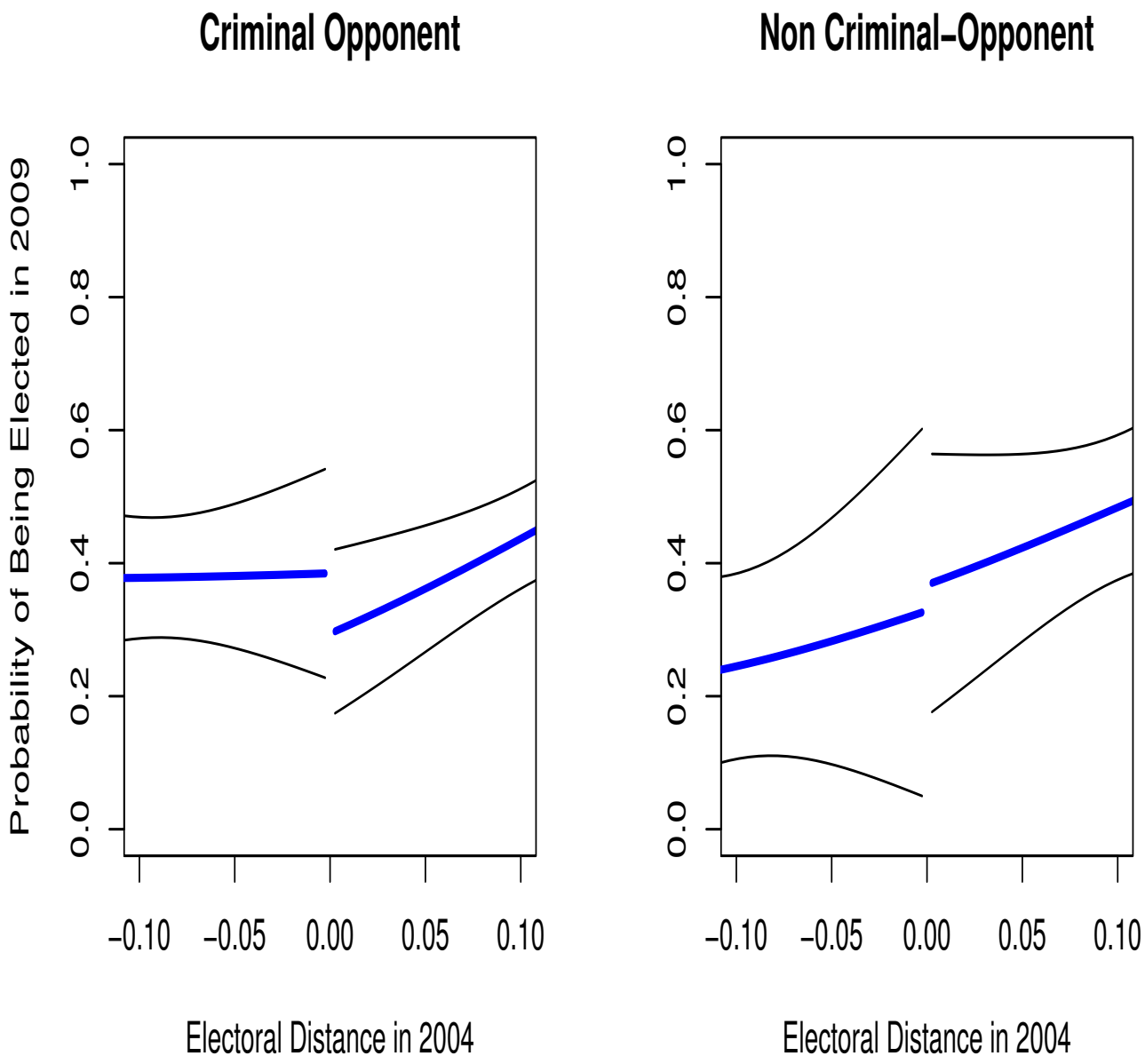
46/22-5/7

Figure 4: Regression Discontinuity Results of the Probability of Being Elected, All Candidates



Notes: Candidates limited to those within 25 percent of winning.

Figure 5: Regression Discontinuity Results of the Probability of Being Elected, Criminal and Non-Criminal Opponent



Notes: Candidates limited to those within 25 percent of winning. Criminal candidates defined as those who report having been convicted of or currently facing criminal charges in their affidavit filed with candidacy papers.