Do Parties and Voters Disagree? An Equilibrium Analysis of Candidate Selection in India*

Arvind Magesan
Economics Department, University of Calgary

Andrea Szabó
Economics Department, University of Houston

Gergely Ujhelyi
Economics Department, University of Houston

December 17, 2025

Abstract

We study how parties choose candidates, a key issue for understanding political selection and policy choices. Do parties select candidates that voters like, or are their choices shaped by other considerations? To study this question, we combine rich candidate-level data from India with a model in which parties trade off candidates' electoral appeal against internal party preferences in a strategic game of candidate selection. We find that parties' preferences deviate from voters'. While parties select likely winners, all else equal they prefer candidates who are not overly popular. Selection decisions are also shaped by strategic considerations and factors independent of voter preferences, such as recruitment costs. Our estimates explain parties' nomination of criminal candidates, their aversion to some incumbents, and, through counterfactual simulations, their likely responses to policies affecting candidate eligibility, such as term limits and restrictions on candidates with a criminal history.

*For useful comments and suggestions we thank Victor Aguirregabiria, Kentaro Fukumoto, Jun Goto, Brad Hackinen, Jean-François Houde, Kei Kawai, Robert Miller, Paul Novosad, Amil Petrin, Yasutora Watanabe, Eric Weese and participants in seminars and conferences at Calgary, Caltech, GRIPS Tokyo, Osaka, Queens University Belfast, Tokyo, University College Dublin, CPEG 2023, the 2024 APSA Virtual Research Meeting, the 7th Banff Empirical Microeconomics Conference and the 2024 Stony Brook Empirical Methods in Game Theory Workshop. Magesan gratefully acknowledges funding from SSHRC through IG grant 435-2019-0437. An earlier version of the paper was titled "Candidate Selection by Parties: Crime and Politics in India."

1 Introduction

Do parties select candidates that voters like, or are their choices shaped by other considerations? In a representative democracy, elections aggregate voter preferences over candidates *who appear on the ballot*. Therefore the question of whom parties choose to run is key to understand political selection, and ultimately policy choices.

Although a sizable literature now explores individuals' own decision to enter politics (Besley 2005; Dal Bó and Finan 2018), we know much less about how parties select their candidates. As Dal Bó and Finan (2018, p.566) write: "political parties likely play a major role in who becomes a politician, and yet we have a very limited understanding of how political parties recruit and screen their candidates."

Parties like to win, which requires running candidates that voters will support. At the same time, there are several reasons why a party may not necessarily want to run the most popular candidate. Party elites may value candidate traits like loyalty, influence or wealth even if these do not lead to more votes. For example, wealthy candidates can bring resources to the party or self-finance their campaign. Similarly, candidates from influential groups can provide valuable connections and access for party leaders. Such benefits have been proposed as an explanation for the widespread nomination of (wealthy and well-connected) criminal politicians in India (Vaishnav 2017).

Sometimes a candidate who enjoys "too much" voter support may in fact threaten the position of party elites, or steer the party's policies in a direction that is inconsistent with the elite's preferences. In the US, these concerns are well illustrated by changes in the Republican party since the 2016 electoral campaign of Donald Trump. In India, where parties are highly centralized, candidates who become too popular can also give rise to internal power struggles and weaken the party leader (Chandra 2016).

In this paper, we study candidate selection in Indian national elections, combining a discrete-choice model of voter preferences over candidates with a strategic game of candidate selection between the two main party alliances. Estimating the model quantifies several factors that cause party choices to systematically deviate from the maximization of winning probabilities (or vote shares). We use our framework to model the impact of policies that affect the set of candidates available to parties, such as term limits or banning candidates with criminal backgrounds.

Focusing on the 2009 and 2014 national elections, we begin by estimating a comprehensive BLP discrete-choice demand system describing voter preferences among candidates (Berry et al. 1995). Voters have preferences over several (endogenous) candidate characteristics, in-

cluding incumbency status, wealth, education, criminal background, and Muslim ethnicity.[1] These preferences are also shaped by constituency characteristics and unobserved candidate characteristics.

Our specification of the "supply side" focuses on candidate choices by the two main party alliances in national politics, the NDA and the UPA, led by the two largest parties, the BJP and the INC, respectively. We use a simultaneous game of incomplete information to model the strategic interaction between these players, and estimate their preferences over candidates using Aguirregabiria and Mira (2007)'s Nested Pseudo-Likelihood procedure.

To make this game and its estimation feasible, we assume that parties' choice sets are comprised of clusters of candidate characteristics, or candidate "types." We approximate the pool of potential candidate types that parties select from in the national election using the set of candidates contesting elections to state legislatures.[2] We then use a clustering algorithm to identify candidate types in this state election data - essentially, these are combinations of candidate characteristics that tend to occur together.[3] This clustering algorithm identifies four candidate types, all of which turn out to have clear interpretations: an "educated type" (educated non-Muslim with no criminal history), an "uneducated type" (uneducated non-Muslim with no criminal history), a "Muslim type," and a "criminal type" (non-Muslim with a criminal history who is also relatively wealthy).

In the candidate selection game, each party chooses one of these candidate types, or (for incumbent parties) chooses to re-nominate the incumbent. Using our demand estimates, we construct counterfactual vote shares corresponding to all combinations of potential choices by the competing parties. This allows us to compute expected vote shares and win probabilities. Parties' objective functions nest these expected vote shares and win probabilities, as well as a set of heterogeneous costs of running different candidates.

Estimating parties' objective functions reveals that while parties prefer candidate types with better chances of winning, all else equal, they prefer candidates with a lower expected vote share. In other words, parties prefer to win with less popular candidates. This finding is consistent with the idea that party elites trade off anticipated electoral performance against the threat of a candidate becoming too powerful and undermining the leadership or its

[1] Muslims are a salient group in Indian politics, but this characteristic is rarely used in academic research due to a lack of data. We create an indicator for Muslim candidates based on their names using text analysis. Throughout we use "Muslim" to describe ethnicity, proxied by name, rather than religion.

[2] Many national election candidates begin their political careers at more local levels of politics, supporting the idea that candidates for state election approximate the pool from which national candidates are drawn.

[3] This approach is similar in spirit to Bandiera et al. (2020) who reduce high dimensional data on CEO activities to a small set of CEO types in order to study how CEO behavior affects firm performance. Hamilton et al. (2021) use a clustering approach to reduce the choice set of patients choosing between different medical treatments.

policies, as suggested by the literature on Indian electoral politics (see Section 2).

We also find that considerations that are independent of voter preferences (and hence win probabilities or vote shares) matter in parties' objective functions. All else equal, candidate types that are more common in the relevant local candidate pool are less costly for parties to recruit and run. Parties also have direct preferences over candidate types: for example, the NDA has a particularly large direct cost of running a Muslim type, which is in line with the BJP's declared Hindu nationalist profile. We estimate large direct costs to incumbent replacement, consistent with incumbents' experience, local networks, and mobilization and fundraising capacity benefiting the party beyond their effect on votes.

Our estimates can be used to understand why we observe specific types of candidates in the data, and to decompose this into voter-driven and party-driven factors. We address two separate long-standing questions in Indian politics: why there are so many politicians with criminal backgrounds, and why incumbents are less successful than in some other countries.

With respect to criminal politicians, the driving forces are voters' positive valuation of these politicians' wealth, and from the parties' perspective, their widespread availability and a strategic complementarity between competing parties' criminal candidates. With respect to "anti-incumbency," we find some disutility to voters from the wealthiest incumbents, but most incumbents are popular with voters. In fact, incumbents can be "too popular" for party leaders, and in such cases are replaced with new candidates.

Our model is well-suited to simulate the impact of policies that place restrictions on the set of candidates that parties can run, such as bans on criminal politicians and term limits. According to our counterfactual analysis, one implication of a criminal ban is to change the distribution of candidate types contesting elections, leading to higher fractions of educated, uneducated, and Muslim types. Interestingly, we find that in many constituencies *both* parties benefit from a criminal ban. The reason for this is the strategic complementarity of criminal candidates. Once their opponent is banned from running a criminal, each party is able to compete with candidates who deliver higher payoffs in equilibrium. This provides an explanation for why political parties could support a ban on criminal candidates while at the same time running such candidates in elections.

Our analysis of term limits shows little impact on the distribution of politician types. Parties forced to drop their incumbent candidates replace them with similar politicians. This calls into question the idea that term limits would broaden the spectrum of characteristics of elected representatives in this context.

Our paper builds on a growing literature on the role of parties in political selection. Caillaud and Tirole (2002) and Mattozzi and Merlo (2015) study the theoretical tradeoffs that parties make between electability and party organization and how this impacts the

quality of candidates. In Galasso and Nannicini (2011) parties find it costly to recruit high-quality candidates and therefore only do this in more competitive districts, which is consistent with data from Italy. Besley et al. (2017) study a model where running the best candidate would jeopardize party leaders' survival, and find evidence consistent with their predictions in Sweden.[4] In the developing country context, two recent papers conduct field experiments to study how party selection decisions depend on candidate nomination procedures. Both Gulzar et al. (2021) (in Nepal) and Casey et al. (2021) (in Sierra Leone) show that providing party elites with information about candidates and their popularity among voters results in parties fielding more candidates preferred by voters, which is consistent with party leaders failing to maximize votes at baseline.

In contrast to papers where candidate quality is measured in terms of education or residuals from a Mincerian wage regression, our approach makes it possible to study selection on multiple dimensions simultaneously. We do not take an a priori stance on what constitutes a high quality candidate, and we also let the data tell us what the relevant types of candidates are in parties' choice sets.[5] Our approach advances the literature by quantifying the different forces that shape candidate selection as well as by explicitly accounting for the strategic interaction between parties. Our structural approach makes it possible to answer questions regarding *counterfactual* policy interventions, including interventions in different settings, which adds to the external validity of our findings.

Conceptually, the papers closest to ours are Iaryczower et al. (2024) and Longuet-Marx (2025), who develop and estimate structural models of candidates' policy positions (ideology) in the context of Brazilian legislative and US congressional elections, respectively. In both of these papers, candidates are exogenously given, and choose policy positions to trade off their own electability against the cost of deviating from an ideal point (such as the party's national platform). By contrast, we focus on the game played by parties as they select their candidates, and in doing so, we take ideology as exogenously given. Unlike the Brazilian or US settings, where candidates in the same party can have disparate ideological positions, in India the central party organization sets the policy platform and elected candidates have little scope to deviate from it (see Section 2). Candidate selection is a discrete game that results in endogenous candidate characteristics over multiple dimensions, and we use estimation methods that account for these features in the analysis.

[4] Casas-Arce and Saiz (2015) show related findings from Spain. In a nondemocratic setting, François et al. (2023) study how party organization shapes political selection within the Chinese Communist Party.

[5] Dal Bó et al. (2017) consider candidate selection on multiple dimensions in their comprehensive study of Swedish politicians. While their focus is on *self*-selection, they also show evidence that parties promote competent individuals (measured by cognitive and leadership skills) independently of their background. In Sweden socioeconomic and ethnic background likely play a considerably smaller role in politics than in our context.

2 Background

2.1 Parties and elections

We study general elections to India's national legislature (the *Lok Sabha*), a near-ideal setting for analyzing strategic candidate selection by political parties. The Lok Sabha is comprised of a large number of single member districts (called *constituencies*). Members are elected via first-past-the-post voting. In contrast to a setting with proportional representation, each competing party selects a single candidate to run in the constituency. Elections are held every 5 years and there are no term limits. Current constituency boundaries were established after nationwide redistricting (*delimitation*) in early 2008.

India is a multiparty system with over 10 candidates contesting the average constituency. In most cases, two of these candidates represent the main competing (pre-election) alliances, the United Progressive Alliance (UPA) and the National Democratic Alliance (NDA) who together win the majority of seats and are led by the two main national parties, the INC and the BJP, respectively. Each alliance runs candidates in almost all constituencies under pre-election agreements that specify which member party will contest the constituency without competition from other members. Because our model will treat alliances as players, we will refer to the two alliances throughout simply as *parties*.

We also use information from elections to states' legislative assemblies. These are separate elections (each national constituency is subdivided into several state constituencies), and in most cases are held in different years from the national election. The set of parties competing in national and state elections can also be different, but the UPA and the NDA are major forces in state elections as well.

2.2 Candidate selection and party goals

Indian parties are known for their centralized organizations in which a central committee, or in some cases a charismatic leader, dictates all major decisions, including candidate selection. Reviewing nomination procedures used in different countries, Farooqui and Sridharan (2014) note that "the USA represents the decentralised extreme, that of party primaries" while "India lies near the other extreme in that most of its major parties are at the completely or near-completely top-down of the six types of party nomination processes, with the national party leadership having the final say." (p.80)[6]

[6]Although both the INC and the BJP have formal consultation procedures that involve local party organizations in the candidate selection process, in practice decisions are ultimately made by each party's central committee (Roy 1966; Farooqui and Sridharan 2014).

Historically, Indian parties have been characterized by internal conflicts as party leaders attempt to maintain their power and subvert factions that emerge to challenge it (Hasan 2010; Chandra 2016). In a particularly consequential example, in 1966 INC party leaders chose Indira Gandhi to take over as prime minister with the hope that she would be a "puppet" under their control (Genovese 1993). Gandhi eventually took complete control of the party, and the former leaders who helped her to power were pushed out to form another party, the INC(O), which is now defunct. In other cases, the challenge to party leaders takes the form of politicians defecting to rival parties of successfully forming new ones. (Chhibber et al. 2014). Party leaders' concern about party-switching led to a 1985 Anti-Defection Law, passed unanimously as part of the 52nd Amendment to the Indian constitution. Under the law, elected members of parliament face disqualification (typically followed by a by-election) for defecting to another party.

In this environment, parties' candidate selection decisions are shaped by factors like loyalty to the party leadership and service to the party organization.[7] As Chandra (2016, p.224) notes:

> "The central leadership's decisions are influenced not only by the anticipated electoral performance of the candidate, but by two intra-party considerations: (1) to ensure the compliance of powerful factions within the party so that they work for the party candidate or at a minimum do not work against the candidate and do not defect from the party and (2) to undercut factions that are becoming too powerful so that they do not threaten the leadership or its loyalists. Expectations of the electoral performance of the candidate are often subordinated to these two criteria."

Of course, loyalty is not the only service that a candidate can provide to the party. Another important factor influencing parties' candidate selection is financial considerations: financial contributions to the party and a candidate's ability to finance their own campaign. Farooqui and Sridharan (2014, p.87) describe, in the case of the BSP party, the process through which candidates effectively bid to receive the nomination. Similarly, Vaishnav (2017) argues that the main appeal of criminal politicians to Indian parties stems from the fact that these individuals can finance their own campaigns, including by breaking campaign finance laws if necessary.

[7] These considerations are often explicit in parties' written procedures on candidate selection (Roy 1966). For example, the INC has declared: "Winnability alone should not be the benchmark for deciding nominees of the party during elections. There should be a balance required between loyalty and winnability." (Jaipur Declaration of the AICC, 2013, quoted in Chandra (2016, p.41)).

2.3 The tradeoff between popularity and party service

Theory suggests that candidates who are more popular among voters yield lower levels of loyalty and party service. In their seminal book on the "personal vote," Cain et al. (1987) contrast candidates' effort to build their own brand and their effort to build the party, and argue that more successful candidates are better positioned to do the former. In Mattozzi and Merlo (2015), "superstar" candidates discourage other party members from providing effort to the party. In Besley et al. (2017), candidates who are popular with voters are more likely to successfully challenge leaders in internal party elections. More generally, popular candidates can make demands on the party and credibly threaten to switch parties if their demands are not met.[8]

A case in point is Sarekoppa Bangarappa, a charismatic politician who won 4 elections with 4 different parties during his 44-year political career. Bangarappa started as a state-level politician in Karnataka but eventually rose to represent the constituency of Shimoga with the BJP, winning the 2004 election with over 50% of the votes. The following year, while a sitting member of parliament, Bangarappa defected to another party (the Samadjwadi Party), forcing a by-election in the constituency, which he won, costing the BJP its seat in the Lok Sabha. Reportedly, one of the reasons for Bangarappa's defection was that he had been passed over for an important position within the party.[9]

Explicitly testing for the relationship between popularity and party service or loyalty is made complicated by the fact that these are equilibrium phenomena reflecting the decisions of voters, candidates, and party leaders. A suggestive statistic comes from a group of "independently popular" politicians - Indian actors who entered politics. In this group, a remarkable 36% switched parties at least once.[10] [11]

In line with these theoretical considerations and anecdotal evidence, our model below will allow parties to have costs (or benefits) from running popular candidates over and above their impact on win probabilities.

[8] Interestingly, the Anti-Defection law mentioned above reinforces the bargaining power of more popular candidates compared to less popular candidates. Because party-switching triggers a by-election, more popular candidates who switch are more likely to keep their seat - which makes their threats to switch more credible.

[9] https://gulfnews.com/world/asia/india/bangarappa-quits-bjp-to-join-mulayams-party-1.280303

[10] From a list of 135 notable Indian actors who are also politicians (https://en.wikipedia.org/wiki/Category:Indian_actor-politicians, accessed February 2025), we were able to confirm both the starting and current party affiliation for 128. Of these politicians, 36% had different starting and current parties, implying that they switched parties at least once.

[11] One of these is iconic actor Shatrughan Sinha, who played in over 150 films before starting a career in politics. Sinha was a member of parliament with the BJP but frequently criticized the party's leadership and its policies during the 2010s. In 2019, he left the BJP, and in 2022 won an election with a different party (the TMC).

2.4 Criminal politicians

Criminality is a salient feature of Indian politics (see Vaishnav (2017) for a detailed discussion). In 2019, 43% of candidates elected to the national legislature had been indicted on criminal charges at some point.[12] By law, criminally convicted politicians are ineligible to run for election until six years after the completion of their sentence. However, due to the slow pace of court proceedings, many indicted politicians run, and win, while undergoing trials that can last for decades. Following the literature, our discussion of criminal backgrounds refers to indictments, i.e., cases where a judge decided to proceed with a criminal charge beyond the initial police investigation and prosecutorial action.

For a sample of 868 UPA and NDA candidates analyzed below, 290 have a criminal background, with a total of 2,916 charges listed under the Indian Penal Code. Grouping the 15 most frequent charges, we find charges related to rioting and unlawful assembly (22% of all charges), violent acts such as wrongful restraint, assault, criminal intimidation and attempted murder (16%), disobedience or violence against public servants (10%), and conspiracy (6%). There are 42 charges of murder or homicide and 21 charges of kidnapping.

Over the last several decades multiple commissions tasked with electoral reform in India have recommended tightening the restrictions on criminal candidates.[13] Perhaps surprisingly, the two main parties themselves have at different times expressed a desire to ban criminal candidates[14] - even while actively running such candidates in elections. Our results below will help rationalize this phenomenon.

3 Data

Elections. Our main data combines official election returns for the 2009 and 2014 national elections with candidate and constituency characteristics. Data from the Election Commission of India (ECI) contain each candidate's name, party, and vote return, as well as turnout (summary statistics are in Appendix Table A.1.) Our specification requires information on local (state legislative) elections matched to "corresponding" national elections. For each state, we assign the first state election held after the 2008 delimitation to the 2009 national

[12] https://adrindia.org/content/nearly-50-cent-mps-new-lok-sabha-have-criminal-records

[13] For example, in 2004 the Election Commission unsuccessfully recommended that indicted politicians be banned from elections In 2013, in Lily Thomas v Union of India, the Indian Supreme Court closed a loophole allowing some politicians to run while appealing their convictions.

[14] https://economictimes.indiatimes.com/news/politics-and-nation/bar-criminals-from-fighting-polls-sonia-gandhi/articleshow/5500935.cms

election, and the second state election to the 2014 national election.[15] We also include data from the 2004 general election and from by-elections held between general elections to correctly identify incumbents in each constituency.

Candidate characteristics. The ECI data contains information on candidates' party, gender, age, and SC/ST status. The ECI also publishes information on broadcast time allowance: these allowances for political advertising on public TV and radio are distributed to parties based on performance in past elections. We supplement this data with information on candidates' education, wealth, and criminal history collected and published by the civil group ADR at www.myneta.info. This is based on affidavits that all Indian candidates for national and state elections are required to file with the ECI. The criminal histories that candidates are required to report include previous criminal convictions and indictments.[16]

Although this dataset on candidates is already quite rich, it does not contain information on one of the most important characteristics in Indian politics: whether the candidate is an ethnic Muslim. We construct a Muslim indicator based on candidate names using text analysis, categorizing candidates based on the "distance" between their name and text fragments commonly found in Muslim names. The details are in Appendix 1.

Identifying incumbents requires matching candidate names across different elections over time. Variations in the spelling of Indian names makes this a complex task. The Trivedi Centre for Political Data made a comprehensive attempt to do this (Bhogale et al. 2019) and we updated this dataset. We define incumbents as the candidates in office at the time of the election in the same constituency.[17] For incumbents who were not re-nominated, we searched for news stories to check whether this was due to exogenous reasons (e.g., death).

Constituency characteristics. In the Indian electoral system, some constituencies are reserved and can only be contested by Scheduled Caste (SC) or Scheduled Tribe (ST) candidates and the ECI data contains indicators for these reserved constituencies. For constituency demographics, we use the SHRUG dataset (Asher et al. 2021), which allows village-level information from the 2011 Indian Census to be matched to constituencies. We use the following characteristics: literacy rate, share of employed population, share of Scheduled Caste and Scheduled Tribe population, whether the village has access to paved roads, and whether the village is located in a rural or urban area. We use both the village level information, and also aggregate it up to the constituency level. This creates some missing constituencies when

[15] In practice this means that state elections held between 2008-2012 are assigned to the 2009 national election and state elections held between 2013-2017 are assigned to the 2014 national election.

[16] Previous studies using the affidavits data include Fisman et al. (2016), Vaishnav (2017), Prakash et al. (2019), Ujhelyi et al. (2021), and Asher and Novosad (2023), among others.

[17] For the 2009 election, the "same" constituency refers to the pre-redistricting constituency that contained the largest area of the constituency (see below).

villages could not be uniquely matched to constituencies (see Asher et al. (2021)).

The April 2008 redistricting changed most constituency boundaries as well as the reservation status of some constituencies. To measure a constituency's exposure to redistricting, we overlap geocoded constituency boundaries before and after 2008, and for each new constituency, find the old constituency with the largest share of its area.

Sample construction. Details of our sample construction are in Appendix 1. We drop constituencies with missing demographic information, and states with very few constituencies. Our final sample includes the 15 largest Indian states, and contains 232 national constituencies and 1629 state constituencies (about half of the constituencies in these states). These national constituencies are contested by a total of 3,208 candidates in 2009 and 3,373 candidates in 2014. The corresponding state constituencies are contested by 17,965 and 18,801 candidates, respectively.

4 Model

We consider a simultaneous Bayesian game of candidate selection between competing parties. Candidates are described by a set of characteristics, and each party weighs its internal preferences against the preferences of voters, and thus the probability of winning.

4.1 Parties

Consider an electoral constituency where competing parties choose which candidate to run. Each party p chooses one candidate out of a set of potential candidates $\mathcal{A}_p$, which may include the option to re-nominate the current incumbent. Let the choice of party p be given by a_p, and the vector of choices of p's opponents by $\mathbf{a}_{-p}$. Voters cast their votes based on the candidates that parties choose to run: let $s_p(a_p, \mathbf{a}_{-p})$ represent the votes cast for party p given a candidate selection profile $(a_p, \mathbf{a}_{-p})$, and let $w_p(a_p, \mathbf{a}_{-p})$ represent its corresponding winning probability (these functions will be derived endogenously below).[18]

A key element of our approach is to allow for the fact that parties may care about the candidate they run beyond its effect on votes. A party may experience costs or benefits from the effort required to recruit certain candidates, those candidates' (dis)loyalty to the party leadership, or their contributions to party finances. To capture these considerations, we specify party p's payoff from choosing candidate $a_p \in \mathcal{A}_p$ as

$$b^w w_p(a_p, \mathbf{a}_{-p}) + b^s s_p(a_p, \mathbf{a}_{-p}) - c_p(a_p) + \varepsilon_p(a_p). \tag{1}$$

[18] Throughout, s_p will correspond to the share of eligible voters who voted for party p (rather than the share among those who turn out to vote). We will refer to s_p as "vote share" for simplicity.

The party cares about its winning probability, with weight b^w, as well as its vote share. The weight b^s on the latter could be positive if, e.g., the party leader's status is enhanced by a large vote share. It could also be negative if popular candidates challenge the leader's authority or defect and form new parties, as discussed in Section 2.3.[19]

The term $-c_p(a_p) + \varepsilon_p(a_p)$ captures payoffs from a_p that are independent of w_p and s_p. In the estimation, $c_p(a_p)$ will be measured relative to the payoff from re-nominating an incumbent. Thus, it may be viewed as the "replacement cost" of choosing a new candidate a_p. While $c_p(a_p)$ is observable to all competing parties, $\varepsilon_p(a_p)$ is party p's private information. The private component $\varepsilon_p(a_p)$ is i.i.d. across parties and candidates, with cdf $G(\cdot)$.

An important factor affecting $c_p(a_p)$ is the pool of potential candidates available to the party. This is determined by who the party's members are, and who among its members has both the motivation and ability to run for office. Finding a Muslim candidate could be easier if the pool includes more Muslims. Running a candidate with a criminal history could be more acceptable within the party if the party has many such candidates in the pool.

As explained below, we will measure a party's pool of potential candidates using characteristics of the party's candidates in state elections. This is motivated by the fact that (i) many Indian parties competing in state elections have clear affiliations to a national party (either the party is the same, or they belong to the same electoral alliance), and (ii) it is common for national politicians to begin their political careers in state elections.[20] To highlight this, write $c_p(a_p) = c_p(a_p, L_p)$, where L_p denotes the pool of party p's candidates in relevant state elections.

Given the presence of private information, this setup gives rise to a simultaneous game of incomplete information between parties competing in the constituency. The solution concept is Bayesian Nash Equilibrium (BNE) in pure strategies.[21] For a realization of the private component $\boldsymbol{\varepsilon}_p \equiv \{\varepsilon_p(a)\}_{a \in \mathcal{A}_p}$, a party chooses candidate $a_p(\boldsymbol{\varepsilon}_p)$. Let $P(\mathbf{a})$ denote the ex ante probability of a profile of choices $\mathbf{a}$. Then given ε_p and $P(\mathbf{a}_{-p})$, in a BNE party p chooses a_p to maximize its expected payoff

$$U_p(a_p, P(\mathbf{a}_{-p})) \equiv E_P[b^w w_p(a_p, \mathbf{a}_{-p}) + b^s s_p(a_p, \mathbf{a}_{-p}) | a_p] - c_p(a_p, L_p) + \varepsilon_p(a_p), \tag{2}$$

[19] Instead of vote shares, these considerations could alternatively be captured by vote margins (difference in vote share relative to the winner) or by the number of votes (implying higher values for larger constituencies). We find that using any of these alternatives makes little difference to the results - see Appendix 7.

[20] The most prominent example is India's current Prime Minister, Narendra Modi, who was a long time MLA in the state of Gujarat. But there are numerous examples of MPs who once ran in elections to the state legislature, including recent examples, Smriti Irani (BJP), and Randeep Singh Surjewala (INC). Matching national politicians' names to all state elections in the same state held since 1990, we find exact matches for 36-47% of BJP and INC candidates.

[21] As long as $G(\cdot)$ is atomless, the existence of a pure-strategy BNE is guaranteed: see Fudenberg and Tirole (1991, Ch 6.8).

where the expectation is over the possible realizations of opponents' choices $\mathbf{a}_{-p}$.

For the purposes of estimation it is convenient to express strategies as *choice probabilities* (CPs). In particular, define payoffs net of the private component as $\tilde{U}_p(a, P(\mathbf{a}_{-p})) \equiv U_p(a, P(\mathbf{a}_{-p})) - \varepsilon_p(a)$, so that a_p maximizes party p's expected payoffs iff $\tilde{U}_p(a_p, P(\mathbf{a}_{-p})) + \varepsilon_p(a_p) \geq \tilde{U}_p(a, P(\mathbf{a}_{-p})) + \varepsilon_p(a)$, $\forall a \in \mathcal{A}_p$. The probability of party p choosing action a_p given the opponent's strategy $P(\mathbf{a}_{-p})$ is then

$$\begin{aligned} P(a_p) &= \int_{\boldsymbol{\varepsilon}_p} \mathbf{1}\Big\{\varepsilon_p(a) - \varepsilon_p(a_p) \leq \tilde{U}_p(a_p, P(\mathbf{a}_{-p})) - \tilde{U}_p(a, P(\mathbf{a}_{-p})), \forall a \in \mathcal{A}_p\Big\} dG(\boldsymbol{\varepsilon}_p) \\ &\equiv \Lambda_p(a_p; \mathbf{P}_{-p}) \end{aligned} \tag{3}$$

Equilibrium in the game is fully characterized by a fixed point in $P(\mathbf{a})$ of the system of equations defined by (3) for all p. Stacking equations by actions and parties, an equilibrium vector of CPs $\mathbf{P}^*$ satisfies $\mathbf{P}^* = \boldsymbol{\Lambda}(\mathbf{P}^*)$.

Under the assumption that $\varepsilon_p(a)$ follows the Type-I Extreme Value Distribution, the equilibrium CPs satisfy

$$\Lambda_p(a_p; \mathbf{P}_{-p}) = \frac{\exp\left\{\tilde{U}_p(a_p, \mathbf{P}(\mathbf{a}_{-p}))\right\}}{\sum_a \exp\left\{\tilde{U}_p(a, \mathbf{P}(\mathbf{a}_{-p}))\right\}} \ \forall p.$$

An important assumption, and limitation, of the above model is that it considers each constituency in isolation. That is, conditional on observables, choosing a candidate in one constituency has no bearing on the candidate chosen in another constituency. In practice it is possible that even after conditioning on observables, party decisions in one constituency affect decisions in another. Allowing for a party to jointly decide on candidates across constituencies (a "Colonel Blotto" type game) would render the model inestimable, as this would leave us with as many markets as we have national elections (two).

4.2 Voters

To model parties' winning probabilities and vote shares, we consider the individual decisions made by a continuum of voters. We assume expressive voting with a flexible specification of voter preferences over candidates' characteristics.[22]

Specifically, each candidate a_p can be described by a vector of characteristics $\mathbf{x}_p = \mathbf{x}(a_p)$, such as their education level or criminal history. Given a set of candidates that parties have

[22]The assumption of expressive voting is supported by extensive survey evidence on Indian voters' motivations. For example, Heath and Ziegfeld (2022) estimate that at most 1.1% of individuals vote strategically.

chosen to run, voter i's utility from voting for the candidate of party p is

$$V_{ip} = \boldsymbol{\beta}_i \mathbf{x}_p + \xi_p + \eta_{ip}. \tag{4}$$

The first term represents voters' (potentially heterogeneous) preferences over the characteristics of p's candidate. The second term, ξ_p, allows for unobserved (to the researcher) shocks to parties' popularity in the given constituency or alternatively, unobserved candidate characteristics valued by voters.[23] The distribution of ξ_p is left unspecified, and it can be correlated with $\mathbf{x}_p$. Finally, η_{ip} are individual preference shocks drawn from a Type-I Extreme Value distribution. To model the sources of preference heterogeneity among voters, write

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\Pi} \mathbf{d}_i, \tag{5}$$

where $\mathbf{d}_i$ is a vector of voter demographics, while $\boldsymbol{\beta}$ and $\boldsymbol{\Pi}$ contain the parameters. The voter's choice set is completed by letting $p = 0$ indicate the option to abstain. The voter's associated utility $V_{i0} = \boldsymbol{\pi}_0 \mathbf{d}_i + \eta_{i0}$, which allows for the utility of abstention (the cost of voting) to vary across voters.

Voter i chooses option p (vote for one of the parties or abstain) if $V_{ip} > V_{ip'}$ for all $p' \neq p$. Thus, voters choose between their options based on the observed and unobserved candidate characteristics, the benefit of abstention, and their idiosyncratic shocks. This implicitly defines the set for which voter i will choose option p, $\{(\mathbf{d}_i, \boldsymbol{\eta}_i) | V_{ip} > V_{ip'} \text{ for all } p' \neq p\}$. Given a distribution of $\mathbf{d}_i$ and $\boldsymbol{\eta}_i$, integrating over this set yields parties' vote shares as a function of their candidate choices. Under the assumed Type-I EV distribution for η_{ip} and given a distribution $F(\mathbf{d}_i)$, these vote shares can be written as

$$s_p(\mathbf{x}_p, \mathbf{x}_{-p}) = \int \frac{\exp\left[\boldsymbol{\beta}_i \mathbf{x}_p + \xi_p - \boldsymbol{\pi}_0 \mathbf{d}_i\right]}{1 + \sum_{q>0} \exp\left[\boldsymbol{\beta}_i \mathbf{x}_q + \xi_q - \boldsymbol{\pi}_0 \mathbf{d}_i\right]} dF\left(\mathbf{d}_i\right). \tag{6}$$

This setup allows the domain of voter preferences to be different from the parties' choice sets (voter preferences are defined over characteristics $\mathbf{x}$ while the parties' choice set is $\mathcal{A}_p$). This is realistic because not every party may have access to candidates with all possible combinations of characteristics.[24]

To model the relationship between party choices and candidate characteristics, we simply assume that, after choices are made, candidate characteristics are drawn from a distribution

[23] Because each party has at most one candidate in a constituency, these two interpretations of ξ_p are observationally equivalent.

[24] This feature also allows our model to potentially be extended to situations where a party leader picking candidates may not have full control over their characteristics (for example, the leader might delegate candidate choice to subordinates). Modeling this explicitly could be an interesting avenue for future work.

$H(\cdot|a)$ for each party. Thus, party p's vote share and winning probability is the expected value over the realizations of these characteristics,

$$s_p(a_p, \mathbf{a}_{-p}) = \int_{\mathbf{x}_p} \int_{\mathbf{x}_{-p}} s_p(\mathbf{x}_p, \mathbf{x}_{-p}) dH_p(\mathbf{x}_p|a_p) dH_{-p}(\mathbf{x}_{-p}|\mathbf{a}_{-p}) \tag{7}$$

$$w_p(a_p, \mathbf{a}_{-p}) = \int_{\mathbf{x}_p} \int_{\mathbf{x}_{-p}} \mathbf{1}\{s_p(\mathbf{x}_p, \mathbf{x}_{-p}) > s_{-p}(\mathbf{x}_p, \mathbf{x}_{-p})\} dH_p(\mathbf{x}_p|a_p) dH_{-p}(\mathbf{x}_{-p}|\mathbf{a}_{-p}) \tag{8}$$

In equilibrium, a party's expected vote share and winning probability takes into account its opponent's strategy, captured by $P(\mathbf{a}_{-p})$ from equation (3):

$$E_P[s_p(a_p, \mathbf{a}_{-p})|a_p] = \sum_{\mathbf{a}_{-p}} s_p(a_p, \mathbf{a}_{-p}) P(\mathbf{a}_{-p}) \tag{9}$$

$$E_P[w_p(a_p, \mathbf{a}_{-p})|a_p] = \sum_{\mathbf{a}_{-p}} w_p(a_p, \mathbf{a}_{-p}) P(\mathbf{a}_{-p}). \tag{10}$$

5 Specification and estimation

5.1 Overview

Our goal is to estimate parameters of the parties' objective function (2). To do this, we proceed in two stages. In the first stage, we estimate voters' utility functions (4) with a BLP procedure. In the second stage, we use the estimated vote share functions to estimate benefit and cost parameters in (2) using a Pseudo-Maximum-Likelihood procedure.

5.2 Estimating voter preferences

Estimation of voter preferences follows the BLP Generalized Method of Moments (GMM) procedure (Berry et al. 1995). Here we modify a previous application to Indian *state* elections in Ujhelyi et al. (2021).

5.2.1 Specification

We focus on five candidate characteristics in the vector $\mathbf{x}$: incumbency, assets, education, Muslim, and crime.[25] In the data, incumbents' assets are twice as large as non-incumbents'

[25] We standardize assets, education, Muslim, and crime to have 0 mean and unit standard deviation to be consistent with the estimation of party objectives below. We also considered three other characteristics observed in the data: SC/ST status, gender, and age. Gender has very little variation (almost all candidates are male). SC/ST has very little variation once we control for constituency reservation. Age does not seem to be an important characteristic for either voters or parties in these elections.

even after restricting the comparison to the major parties. To allow for corresponding heterogeneity in voter valuations, we also include in $\mathbf{x}$ an interaction between incumbency and assets. In addition, we include in (4) the broadcast allowance allocated to each party, as well as fixed effects for party, alliance, state, year, imputed characteristics, and reserved constituencies.

In our main specification we treat ideology as a party-level characteristic captured by the party and alliance fixed effects. This reflects the context of India's highly centralized parties where party ideology and policy are set by the party elite.[26] In Section 7 we extend our specification using several explicit measures of ideology.

To deal with the presence of many small parties and independent candidates, we follow Ujhelyi et al. (2021) and aggregate these candidates in each constituency. Individual voter characteristics $\mathbf{d}_i$ are drawn from the village-level SHRUG data. The average constituency has 1,310 villages. We draw 400 villages per constituency, with replacement, using village size as weights. For each draw, we assign a vector of voter characteristics using the village-specific distribution.

5.2.2 Identification of endogenous characteristics

The BLP procedure requires the use of instrumental variables (IV) for two reasons: to identify the "nonlinear" parameters $\mathbf{\Pi}$, and to identify the parameters on any variables in $\mathbf{x}$ that parties can adjust in response to the popularity shocks ξ_p (the usual source of endogeneity in the identification of demand). We discuss the instruments for endogenous characteristics here and the identification of nonlinear parameters in the next subsection.

Because the focus of our study is parties' choice of their candidates, we treat all five candidate characteristics as endogenous. Variables that affect parties' choices but do not directly enter voter preferences are valid instruments. In our specification, the characteristics of a party's candidates in state elections, L_p, satisfy this condition as they enter the party objective functions (2) but not voter utility.[27] We create these instruments as a function of the party alliance (UPA/NDA/neither) and the state election constituencies overlapping with the national election constituency. For example, we instrument the assets of a UPA candidate with the average assets of all UPA candidates running in the state constituencies contained in the given national constituency. We create these instruments both for the current election and for the other election in the data.

[26] See Section 2. Based on a provision in the 1985 Anti-Defection Law, members of parliament can in principle lose their seat for voting against their party in a legislative session.

[27] L_p, which affects recruitment costs, plays the same role as a cost shifter that affects firm profits but not consumer utilities in Industrial Organization applications.

The characteristics L_p of a party's candidates in state elections are valid instruments as long as they are uncorrelated with voter valuations ξ_p in the national election. This is similar to the standard identifying assumption in IO applications that rely on variables from other markets to instrument a firm's choices in a given market. The most common versions of this strategy use neighboring markets such as nearby cities ("Hausman instruments," e.g., Nevo (2001)) or sub-markets such as counties within a newspaper's circulation area ("Waldfogel instruments," e.g., Fan (2013)). Both of these versions require *assuming* that markets can be separated (e.g., that demand across cities evolves independently, or that demand in a sub-market evolves independently from demand on the larger market). In our case, there is an *administrative* separation between markets because state elections are separate from national elections. The two elections feature a different set of candidates, different stakes (because state legislatures decide on different issues than the national legislature), and are typically held in different years. This makes the assumption that voter valuations in a given market are independent of candidate characteristics in other markets (from a different election) more plausible.[28]

One potential threat to the identifying assumption is if L_p reflects party choices in the state election that are also affected by popularity shocks, and shocks at the national and state level are correlated. For example, a scandal involving a criminal politician could reduce the popularity of parties running criminal candidates at both the state and national level, and parties could respond by reducing the number of criminal politicians in state elections. The separation between elections attenuates this possibility - for example, the correlation between the popularity shock in the national election and the popularity shock in a state election held in a different year may not be strong enough to change the *average* criminality of candidates chosen by the party in the state constituencies. Below, we directly evaluate the sensitivity of our estimates to potential violations of the exogeneity of the instruments and in each case find that the bias would be quantitatively small.

To instrument for incumbency, we use two sources of variation in the *availability* of an incumbent candidate for the party. The first is the 2008 redistricting: the breaking up of a constituency into more areas makes it less likely that the previous incumbent will be available in one of the new constituencies. To measure this, we use *redistricting overlap*, the largest share of a new constituency's area in an old constituency.[29] The second source

[28] In the context of political markets, Iaryczower et al. (2024) use a similar strategy, instrumenting the ideology of a candidate with the ideology of the same party's candidates in local elections.

[29] Since redistricting took place in 2008, this measure varies across constituencies in 2009 and is always 1 in 2014. Note that we use redistricting as a shock to the availability of an incumbent in a given constituency in order to identify whether voters like incumbents. We do not use it as a shock to the distribution of voters to identify *why* voters like incumbents (such as the incumbent's history with a constituency vs. his general qualifications), which would be problematic (Sekhon and Titiunik 2012).

of variation proxies the likelihood that the party won the previous election (and hence has an incumbent available). To measure this *competitiveness*, we divide the statewide vote share of the candidate's party by the effective number of parties running in the constituency, both in the previous election.[30] We also use an interaction of the competitiveness measure with an indicator for 2014 (to allow for differential impacts after redistricting) and another interaction with the asset instrument (to instrument for the Incumbent $\times$ Asset interaction).

To get a preliminary sense of the relevance and strength of our instruments, we regress each characteristic on the corresponding instruments (Appendix Table A.4). For each characteristic, we find a significant positive correlation with its prevalence in the candidate pool. These correlations, which are interesting in their own right, provide support for the idea that characteristics of the pool of candidates affect whom a party chooses to run.[31] We also confirm the relevance of the incumbency instruments: a higher likelihood that the incumbent is in the party's choice set, proxied either by its past competitiveness or the stability of constituency boundaries, makes it more likely that the candidate will be an incumbent.

5.2.3 Differentiation IVs and specification choice

In order to identify nonlinear parameters, we use the "differentiation IVs" proposed by Gandhi and Houde (2019). The idea behind these instruments is to identify preference heterogeneity using the menu of choices available to different decision-makers. In our application, preference heterogeneity among voters for a candidate's wealth (say) is identified based on how many candidates in a voter's choice set have similar wealth.[32]

To guide our specification choice and evaluate the strength of our instruments, we use the methods proposed by Gandhi and Houde (2019). The details are in Appendix 2.1. The idea is to evaluate whether the instruments are "strong enough" to reject the linear (Logit) specification, i.e., $\mathbf{\Pi} = \mathbf{0}$. First, we enter the differentiation IVs as controls in a Logit specification. We find that the differentiation IVs for Muslim and crime are statistically significant while the differentiation IVs for assets and education are not. This suggests that the former two are capable of capturing departures from the Logit model. As an alternative diagnostic, we also run a specification that includes the differentiation IVs as instruments

[30] The effective number of parties, a widely used measure of concentration in political markets, is $1/\sum_p(S_p^2)$ where S_p is the vote share of party p.

[31] The only exception is one of the IVs for assets, where the correlation was small and insignificant. We therefore decided not to use this IV in the analysis that follows.

[32] To construct the differentiation IVs, we first predict each endogenous candidate characteristic using the instruments from the previous section. Then, for each characteristic of a candidate that enters the nonlinear part of voter utility, we compute the number of candidates in the constituency whose corresponding (predicted) characteristic is within one standard deviation. We cannot use differentiation IVs for incumbents because there can be at most one incumbent candidate in a constituency. We therefore focus on nonlinear parameters on the other four candidate characteristics.

instead of controls. The overidentification J-test clearly rejects this specification, which also provides support for focusing on the nonlinear specifications (Gandhi and Houde 2019).

Based on these specification checks, we focus on identifying the sources of preference heterogeneity in voters' valuation of Muslim and Crime. Our main demographic variables are literacy, rural population, presence of paved roads, employed population, and SC/ST population. We interact the differentiation IV for Muslim and Crime with the average value of each of these variables in the constituency, and again evaluate these instruments using Logit specifications. This supports using the interaction of the Crime differentiation IV with the share of SC/ST population (Table A.6 of the Appendix). For Muslim, no clear guidance emerges, we therefore estimate specifications that add each nonlinear term (and corresponding instruments) one at a time in Table 1. Column 1 is a clear favorite: this specification passes the overidentification J test, and results in nonlinear coefficients that are jointly statistically significant based on the Newey-West test.[33]

5.2.4 Voter preference estimates

According to the estimates in column 1 of Table 1, urban voters have a preference for non-Muslim candidates. This is consistent with animosity towards Muslims historically being more pronounced in urban areas than rural areas (Farooqui 2020). Voters also value wealth, but only among non-incumbents and not among incumbents. As a result, while we estimate a preference for most incumbents, the wealthiest incumbents are disliked (the marginal utility of Incumbent becomes negative at the 80th percentile of the asset distribution among incumbents). Finally, SC/ST voters dislike criminal politicians. This is in line with Vaishnav (2017)'s observation that there are fewer criminal politicians in reserved constituencies. Our estimates indicate that heterogeneity in preferences between SC/ST and non-SC/ST voters is an important source of this aggregate pattern.[34]

[33] In Appendix Table A.7 we experiment with adding different combinations of Normally distributed random coefficients to this specification and find that these are always statistically insignificant.

[34] Vaishnav (2017) argues that some voters value criminals because of their ability to "get things done." Under this interpretation, our estimates may indicate that SC/ST voters benefit less from these services.

Table 1: Voter preference parameter estimates

	(1)	(2)	(3)	(4)	(5)
Incumbent	5.32	4.40	4.49	3.88	3.42
	(1.80)	(1.44)	(1.37)	(2.19)	(1.30)
Assets	2.85	2.60	2.44	2.69	2.77
	(0.90)	(0.73)	(2.11)	(0.85)	(0.61)
Incumbent × Assets	-4.06	-3.61	-3.42	-3.42	-3.17
	(1.69)	(1.38)	(3.77)	(1.38)	(1.34)
Education	-1.56	-1.51	-1.23	-1.45	-1.57
	(1.09)	(0.91)	(2.07)	(0.78)	(0.76)
Muslim	-10.82	-0.51	0.10	-0.52	-0.67
	(6.22)	(5.51)	(5.18)	(4.00)	(0.68)
Crime	0.01	0.40	0.38	0.43	0.42
	(0.67)	(0.50)	(0.46)	(0.48)	(0.46)
Nonlinear parameters:					
Crime × SC/ST	-4.44	-4.56	-4.47	-3.50	-2.87
	(2.49)	(2.44)	(2.07)	(2.49)	(2.77)
Muslim × Rural	10.88				
	(6.27)				
Muslim × Literacy		0.18			
		(8.84)			
Muslim × Roads			-0.69		
			(8.14)		
Muslim × Work				0.44	
				(9.23)	
Muslim × SC/ST					1.13
					(2.09)
J	4.67	9.60	14.21	9.86	8.97
p-value	0.46	0.09	0.01	0.08	0.11
Newey-West p-value	0.07	0.21	0.16	0.36	0.39

Notes: BLP estimates. Specifications include broadcast allowance and fixed effects for state, year, party, alliance, imputed characteristics, and reserved constituencies. J is the overidentification J-statistic (df=5) with its p-value. The bottom row shows the p-value of the Newey-West D-test for the null that all nonlinear parameters are jointly 0. Robust standard errors clustered by constituency in parentheses. N = 2,649.

Figure A.1 in the Appendix shows the strength of voter preferences for each candidate characteristic, measured by the change in indirect utility from varying that characteristic. In Appendix 2.2, we quantify the bias that would result from failures of the exogeneity of our instruments, using the sensitivity measure proposed by Andrews et al. (2017). In each case find that the bias in our parameter estimates would be very small.

5.3 Specifying the parties' choice sets

5.3.1 Clustering

In our model, voters view candidates a_p as a bundle of characteristics $\mathbf{x}_p$. Applied directly to parties' problem, this would imply very large choice sets $\mathcal{A}_p$, containing all the possible combinations of characteristics. This is neither practical for estimation, nor realistic as a model of party choices (e.g., it is unlikely that a party views two otherwise identical candidates with slightly different assets as substantively different options).

For a more conceptually appealing model of parties' problem, we assume the existence of a smaller set of candidate "types" that parties consider when choosing who to run. For example, a type could be an "educated non-Muslim with a criminal history in the second quartile of the asset distribution." Non-incumbent parties choose among these types, while incumbent parties have the additional option to retain the incumbent. Rather than making ad hoc assumptions about what the relevant candidate types are, we use a data driven clustering approach to identify them.

We define candidate types based on the characteristics of the candidates running in state elections. Thus, we use information from the pool of potential national candidates as opposed to the set of candidates actually selected by the parties. There are many state candidates (36,766 in our data), which yields rich variation in candidate characteristics in the pool of potential candidates. As above, we use the candidate characteristics assets, education, Muslim, and criminal history.[35]

We construct candidate types using k-means clustering, an unsupervised iterative algorithm. In our application, the algorithm assigns each candidate to the nearest cluster centroid, updates the centroids to be the mean of the characteristics of the assigned candidates, and iterates until the clusters stabilize. The resulting clusters are our candidate types.

The number of clusters, K, is the key choice when applying k-means. In Appendix 4 we show that two standard measures from the machine learning literature for choosing K,

[35] Of these, assets, education, and criminal history have missing values. In order to assign each candidate to a type, we impute these missing characteristics as described in Appendix 1. We standardize all variables to have 0 mean and unit standard deviation.

the Within Cluster Sum of Squares (which measures the similarity of units within the same cluster) and the Silhouette Coefficient (which measures how far apart the clusters are from one another), strongly support $K = 4$ types of candidates in our application. Intuitively, this value results in densely packed clusters that are also far from each-other. Figure 4 in the Appendix also shows that k-means is superior to the most commonly used alternative, Hierarchical Clustering Analysis in our application.[36]

5.3.2 Candidate types

Table 2 shows the centroids of the four candidate types resulting from the clustering algorithm. The types turn out to have clear interpretations. Types 1 and 2 are both non-Muslim candidates with no criminal history, but Type 1 has high education while Type 2 has low education. Type 3 contains all the Muslim candidates. Type 4 contains all the non-Muslim candidates with criminal history, and these candidates are also wealthier than the other three types. For simplicity, we will refer to the four types as *educated*, *uneducated*, *Muslim*, and *criminal* types, respectively. These candidate types appear quite sensible. Given the salience of the Muslim characteristic in Indian politics, it is not surprising that one of the types we obtain is defined by this characteristic. The positive correlation between criminality and wealth, reflected in our criminal type, is consistent with Vaishnav (2017) and Asher and Novosad (2023). In sum, the types resulting from the clustering reflect meaningful choices for the parties, and not simply an artificial combination of candidate characteristics.[37]

Table 2: Centroids of candidate types

	Assets	Crime	Education	Muslim
Type 1	14.35	0.00	1.00	0.00
Type 2	14.27	0.00	0.00	0.00
Type 3	14.30	0.17	0.44	1.00
Type 4	15.06	1.00	0.68	0.00

Notes: Centroids resulting from k-means clustering. The algorithm is run on standardized variables; the table shows values transformed back to the original scale for ease of interpretation. Assets: real Rp in logs; Crime: 1 if has at least one criminal case; Education: 1 if completed high school; Muslim: 1 if Muslim name.

Table A.8 in the Appendix shows the distribution of candidate types in the data. Most UPA and NDA candidates are either educated or criminal types. Compared to the UPA,

[36] Trebbi and Weese (2019) develop an interesting alternative method to identify the number of organized insurgent groups using correlations in the timing of attacks over space. It is not obvious how to apply their method in our setting however, as identification of the number of clusters would require covariates that vary independently across our candidate characteristics, which we do not have.

[37] In Appendix 4 we ask how robust the types are to using the other clustering algorithm, HCA, mentioned above. Remarkably, we find that the two algorithms perfectly agree on the Muslim and criminal types, and only differ slightly on the other two.

the NDA has a somewhat higher share of Uneducated types and (not surprisingly) a smaller share of Muslim types.

5.3.3 Types in the national elections

We assign national candidates to the candidate types created from the state elections data based on their characteristics.[38] Compared to state candidates, national elections have relatively more candidates of the educated type (Appendix Table A.8). There are also more of the criminal type, but this difference between state and national elections is less pronounced for candidates of the UPA and the NDA. As in the state elections, the most pronounced difference between the NDA and the UPA is the former's lower share of the Muslim type and, to a lesser extent, its higher share of the uneducated type.

Table 3: Electoral performance and frequency of different candidates in the raw data

	Winner	Closeness to winner	Vote share	Choice share
Incumbent not available ($N = 612$)				
Type 1	0.29	0.61	0.17	0.53
Type 2	0.29	0.57	0.16	0.07
Type 3	0.17	0.56	0.17	0.08
Type 4	0.29	0.64	0.17	0.31
Incumbent available ($N = 256$)				
Type 1	0.35	0.80	0.23	0.13
Type 2	0.29	0.76	0.22	0.03
Type 3	0.50	0.93	0.29	0.02
Type 4	0.50	0.83	0.24	0.10
Incumbent	0.42	0.78	0.23	0.72

Notes: Average values in the data for UPA and NDA candidates. Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. *Winner* is an indicator for candidates who won. *Closeness to winner* is the candidate's vote share divided by the winner's vote share. *Vote share* is the vote share observed in the data. *Choice share* is relative frequency observed in the data. The incumbent is not an available choice either if the party did not win the previous election, or if the incumbent cannot run again for exogenous reasons (e.g., death).

Using these candidate types, Table 3 presents information on parties' choices and resulting electoral performance in the raw data. When the incumbent is not an available option, parties choose a new candidate of Types 1-4. When the incumbent is available, parties choose a new candidate of Types 1-4 or choose the incumbent. The first two columns are measures of winning probability: "winner" is an indicator equal to 1 if the candidate won,

[38] Based on Table 2, we assign educated, non-Muslim, non-criminal candidates to Type 1, uneducated, non-Muslim, non-criminal candidates to Type 2, Muslim candidates to Type 3, and all the non-Muslim criminal candidates to Type 4. We obtain a very similar assignment if we instead assign each candidate to the type with the closest centroid: only 3 UPA or NDA candidates are assigned to a different type in this case.

and "closeness to winner" is the candidate's vote share divided by the winner's vote share. These values suggest that different choices imply different winning probabilities. In spite of these differences, as shown in the third column, the vote shares associated with different choices tend to be very similar on average. A possible explanation is that parties do not attempt to increase their vote shares beyond the minimum necessary to win, in line with the discussion in Section 2. The choice frequencies, shown in the last column, indicate a very different pattern from either win probabilities or vote shares. This is suggestive of the fact that parties' choices are driven by other considerations.

5.4 Specification, identification, and estimation of party objectives

We focus on candidate selection by the UPA and the NDA. Other parties are included in the model as non-strategic players, with their candidates' characteristics fixed at their values observed in the data. Each party can either re-nominate an incumbent (if the party has one available) or choose a new candidate from one of the 4 types described above.[39] Based on (2), we specify party p's objective function in constituency c as

$$U_p(a_{pc}, P(\mathbf{a}_{-p,c})) = \sum_{\mathbf{a}_{-p,c}} \left[b^w w_{pc}(a_{pc}, \mathbf{a}_{-p,c}) + b^s s_{pc}(a_{pc}, \mathbf{a}_{-p,c})\right] P(\mathbf{a}_{-p,c}) - \sum_{k=1}^{4} (c^0_{kp} + \mathbf{c}_k \mathbf{L}_{pc}) \mathbf{1}\{a_{pc} = k\} + \varepsilon_{pc}(a_{pc}), \tag{11}$$

where $\mathbf{a}_{-p,c}$ includes all other parties' choices, $\mathbf{L}_{pc}$ contains the proxies for the pool of candidate characteristics (the average of candidate education, assets, Muslim and criminal history in the state election constituencies corresponding to constituency c), and c^0_{kp} and $\mathbf{c}_k$ are cost parameters associated with type k. As is typical in discrete choice models, we require a normalization. The formulation in (11) sets the cost of selecting the incumbent to 0, the payoff parameters c^0_{kp} and $\mathbf{c}_k$ may thus be viewed as the *replacement costs* of choosing a type k candidate instead of the incumbent. We will refer to the costs that depend on the candidate pool, $\mathbf{c}_k$, as *recruitment costs* and the costs c^0_{kp} as *direct costs*.[40] Our goal is to estimate the parameters $\boldsymbol{\theta} = (b^w, b^s, \{c^0_{kp}, \mathbf{c}_k\}_{k=1}^4)$.

To compute the vote shares and winning probabilities corresponding to different hypothetical candidate choices by the parties, we use our demand estimates. We compute the vote

[39] Unless otherwise stated, in what follows when we refer to a Type 1-4 candidate we mean a non-incumbent candidate of these types.

[40] In principle, direct costs could also depend on the characteristics of the incumbent. Allowing for incumbent-type-specific costs results in similar estimates to those reported below.

shares and associated winning probabilities for all possible action profiles $(a_{pc}, \mathbf{a}_{-p,c})$ of the players in a given constituency, holding fixed all exogenous variables (including non-strategic parties' choices) as well as the popularity shocks ξ_{pc}.

To use our demand estimates, we also need to assign a vector of characteristics to each action. As described in Section 4.2, when parties make their choices, there is some uncertainty over the actual realization of characteristics. For non-incumbents, we capture this uncertainty using the distribution of candidate characteristics for each type in the state election data. For incumbents, we use the distribution of current characteristics conditional on past characteristics.[41] For each combination of actions, we draw 1000 values of characteristics, and compute vote shares and winning probabilities as expected values across these draws (equations (7) and (8), respectively).

We provide a formal discussion of the identification of the party objective function parameters in Appendix 3, but it is worth briefly discussing the sources of identification here. Cost parameters are identified up to differences with respect to a reference alternative. The identification of these differences, for given parameters (b^w, b^s), is standard and depends on the magnitude of the observed probability of selecting type k, $P(a_{pc} = k)$, relative to the probability of selecting the reference type K. By contrast, the benefit parameters are pinned down in levels, not differences. To see this, note that given the Type-I Extreme Value assumption we can write $\ln(P(a_{pc} = k)) - \ln(P(a_{pc} = K)) = b^w(E[w_{pc}(k, \mathbf{a}_{-p,c})] - E[w_{pc}(K, \mathbf{a}_{-p,c})]) + b^s(E[s_{pc}(k, \mathbf{a}_{-p,c})] - E[s_{pc}(K, \mathbf{a}_{-p,c})]) - (c_k - c_K) + \mu_{pc}(k)$, where we have assumed a single type-specific cost term c_k for simplicity. As we can treat choice probabilities and win probabilities as known, this can be viewed as a regression of $\ln(P(a_{pc} = k)) - \ln(P(a_{pc} = K))$ on $(E[w_{pc}(k, \mathbf{a}_{-p,c})] - E[w_{pc}(K, \mathbf{a}_{-p,c})])$ and $(E[s_{pc}(k, \mathbf{a}_{-p,c})] - E[s_{pc}(K, \mathbf{a}_{-p,c})])$ where the intercept is the cost difference $c_k - c_K$. For instance, b^w is, loosely, identified as the covariance across constituencies between the probability of selecting type k relative to the reference option K, and the difference in expected win probabilities between k and K. If the parties tend to select candidate k in the constituencies where they are likely to win, b^w will be positive.

Estimation proceeds by recursively updating the choice probabilities using (pseudo) maximum likelihood estimates of the parameter vector $\boldsymbol{\theta}$ up to convergence as in Aguirregabiria and Mira (2007). Specifically, consider an initial choice probability estimate $\hat{P}^0(\mathbf{a}_{-p,c})$. In constituency c, party p chooses a_{pc} to maximize (11). Again defining utility net of the

[41] We assume that Muslim and education do not change over time. For criminality and assets, we regress each of these on past characteristics, and draw from the estimated error distribution.

unobservable as $\tilde{U}_p(a, P(\mathbf{a}_{-p,c}); \boldsymbol{\theta})$, the implied probability that $a_{pc} = a$ is:

$$P(a|\hat{\mathbf{P}}^0, \boldsymbol{\theta}) = \frac{\exp\left\{\tilde{U}_p(a, \hat{\mathbf{P}}^0(\mathbf{a}_{-p,c}), \boldsymbol{\theta})\right\}}{\sum_{a'} \exp\left\{\tilde{U}_p(a', \hat{\mathbf{P}}^0(\mathbf{a}_{-p,c}), \boldsymbol{\theta})\right\}} \tag{12}$$

where we emphasize the fact that the choice probabilities are a function of the estimates $\hat{\mathbf{P}}^0$ as well as the parameters $\boldsymbol{\theta}$. Denoting observed party choices by a^*_{pc}, the estimates $\hat{\boldsymbol{\theta}}^0$ maximize the log likelihood

$$\ell(\boldsymbol{\theta}, \hat{\mathbf{P}}^0) = \sum_{pc} \sum_{a} \mathbf{1}\{a = a^*_{pc}\} \ln P(a|\hat{\mathbf{P}}^0, \boldsymbol{\theta}).$$

With these estimates, we can construct updated estimates of the choice probabilities, $\hat{\mathbf{P}}^1$, by substituting $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^0$ into the right hand side of Equation (12). This yields an updated likelihood and new parameter estimates $\hat{\boldsymbol{\theta}}^1$. We continue in this fashion until $\hat{\boldsymbol{\theta}}^{h+1} \simeq \hat{\boldsymbol{\theta}}^h$. This yields the estimator, $\hat{\boldsymbol{\theta}}_{NPL}$, which is consistent for $\boldsymbol{\theta}$ regardless of the initial guess $\hat{\mathbf{P}}^0$.[42] The asymptotic properties of $\hat{\boldsymbol{\theta}}_{NPL}$ are attractive as recursive updating improves efficiency of the estimates by imposing equilibrium conditions on the data (Aguirregabiria and Mira 2007). Standard errors are obtained from the inverse Hessian of the likelihood function evaluated at $\hat{\boldsymbol{\theta}}_{NPL}$.

6 Estimation results

6.1 What do parties maximize?

In Table 4, we first estimate a version of the model where parties care only about their expected winning probability and vote share. We then introduce the recruitment costs $\mathbf{c}_k \mathbf{L}_{pc}$ in column 2, and also add the direct costs c^0_{kp} in column 3. Allowing for these costs substantially improves the model's fit: as we move from column 1 to 3, the log likelihood increases by 33.2%. Costs, which are independent of voter preferences, are important in explaining parties' choices of which candidates to run.[43]

According to our estimates, parties like to win ($b^w > 0$), but, all else equal, choosing candidates with higher vote shares has a disutility ($b^s < 0$). This is consistent with the idea

[42] The procedure need not be initialized with a consistent estimate of choice probabilities, a typical requirement in other computationally efficient approaches. We initialize the NPL procedure with uniform probability across the possible choices: $\widetilde{P}^0(a_{pc}) = 1/5$ for all 5 possible choices of a_{pc} for each p, c.

[43] Likelihood ratio tests for the joint significance of costs yield p-values below 0.01 in all specifications we report. See Section 6.3 below for a more detailed evaluation of model fit.

that party elites balance anticipated electoral performance of a popular candidate against the possibility that the candidate becomes too powerful and threatens the elite, as discussed in Section 2.

Table 4: Party objective function estimates

	(1)	(2)	(3)		(1)	(2)	(3)
b^w	7.46	4.89	4.39	c_4^{educ}		2.30	0.37
	(0.68)	(0.81)	(0.83)			(0.38)	(0.44)
b^s	-10.90	-10.02	-11.09	c_1^{crime}		1.64	0.90
	(1.03)	(1.17)	(1.21)			(0.30)	(0.30)
$c_{1,NDA}^0$			3.00	c_2^{crime}		1.70	0.73
			(0.53)			(0.35)	(0.38)
$c_{2,NDA}^0$			3.55	c_3^{crime}		1.61	0.42
			(0.52)			(0.34)	(0.37)
$c_{3,NDA}^0$			4.90	c_4^{crime}		0.89	-0.16
			(0.58)			(0.30)	(0.31)
$c_{4,NDA}^0$			4.17	c_1^{asset}		1.87	0.53
			(0.55)			(0.30)	(0.35)
$c_{1,UPA}^0$			2.24	c_2^{asset}		2.53	0.54
			(0.44)			(0.33)	(0.42)
$c_{2,UPA}^0$			3.55	c_3^{asset}		2.95	0.98
			(0.51)			(0.35)	(0.43)
$c_{3,UPA}^0$			3.44	c_4^{asset}		2.01	0.10
			(0.50)			(0.30)	(0.35)
$c_{4,UPA}^0$			3.35	c_1^{Muslim}		0.30	0.25
			(0.46)			(0.42)	(0.42)
c_1^{educ}		1.23	0.21	c_2^{Muslim}		0.06	0.24
		(0.38)	(0.43)			(0.52)	(0.58)
c_2^{educ}		2.69	0.68	c_3^{Muslim}		-1.29	-1.09
		(0.44)	(0.54)			(0.44)	(0.47)
c_3^{educ}		2.94	0.46	c_4^{Muslim}		-0.19	-0.00
		(0.43)	(0.54)			(0.43)	(0.45)
Log likelihood					-1188.4	-872.86	-794.86

Notes: Estimates of party objective functions in (11). Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. The cost parameters c are measured relative to the excluded category, incumbents. $c_{k,p}^0$ is party p's direct cost of choosing a type k candidate. c_k^l is the impact of characteristic l in the pool of candidates on the recruitment cost of a type k candidate. Standard errors in parentheses. Number of markets: 434.

To interpret the magnitude of the estimates, we use the standard deviation of parties' equilibrium payoffs, which is 1.56. All else equal, winning raises parties' payoff by 2.8 standard deviations (4.39/1.56). Winning with a one standard deviation higher vote share (8 percentage points in the data) lowers payoffs by about 0.6 standard deviations.

The recruitment cost parameters $\mathbf{c}_k$ show that a higher prevalence of some candidate characteristic in a party's candidate pool makes candidate types with that characteristic relatively less costly for the party. More educated candidates in the pool make the educated Type 1 relatively less costly (c_1^{educ} is the lowest among types), more criminals and wealthier candidates in the pool make the criminal Type 4 relatively less costly (c_4^{crime} and c_4^{asset} are the lowest), and more Muslim candidates in the pool make the Muslim Type 3 relatively less costly (c_3^{Muslim} is the lowest). The supply of candidate characteristics available to parties affects who they choose to run, mirroring the patterns seen in Table A.4.

Column 3 shows that parties' costs of running specific candidates are not restricted to recruitment costs. All direct cost estimates are positive, showing that both parties have a direct preference for running an incumbent candidate. This may reflect factors such as connections and experience that benefit the party beyond their effect on votes. The NDA has the largest direct cost from running a Muslim type (Type 3): running such a candidate lowers its payoff by 3.1 standard deviations, all else equal. This is in line with the Hindu nationalist profile of the BJP, the NDA's leading party, and indicates that the NDA's aversion to running Muslim types is not due simply to the preferences of its voters.

Figure 1 plots the recruitment costs for each type, and compares them to the direct costs (indicated with vertical lines). Both recruitment costs and direct costs matter, though recruitment costs tend to be smaller than direct costs. This is especially the case for the criminal Type 4, which has the lowest recruitment cost. In most constituencies, parties incur relatively low costs from selecting criminal types compared to other types due to the ample supply of criminality in the candidate pool.

6.2 Implications: party organization and party choices

6.2.1 Parties' aversion to candidates who win big

The finding of a disutility from large vote shares in parties' objective function is consistent with the idea that relatively successful candidates are more likely to be disloyal to their party. Although candidates who barely lose and candidates who win big can both use the threat of disloyalty to formulate demands on the party, there is an important asymmetry between the two. By law, winners who switch parties trigger a by-election, putting the party's legislative seat in jeopardy. This suggests that parties' aversion to successful candidates should be

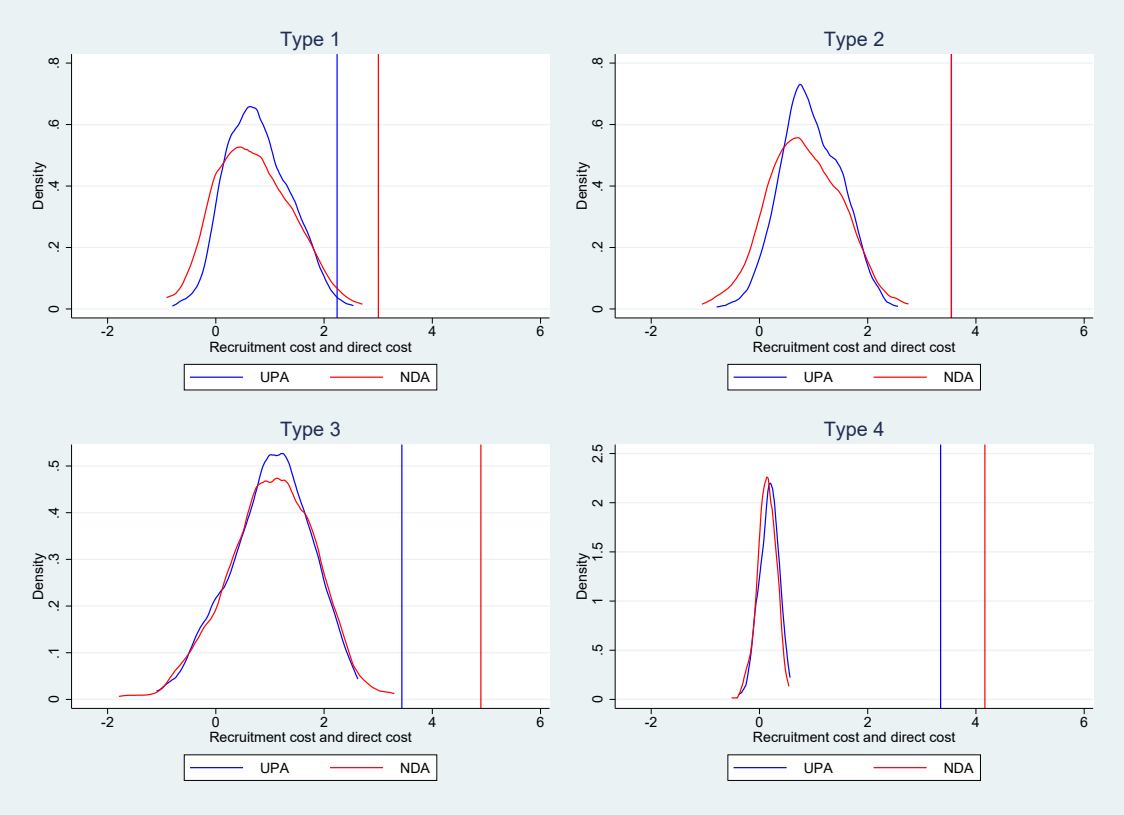


Figure 1: Distribution of each party's costs of different types across constituencies

Based on column 3 of Table 4. Kernel density plots of recruitment costs $\mathbf{c}_k\mathbf{L}_{pc}$; vertical lines are direct costs c^0_{kp}. Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. Costs are measured relative to incumbents.

larger if these candidates are more likely to win.

To study this, in Appendix Table A.13 we estimate a version of our model where we interact expected win probability and vote share in parties' objective function. We find that parties' disutility from large vote shares is decreasing in their expected win probability. Parties are averse especially to candidates who are expected to win big. This is consistent with the interpretation that $b^s < 0$ reflects a party's cost from running candidates who can then credibly threaten the leadership with disloyalty.

6.2.2 Why do parties choose criminal candidates?

Our estimates reveal a set of reasons that guide parties' choice of criminal candidates, formalizing previous explanations in the literature, and adding some new considerations.

First, since parties like to win, they have an incentive to choose candidates who appeal to voters all else equal. According to our demand estimates in Table 1, criminal candidates are indeed appealing to voters, though this is driven by their wealth, rather than their criminality. The importance of money in the valuation of criminal politicians in India echoes the discussion in Vaishnav (2017).

Second, however, the fact that criminal types can generate large vote shares also creates a disutility for party leaders given $b^s < 0$. This reduces their incentive to run such a candidate compared to some of the other types.

Third, criminal types also affect party payoffs independently of votes. We found that in most constituencies recruitment costs favor this candidate type (Figure 1): the large supply of criminal candidates makes it relatively cheap to run them.

Fourth, in our model a party's choice is a strategic response to the candidate chosen by its opponent. A party could run a criminal simply because its opponent is doing that as well: criminals may be strategic complements.[44] In Appendix 6, we compute $\frac{\partial P_p(4)}{\partial P_{-p}(4)}$, the change in party p's probability of choosing a criminal in response to a change in its opponent's probability of choosing a criminal. A positive derivative for both parties indicates strategic complementarity. Among the constituencies where the incumbent is not an option, we find that criminals are strategic complements in 84% of the cases.

[44] Intuitively, when a party's opponent is running a Type 4 candidate, running a different type would yield a low probability of winning, while running a Type 4 increases the probability of winning without generating vote shares that are too large.

6.2.3 Understanding "anti-incumbency"

Developing country politicians in office are less likely to run for re-election than their US counterparts, and those who do run are not as successful (e.g., De Magalhaes (2015)).[45] These patterns are often described as "anti-incumbency." Close-election Regression Discontinuity (RD) studies confirm that incumbents face a significant electoral disadvantage relative to their opponent in India (Uppal 2009; Lee 2020).

Incumbents' electoral fortunes are equilibrium outcomes. What looks like anti-incumbency by voters could in fact reflect parties' choice to run incumbents who are relatively less popular.[46] In what sense are voters and / or parties pro or anti-incumbent? Our findings provide several insights into this question.

First, our estimates reveal that voters' valuation of incumbency is heterogeneous across candidates. Voters value incumbency negatively for the wealthiest candidates but positively for everyone else (Table 1).

Second, because party leaders experience disutility from candidates who are expected to be overly successful (Table 4), parties and voters may disagree on which incumbents to replace. We find that this can explain most cases where incumbents are replaced: if we remove the disutility from high vote shares (by setting $b^s = 0$), 97% of incumbents are retained by the parties, compared with 72% in the baseline equilibrium.

Third, in most constituencies we estimate positive direct costs for parties from replacing the incumbent. This shows that parties benefit from incumbents, perhaps through their experience or connections, independently of votes.

These observations highlight the importance of considering endogenous choices by parties when comparing incumbents and non-incumbents who run for election.

6.3 Robustness and model fit

In Appendix 7 we present a series of robustness checks on our estimates. We replace the vote share terms $b^s s(\cdot)$ in parties' objective function with terms allowing for further nonlinearities, with vote margins, or with the number of votes. We find that these changes make little difference to our results. We show that controlling for several sources of observable constituency level heterogeneity, including election year, reserved constituencies, or the share of rural population, does not change our estimates in a material way. We also apply the method of Bonhomme et al. (2022) to control for *unobservable* constituency level heterogeneity, and show that this also leaves our main estimates unaffected.

[45] In our data, 67% of incumbents run again and 49% of these win - in the US House of Representatives, both of these shares are around 95%.

[46] See De Magalhaes (2015) for a careful discussion of this point in the context of RD studies.

To explore the model's ability to fit the data, we first use the estimated model to simulate party choices in every constituency. We draw actions using the equilibrium choice probabilities 100 times, and compare the average over simulations with the actual choices observed in the data for each type and each party. The result is in Panel A of Table 5. The model performs well.

Table 5: Model fit

	UPA actual	UPA predicted	NDA actual	NDA predicted	All actual	All predicted
Panel A: Actual and predicted choices						
Type 1	167	166.73	191	191.92	358	358.65
Type 2	17	16.57	35	34.37	52	50.94
Type 3	38	38.24	18	17.56	56	55.80
Type 4	111	111.58	106	106.28	217	217.86
Incumbent	101	100.88	84	83.87	185	184.75
Panel B: 5-fold cross validation						
Type 1	31.60	34.60	36.20	35.84	67.80	70.44
Type 2	3.40	2.96	6.80	6.76	10.20	9.72
Type 3	7.00	7.28	3.60	2.76	10.60	10.04
Type 4	21.40	19.48	20.20	20.64	41.60	40.12
Incumbent	19.60	18.68	16.20	17.00	35.80	35.68

Notes: Panel A shows the number of candidates of each type observed in the data and predicted by the model (average across 100 simulations). Panel B uses a 5-fold cross validation procedure as described in the text (values shown are averages across 20% subsamples). Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively.

To evaluate the model's "out of sample" performance, we use a cross-validation procedure. We repeat the following 5 times. First, we hold out the first 20% of the sample, and estimate the model using the remaining 80% of observations. We then solve for the equilibrium in each constituency in the 20% hold-out sample, and use this to simulate party choices and evaluate the model's fit on that sample. We repeat this process for the next 20-80 split, etc., and take the average across the 5 sets of predictions (this is essentially a k-fold cross validation, with $k = 5$). The results are in Panel B of Table 5. While there is variation across the folds in predictive ability, on average over the k-folds the model does just as well in predicting outcomes as in Panel A.

7 Ideology

Ideology is an important consideration for both the demand and supply side of politics in most party systems (Iaryczower et al. 2024; Longuet-Marx 2025). In India, observers have historically downplayed the role of ideology, while a recent literature argues that ideologies

matter but do not fit on the left-right spectrum familiar in Western party systems.[47]

So far in our specification of the supply side, ideological differences between parties are held fixed based on the idea that individual candidates will largely toe the party line. On the demand side, these ideological differences are captured by the party fixed effects.[48] This raises two questions: Do voters care about ideology, and if yes, is there substantial heterogeneity in these preferences along the demographics we study?

We investigate these questions using two complementary approaches to measuring ideology. (The full discussion of the data creation and sources are in Appendix 1.4.) First, we create two measures using Wikipedia. The first places parties on a Left-Right scale, and the second creates indicators for various keywords that describe parties' ideology. Both of these rely on consistent headers across parties in the standardized structure of Wikipedia entries. If a party is in the NDA (respectively, UPA) alliance, we take its candidates' ideology to be a combination of the party's and the BJP's (respectively, INC's) ideology. Because there is variation in alliances over time and across constituencies, this creates a measure of candidate ideology that is not captured by the party fixed effects in our main specification.

Our second approach to measuring ideology is based on expert surveys from the Democratic Accountability and Linkages Project. These score parties on various dimensions, and we create two alternative measures based on them. First, we use the first principal component of the different scores (which we find captures over 80% of the variation in the data). Second, we use two individual scores, one measuring recognition of disadvantaged groups and the other the role of the state in regulating social norms and redistribution (statism). Chhibber and Verma (2018) argue that these are the two dimensions that define contemporary political ideology in India.

We first investigate whether Indian voters care about ideology by including the above measures as an additional (exogenous) candidate characteristic in our main demand specification. The results in Table 6 provide some evidence that voters care about ideology. Voters exhibit a preference for more right-wing candidates in columns 1 and 3, and for candidates rated low on Recognition in column 4. Importantly, the inclusion of ideology, regardless of the measure used, has little impact on the estimates of the other demand parameters. This suggests that the variables included in our main specification already capture much of the variation in voter preferences that is due to these ideological differences.

Next we check for heterogeneous voter preferences for ideology using the Left-Right ideology measure, the DALP principal component, and the two alliance fixed effects. To identify

[47] See Chhibber and Verma (2018) who challenge "the consensus that Indian political parties are nonideological" (p9).

[48] Because of the large number of parties, these fixed effects are likely to capture much more fine-grained ideological differences between candidates that would be the case in a two-party system like the US.

Table 6: Voter preference parameter estimates with different ideology measures

	Left-Right	Terms	DALP	CV
Incumbent	5.02	4.75	5.13	5.42
	(1.86)	(1.87)	(1.81)	(1.94)
Assets	3.14	3.26	3.06	3.14
	(0.91)	(0.88)	(0.89)	(0.91)
Incumbent × Assets	-4.45	-4.56	-4.45	-4.71
	(1.74)	(1.68)	(1.74)	(1.84)
Education	-1.38	-1.72	-1.37	-1.58
	(1.13)	(1.07)	(1.15)	(1.16)
Muslim	-11.93	-12.77	-11.59	-11.86
	(6.85)	(7.52)	(6.68)	(6.83)
Crime	0.22	0.12	0.24	0.13
	(0.69)	(0.70)	(0.68)	(0.69)
Ideology	2.23		0.68	3.14
	(1.53)		(0.27)	(1.81)
Ideology 2				-0.52
				(0.60)
Nonlinear parameters:				
Muslim × Rural	11.97	12.78	11.63	11.91
	(6.89)	(7.50)	(6.72)	(6.87)
Crime × SC/ST	-4.91	-4.16	-5.14	-4.77
	(2.48)	(2.78)	(2.43)	(2.52)
J	3.88	(2.75)	4.03	3.31
p-value	0.57	0.74	0.55	0.65
Newey-West pval	0.07	0.13	0.05	0.08

Notes: BLP estimates. Ideology measures are: Left–Right (Wikipedia-based scale), Terms (12 keyword indicators from party descriptions), DALP (expert survey from DALP), and CV (Chhibber–Verma dimensions). The latter includes two variables, Ideology (low recognition) and Ideology 2 (low statism). Parameter estimates for Terms are not shown.

each of these parameters, we create differentiation IVs similarly to the other characteristics.[49] Table A.9 and A.10 in the Appendix show some of these specifications.[50] We find that, in specifications that are not rejected by the J test, heterogeneity in preferences for ideology is insignificant, while the parameters in our main specification generally remain significant and relatively stable.

[49]For Left-Right and DALP, we count the number of candidates in the constituency with ideology within one standard deviation of the given candidate. For heterogeneity by voter demographics, we take the product of the differentiation IV with the corresponding average demographic in the constituency. To identify the pairs of random coefficients on the alliance fixed effects, we use the differentiation IV measure of Left-Right ideology and its square.

[50]For each measure, we estimated 6 specifications, with heterogeneity captured through Normal random coefficients, Literacy, Rural, Muslim, Roads, Employed, or SC/ST. We present the specifications with Normal random coefficients, Roads, and SC/ST for each of the 3 measures (Roads and SC/ST showed the strongest evidence for the possible presence of heterogeneous preferences).

To investigate how the inclusion of voter preference for ideology affects the supply estimates, we re-estimate the supply model using two of these demand specifications in which ideology appeared to matter: including Left-Right ideology in mean voter utility only (from column 1 of Table 6) or also allowing for heterogeneity by Roads (from column 2 of Appendix Table A.10). The results are reported in Appendix Table A.11. The main supply parameters remain qualitatively unchanged.

8 Policy experiments

8.1 Banning the criminal type

What is the impact of banning candidates with a criminal history from elections? To model such a policy, we consider a counterfactual scenario where we make it prohibitively costly to choose criminal type (Type 4) candidates - both incumbents and non-incumbents.[51] Using our parameter estimates, we compute a new equilibrium.

Table 7 compares average choice probabilities in the criminal ban counterfactual equilibrium and the baseline equilibrium. We find that the choice probabilities of candidate types 1-3 all increase.[52] The increase is largest for the educated type (Type 1), whose average choice probability increases by 24 percentage points. We also find that the share of re-nominated incumbents declines. This is not simply a result of the replacement of criminal incumbents. In fact, when a party has a non-criminal incumbent available, banning criminal candidates *increases* the choice probability of incumbents, reflecting both substitution away from criminal type candidates and equilibrium responses to changes in the opponent's strategy.

Computing the changes in parties' expected equilibrium payoffs, we find that on average, both the UPA's and the NDA's payoffs decrease when the criminal type is banned (Appendix Figure A.4). In 17% of the constituencies, both parties' payoffs decline. However, eliminating the criminal option can also increase both parties' payoffs: this is the case in 45% of the constituencies. This reflects the strategic complementarity discussed in Section 6.2.2. When a party chooses a criminal because its opponent chooses a criminal, banning criminals allows

[51] Although Type 3 also contains a small fraction of candidates with criminal history, Type 4 candidates *always* have a criminal history. They are also wealthier, and thus more closely match the kind of criminal candidates who are considered problematic in the Indian context (Vaishnav 2017). We also considered versions of this counterfactual where we removed third parties that run criminal candidates before imposing the criminal ban, or where we also removed criminal candidates from the state-election pool (i.e., when computing $\mathbf{L}_p$). In both cases the ban's impact on party choices was similar.

[52] Unlike in a single-agent model, here the increase in choice probabilities when one of the options is removed is not simply mechanical because parties play best responses to each-other's strategy.

parties to profitably choose different candidates. Even if these candidates lower a party's probability of winning, this can be offset by lower recruitment costs and a lower disutility from excessive popularity.

The increase in payoffs suggests that, in some cases, the main Indian parties may be willing to collectively support a ban on criminal candidates. This helps explain instances when these parties express support for a ban on criminal candidates while at the same time running such candidates in elections.

Table 7: Party choices with criminal ban

	All parties		No incumbent		Criminal incumbent		Non-criminal incumbent	
	Baseline	Ban	Baseline	Ban	Baseline	Ban	Baseline	Ban
Type 1	0.41	0.65	0.53	0.77	0.13	0.73	0.15	0.18
Type 2	0.06	0.10	0.08	0.11	0.02	0.16	0.03	0.03
Type 3	0.06	0.11	0.08	0.13	0.02	0.12	0.02	0.03
Type 4	0.25	0	0.32	0	0.10	0	0.09	0
Incumbent	0.21	0.15	0	0	0.73	0	0.72	0.76
N	868	868	612	612	87	87	169	169

Notes: Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. Values shown are the average choice probabilities across parties (for all parties, parties with no incumbent available, and parties with a criminal or a non-criminal incumbent available).

8.2 Term limits

What is the impact of term limits on political selection - in particular, on the characteristics of candidates chosen by the parties? Advocates argue that term limits have the potential to bring "a new breed" of politician into government (Carey et al. 2006), but the evidence is mixed. In the US, term limits in sate legislatures did not appear to significantly affect the characteristics of those elected (Carey et al. 2006), while in the Philippines term limits for mayors increased both the share of women running and winning (Labonne et al. 2021).[53]

In our data, incumbents running for re-election are more likely to have a criminal background than non-incumbents (37.8 vs. 31.8%) and less likely to be educated (47.0 vs. 52.4%). Suppose a term limit removes incumbents from parties' choice sets. Will the resulting distribution of candidate characteristics change, or will parties simply replace incumbents with new candidates who are similar to them?

To answer this question, we use our estimated model to simulate a counterfactual equilibrium where we remove incumbents from parties' choice sets.[54] The results are shown

[53]Most studies of term limits focus on the behavior of politicians. The question considered here regarding their impact on political selection is complementary.

[54]We also ran a version of this counterfactual where we removed third party incumbents from the data,

in Table 8. The first column shows the average choice probabilities at baseline. To make comparisons meaningful, the second column presents the baseline choice probabilities when incumbent candidates are also assigned one of the Types 1-4 (based on their last observed characteristics). The last column shows the counterfactual choice probabilities under term limits. Comparing the second and the third columns shows a small amount of substitution away from the educated type (-2.2 ppoints) and mostly towards the uneducated type (+1.7 ppoints). However, overall, the distribution of choice probabilities changes little. Parties tend to replace incumbents of each type by non-incumbents of the same type.

Based on these results, term limits would not meaningfully alter the distribution of candidate types in this context.

Table 8: Choice probabilities with and without term limits

	Baseline	Baseline with types for incumbents	Term limit
Type 1	0.34	0.52	0.50
Type 2	0.05	0.08	0.09
Type 3	0.05	0.07	0.08
Type 4	0.20	0.33	0.33
Incumbent	0.36	-	0

Notes: Column 1 shows average choice probabilities in the baseline equilibrium among constituencies where the incumbent is a possible choice. Column 2 shows these probabilities when incumbents are also assigned a type 1-4. Column 3 shows the average choice probabilities when incumbents are removed from the players' choice sets.

9 Conclusion

We estimate a model of candidate selection by political parties to study why parties in a representative democracy select the candidates they do. Our setting is India, where we combine a rich demand side specification of voter preferences with a supply side game between parties that incorporates direct payoffs from candidate selection.

We find that, while parties value candidates who are more likely to win, selection decisions are also shaped by other considerations. All else equal, parties would prefer to win with candidates who are not overly popular. This is consistent with the view that the major Indian parties are reluctant to select candidates who might become too powerful and threaten the party and its elites. In addition, we find that selection decisions are also shaped by considerations that are independent of voter preferences, including recruiting cost and other direct costs from replacing incumbents with specific candidate types.

and obtained very similar results.

Our estimates provide a detailed explanation for why parties run criminal candidates. They also provide insights into "anti-incumbency" in India, showing that much of this can be explained by parties' reluctance to renominate the most popular incumbents.

Our approach is well-suited to study the impact on political selection of counterfactual policies that create changes in parties' choice sets, and we illustrate this through analyses of criminal bans and term limits.

References

Aguirregabiria, V. and P. Mira (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75, 1–53.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *The Quarterly Journal of Economics*, 132, 1553–1592.

Asher, S., T. Lunt, R. Matsuura, and P. Novosad (2021): "Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in India using the SHRUG open data platform," *The World Bank Economic Review*, 35, 845–871.

Asher, S. and P. Novosad (2023): "Rent-Seeking and Criminal Politicians: Evidence from Mining Booms," *The Review of Economics and Statistics*, 105, 20–39.

Bandiera, O., A. Prat, S. Hansen, and R. Sadun (2020): "CEO behavior and firm performance," *Journal of Political Economy*, 128, 1325–1369.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, 63, 841–890.

Besley, T. (2005): "Political Selection," *Journal of Economic Perspectives*, 19, 43–60.

Besley, T., O. Folke, T. Persson, and J. Rickne (2017): "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden," *American Economic Review*, 107, 2204–2242.

Bhogale, S., S. Hangal, F. R. Jensenius, M. Kumar, C. Narayan, B. U. Nissa, and G. Verniers (2019): *TCPD Indian Elections Data v1*, Trivedi Centre for Political Data, Ashoka University.

Bonhomme, S., T. Lamadon, and E. Manresa (2022): "Discretizing unobserved heterogeneity," *Econometrica*, 90, 625–643.

Caillaud, B. and J. Tirole (2002): "Parties as Political Intermediaries," *The Quarterly Journal of Economics*, 117, 1453–1489.

Cain, B., J. Ferejohn, and M. Fiorina (1987): *The personal vote: Constituency service and electoral independence*, Harvard University Press.

Carey, J. M., R. G. Niemi, L. W. Powell, and G. F. Moncrief (2006): "The effects of term limits on state legislatures: a new survey of the 50 states," *Legislative Studies Quarterly*, 31, 105–134.

Casas-Arce, P. and A. Saiz (2015): "Women and power: unpopular, unwilling, or held back?" *Journal of political Economy*, 123, 641–669.

Casey, K., A. B. Kamara, and N. F. Meriggi (2021): "An experiment in candidate selection," *American Economic Review*, 111, 1575–1612.

Chandra, K. (2016): *Democratic dynasties: State, party, and family in contemporary Indian politics*, Cambridge University Press.

Chhibber, P., F. R. Jensenius, and P. Suryanarayan (2014): "Party organization and party proliferation in India," *Party Politics*, 20, 489–505.

Chhibber, P. K. and R. Verma (2018): *Ideology and identity: The changing party systems of India*, Oxford University Press.

Dal Bó, E. and F. Finan (2018): "Progress and perspectives in the study of political selection," *Annual Review of Economics*, 10, 541–575.

Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne (2017): "Who becomes a politician?" *The Quarterly Journal of Economics*, 132, 1877–1914.

De Magalhaes, L. (2015): "Incumbency effects in a comparative perspective: evidence from Brazilian mayoral elections." *Political Analysis*, 23, 113–26.

Fan, Y. (2013): "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market," *American Economic Review*, 103, 1598–1628.

Farooqui, A. (2020): "Political representation of a minority: Muslim representation in contemporary India," *India Review*, 19, 153–175.

Farooqui, A. and E. Sridharan (2014): "Incumbency, internal processes and renomination in Indian parties," *Commonwealth & Comparative Politics*, 52, 78–108.

Fisman, R., F. Schulz, and V. Vig (2016): "Financial disclosure and political selection: Evidence from India," *working paper*.

François, P., F. Trebbi, and K. Xiao (2023): "Factions in Nondemocracies: Theory and Evidence From the Chinese Communist Party," *Econometrica*, 91, 565–603.

Fudenberg, D. and J. Tirole (1991): *Game theory*, MIT press.

Galasso, V. and T. Nannicini (2011): "Competing on good politicians," *American Political Science Review*, 105, 79–99.

Gandhi, A. and J.-F. Houde (2019): "Measuring substitution patterns in differentiated-products industries," *NBER Working paper w26375*.

Genovese, M. A. (1993): *Women as National Leaders*, Sage.

Gulzar, S., Z. Hai, and B. K. Paudel (2021): "Information, candidate selection, and the quality of representation: Evidence from Nepal," *The Journal of Politics*, 83, 1511–1528.

Hamilton, B. H., A. Hincapié, R. A. Miller, and N. W. Papageorge (2021): "Innovation and Diffusion of Medical Treatment," *International Economic Review*, 62, 953–1009.

Hasan, Z. (2010): "Political Parties," in *The Oxford companion to politics in India*, ed. by N. G. Jayal and P. B. Mehta, Oxford University Press Delhi.

Heath, O. and A. Ziegfeld (2022): "Why So Little Strategic Voting in India?" *American Political Science Review*, 116, 1523–1529.

Iaryczower, M., G. Kim, and S. Montero (2024): "Representation Failure," *working paper*.

Labonne, J., S. Parsa, and P. Querubin (2021): "Political dynasties, term limits and female political representation: Evidence from the Philippines," *Journal of Economic Behavior & Organization*, 182, 212–228.

Lee, A. (2020): "Incumbency, parties, and legislatures: Theory and evidence from India," *Comparative Politics*, 52, 311–331.

Longuet-Marx, N. (2025): "Party Lines or Voter Preferences? Explaining Political Realignment," *working paper*.

Mattozzi, A. and A. Merlo (2015): "Mediocracy," *Journal of Public Economics*, 130, 32–44.

Nevo, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69, 307–342.

Prakash, N., M. Rockmore, and Y. Uppal (2019): "Do criminally accused politicians affect economic outcomes? Evidence from India," *Journal of Development Economics*, 141, 102370.

Roy, R. (1966): "Selection of Congress candidates. I: The formal criteria," *Economic and Political Weekly*, 833–840.

Sekhon, J. S. and R. Titiunik (2012): "When Natural Experiments Are Neither Natural nor Experiments," *American Political Science Review*, 106, 35–57.

Trebbi, F. and E. Weese (2019): "Insurgency and small wars: Estimation of unobserved coalition structures," *Econometrica*, 87, 463–496.

Ujhelyi, G., S. Chatterjee, and A. Szabó (2021): "None of the Above: Protest Voting in the World's Largest Democracy," *Journal of the European Economic Association*, 19, 1936–1979.

Uppal, Y. (2009): "The disadvantaged incumbents: estimating incumbency effects in Indian state legislatures," *Public choice*, 138, 9–27.

Vaishnav, M. (2017): *When crime pays: Money and muscle in Indian politics*, Yale University Press.

Online Appendix to "Do Parties and Voters Disagree? An Equilibrium Analysis of Candidate Selection in India"

Arvind Magesan
Economics Department, University of Calgary

Andrea Szabó
Economics Department, University of Houston

Gergely Ujhelyi
Economics Department, University of Houston

December 17, 2025

Contents

1 Data and sample **3**
1.1 Construction of the Muslim characteristic 3
1.2 Sample construction 3
1.3 Missing characteristics 4
1.4 Ideology measures 6

2 Demand estimation: further details **7**
2.1 Specification choice 7
2.2 Sensitivity of the demand estimates 8

3 Identification of model parameters **13**

4 Clustering and candidate types **15**

5 Additional results with ideology measures **16**

6 Criminals as strategic complements **20**

7 Additional robustness checks **20**
7.1 Alternative payoff specifications 20
7.2 Constituency-level heterogeneity 24

1 Data and sample

1.1 Construction of the Muslim characteristic

We construct a Muslim indicator based on candidates' names. First, we used the candidate affidavit data for all elections between 2008-2019 to obtain a library of Muslim names and common substrings ("name fragments") found in Muslim names. Indian Muslim names have Arabic or Farsi origins but are often spelled differently or modified from the original name in some other way. We create a similar library for non-Muslim names. There are a total of 470 names and fragments in the Muslim library and 1210 in the non-Muslim library.

For every name in our data which we wish to classify, we compute the "distance" between the name and each of the two libraries. Denote the library of Muslim names as M and the library of non-Muslim names as H. Then, for every candidate name $name_i$ in our data, we calculate the Levenshtein distance to every item in each of M and H. We take the distance of $name_i$ and a library (M or H) to be the minimum of these distances across names in the library. Let $d(name_i, M)$ and $d(name_i, H)$ denote these distances.

Next, we use the name fragments to construct another measure. We count, for every name in the data, how many Muslim fragments and how many non-Muslim fragments appear in the name, and divide this by the number of fragments in the respective library to get a frequency. Denote these as $frag(name_i, M)$ and $frag(name_i, H)$ respectively.

Finally, $name_i$ is assigned "Muslim" identity if either $frag(name_i, M) > frag(name_i, H)$ or $\{frag(name_i, M) = frag(name_i, H)$ and $d(name_i, M) < d(name_i, H)\}$, and it is assigned "non-Muslim" identity otherwise.

1.2 Sample construction

Our dataset is limited by the state constituencies for which we can obtain demographic information from the SHRUG (Asher et al. 2021). Because we analyze national elections, we also need to aggregate the state constituencies up to the national constituency level. There are reasons to believe that state constituencies that are missing in a national constituency are systematically different from other state constituencies (e.g., they are more likely to be large urban areas). Therefore we only include in our analysis national constituencies for which we have information on all the state constituencies they contain. This drops from the sample several states in their entirety (mostly small states with only a few national constituencies).[1]

[1]The states excluded and the total number of their national constituencies are: Arunachal Pradesh (2), Goa (2), Manipur (2), Meghalaya (2), Mizoram (1), Nagaland (1), Puducherry (1), Punjab (13), Sikkim (1), Tripura (2), Uttarakhand (5), Delhi NCT (7).

From the remaining 18 states, we drop 3 because we either only have less than 20% of their constituencies or because they have very few constituencies to begin with.[2]

The remaining 15 states contain 478 of the 538 constituencies in India, and we have constituency characteristics from the SHRUG for 234 of these. We drop 2 constituencies because some of their candidates have unrealistically high numbers of criminal convictions,[3] leaving us with a total of 232 constituencies in the dataset.

1.3 Missing characteristics

The specification of both voters' and parties' choices requires that all relevant candidate characteristics be observed. In the data, education, assets, and criminal history have missing values, and we impute these characteristics based on the candidate's gender, age, and SC/ST status (which have no missing values). Specifically, we impute assets, number of criminal cases, and number of completed years of education using the average by gender, SC/ST and age range (+/- 1 year relative to the candidate's age). For example, a 30 year old male non-SC/ST candidate's imputed asset is the average of all male non-SC/ST candidates aged 29-31. We then create the variables used in the estimation and the clustering algorithm (log(assets + 1), an indicator for at least one criminal case, and an indicator for at least twelve years of education). In practice, the main impact of imputing these characteristics is that we can use all candidates in the data when creating types.[4]

The state election data has 36,766 observations. We drop 2 observations because they have missing characteristics and have no similar candidates (based on gender, SC/ST and age) that we could use for imputing these missing values. The asset variable has 6,301 missing values and 562 zeros which we also treat as missing. Criminal history has 6,301, and education 7,629 missing values. The national election data has 6,581 observations. We drop 3 observations with missing characteristics that cannot be imputed. After aggregating small parties and independent candidates, we have 2,649 observations. Assets, criminal history, and education have, respectively, 119, 93 and 172 missing values. Among the 897 candidates of the UPA and the NDA, the corresponding numbers are 38, 36, and 53.

[2]Specifically, we drop Chhattisgarh, with only 2 out of 11 constituencies, Himachal Pradesh, with 2 out of 4 constituencies, and Jammu & Kashmir, with 1 out of 6 constituencies.

[3]Both of these are in Tamil Nadu, and both have a candidate with close to 400 criminal cases.

[4]An alternative approach we considered is to use a Missing indicator as an additional characteristic. However, there are differences between missing values in the state and national election data (for example, most independent candidates' education, assets, and criminal history is missing in the former). Thus, using Missing as an additional characteristic would mechanically make candidate types in the two datasets less comparable.

Table A.1: Summary statistics

	N	Mean	Std. Dev.	Median	10%	90%
A. Candidate characteristics - state elections						
UPA (0/1)	36766	0.10				
NDA (0/1)	36766	0.11				
Education (0/1)	29137	0.62				
Muslim (0/1)	36766	0.11				
Crime (0/1)	30465	0.20				
Assets (log)	29903	13.90	2.36	14.00	10.84	16.83
Male (0/1)	36766	0.93				
Age	36766	44.76	11.27	44	31	61
B. Candidate characteristics - national elections						
UPA (0/1)	6581	0.07				
NDA (0/1)	6581	0.07				
Education (0/1)	3425	0.73				
Muslim (0/1)	6581	0.12				
Crime (0/1)	3611	0.23				
Assets (log)	3544	14.63	2.51	14.76	11.46	17.60
Male (0/1)	6581	0.93				
Age	6581	46.25	11.96	45	31	63
SC or ST (0/1)	6581	0.35				
Incumbent (0/1)	6581	0.04				
Broadcast allowance (log)	6581	1.67	2.38	0.00	0.00	5.86
C. Constituency characteristics - national elections						
Eligible voters (1000)	464	1421.50	195.59	1426.24	1173.14	1685.34
Turnout (%)	464	64.97	12.39	65.99	47.57	81.05
N. of candidates before aggregation	464	14.18	6.06	14	7	22
N. of candidates after aggregation	464	5.71	1.46	5	4	8
Reserved constituency (0/1)	464	0.28				
Literate population (share)	464	0.61	0.10	0.61	0.49	0.74
ST and SC population (share)	464	0.27	0.14	0.24	0.13	0.46
Rural population (share)	464	0.82	0.11	0.83	0.68	0.94
Population with paved roads (share)	464	0.83	0.19	0.90	0.58	1.00
Employed population (share)	464	0.42	0.07	0.43	0.32	0.50
Redistricting overlap (share)	464	0.87	0.17	1.00	0.59	1.00

Notes: Education: 1 if completed high school. Crime: 1 if has at least one criminal case. Assets (log): log of real assets in Rp. Incumbent: 1 if held the same seat just before the election. N. of candidates after aggregation: number of candidates after small-party candidates are aggregated. Redistricting overlap: largest share of a constituency's area in an old constituency.

1.4 Ideology measures

Wikipedia-based measures. We start from https://en.wikipedia.org/wiki/List_of_political_parties_in_India and the individual party websites linked from this page, and we collect information on the ideology of the 32 parties that enter our dataset as individual parties (i.e., are not in the "aggregate" party).

Table A.2: Coding of Left–Right ideology positions from Wikipedia

Description	Our code	N. parties	Description	Our code	N. parties
"far left"	-1	1	"center" or "syncretic"	0	8
"left wing to far left"	-0.75	1	"right wing" or "center right"	0.5	6
"left wing" or "center left"	-0.5	11	"right wing to far right"	0.75	1
"center to center left"	-0.25	3	"far right"	1	1
Total					32

Notes: Ideology categories from party descriptions on Wikipedia.

The first measure, Left-Right position, is coded as in Table A.2. If a party is in the NDA (respectively, UPA) alliance, we take its candidates' ideology to be the average of the party's and the BJP's (respectively, INC's) ideology. This reflects the idea that a party allied with one of these major parties is likely to campaign on a shared platform and present a blended ideological stance to voters.

The second measure uses Wikipedia's list of up to six specific ideological keywords for each party. Across the 32 parties in our sample, there are a total of 59 unique terms. We construct indicators for each ideological term, applying minimal grouping only when we are confident the terms refer to the same concept. If a party is in an alliance, we take the union of its ideology terms and those of the BJP (if NDA) or the INC (if UPA). After dropping perfectly co-linear variables, we are left with 12 unique ideological terms, shown in Table A.3.

Table A.3: Ideological keyword measures from Wikipedia

Term	N. parties	N. candidates	N. parties / alliances	Term	N. parties	N. candidates	N. parties / alliances
Regionalism	13	260	18	Integral	1	438	14
Agrarian	2	4	3	Islamic	1	2	1
Caste politics	9	121	11	Liberalism	1	455	12
Left	21	1574	29	Populism	5	308	7
Right	7	511	19	Progressivism	3	166	5
Dravidianism	3	42	4	Secularism	16	952	27

Notes: Keywords from Wikipedia party descriptions. N. parties/alliances shows the frequency of each term across the union of parties in an alliance.

Expert assessment based measures. We use information from the Democratic Accountability and Linkages Project (DALP; Kitschelt and Smidt (2013)), a dataset based on a 2009 survey of experts that scores parties along 5 dimensions (social spending on the disadvantaged, statism, public spending, national identity, traditional authority). These measures are only available for 15 of the 32 parties, and we assign values to the missing parties by predicting their scores based on the full set of Wikipedia terms described above.

First, we create an ideology index using principal components analysis (we find that the first principal component explains almost 80% of the variation in the data). Second, we focus on individual questions from the survey. In particular, Chhibber and Verma (2018) argue that there are two meaningful ideological dimensions among Indian parties, Recognition and Statism. We approximate the former with "Social spending on the disadvantaged," measured on a 1-10 scale where 1 is "Party advocates extensive social spending redistributing income to benefit the less well-off in society" and 10 is "Party opposes" such spending. To approximate Statism, we use "State role in governing the economy," measured on a 1-10 scale where 1 is "Party supports a major role for the state in regulating private economic activity to achieve social goals, in directing development, and/or maintaining control over key services." and 10 is "Party advocates a minimal role for the state in governing or directing economic activity or development." As above, when parties are in an alliance, we average their score with the BJP's (if NDA) or the INC's (if UPA).

2 Demand estimation: further details

2.1 Specification choice

Following Gandhi and Houde (2019), we first enter the differentiation IVs as controls in a Logit specification. This specification includes all the control variables, and instruments the endogenous characteristics as described in the paper. The results are in Table A.5. In column 1, the differentiation IVs for Muslim and Crime are statistically significant while the differentiation IVs for Education and Assets are not. This suggests that the former two are capable of capturing departures from the Logit model. As an alternative diagnostic, we also run a specification that includes the differentiation IVs as instruments instead of controls. The last row of the table (IIA p-val) shows the p-value of the overidentification J-test for this specification. The fact that this specification is clearly rejected also provides support for focusing on the nonlinear specifications (Gandhi and Houde 2019). In column 2 we use only the differentiation IVs for Muslim and Crime and obtain similar conclusions. Column 3 and 4 show corresponding estimates when Incumbent $\times$ Assets is included as another

Table A.4: Characteristics regressed on instruments

Dep. var.:	Education		Muslim		Crime		Assets		Incumbent		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
IV^1	0.16	0.11	0.54	0.33	0.30	0.23	0.30	0.30	0.02	0.01	0.01
	(0.05)	(0.05)	(0.07)	(0.08)	(0.06)	(0.06)	(0.05)	(0.05)	(0.00)	(0.00)	(0.00)
IV^2		0.18		0.38		0.21		0.02		0.01	0.01
		(0.05)		(0.08)		(0.06)		(0.05)		(0.00)	(0.00)
IV^3											0.14
											(0.05)
Adj. R^2	0.15	0.16	0.09	0.10	0.14	0.14	0.36	0.36	0.15	0.16	0.16
F	10.07	11.14	61.04	45.21	23.78	17.44	41.95	20.27	83.02	48.32	39.39
Mean	0.31	0.31	-0.03	-0.03	0.08	0.08	-0.02	-0.02	0.10	0.10	0.10
Std. dev.	0.81	0.81	0.88	0.88	0.98	0.98	0.96	0.96	0.30	0.30	0.30

Notes: For Education, Muslim, Crime and Assets, IV^1 is the average of the characteristic among a party's candidates in state constituencies in the given year, and IV^2 is the same variable for the other election in the data (2009 for 2014 and vice versa). For Incumbent, IV^1 is the party's competitiveness in the previous election, $IV^2 = IV^1 \times \mathbf{I}(year = 2014)$, and IV^3 is redistricting overlap. Regressions include broadcast allowance, and fixed effects for state, year, party, alliance, imputed characteristics, and reserved constituencies. F is the Olea and Pflueger (2013) effective F statistic. The last rows show the mean and standard deviation of the dependent variable. Robust standard errors in parentheses. N = 2,649.

(endogenous) characteristic.

2.2 Sensitivity of the demand estimates

We follow the method proposed by Andrews et al. (2017) to evaluate the extent to which our demand estimates are sensitive to failures of the moment conditions, i.e., the exogeneity of our instruments. We use their measure to compute the bias in our main estimates assuming a violation in each moment condition where a 1 standard deviation change in the instrument results in a 1 unit (approximately one third standard deviation) change in the econometric error ξ. We compute this bias for each instrument used to identify the endogenous characteristics, and display the impact on 8 of our key parameters in Figure A.2. The vertical lines show the parameter estimates (column 1 of Table 1 in the paper) and the dots the "true" parameter once the implied bias is taken into account. Entries on the Y-axis correspond to the 11 instruments (see section 5.2.2 in the paper for details).

The figure shows that none of our parameter estimates is very sensitive to this kind of mis-specification. For example, our point estimate for the marginal utility of Assets, 2.8485, would be biased away from values between 2.848 and 2.851 if the exogeneity assumption on the corresponding instrument failed in the manner described above.

Table A.5: Specification choice: differentiation IVs

	(1)	(2)	(3)	(4)
Incumbent	3.15	3.19	3.81	3.78
	(0.97)	(0.96)	(1.22)	(1.20)
Assets	1.51	1.54	2.36	2.40
	(0.39)	(0.41)	(0.60)	(0.64)
Incumbent × Assets			-3.36	-3.06
			(1.37)	(1.29)
Education	-0.46	-0.52	-1.54	-1.57
	(0.45)	(0.48)	(0.70)	(0.75)
Muslim	-0.39	-0.41	-0.41	-0.43
	(0.18)	(0.18)	(0.23)	(0.24)
Crime	0.37	0.28	0.04	-0.04
	(0.30)	(0.30)	(0.39)	(0.39)
diffIV(Educ)	-0.02		-0.01	
	(0.03)		(0.04)	
diffIV(Muslim)	-0.07	-0.08	-0.05	-0.07
	(0.03)	(0.03)	(0.03)	(0.03)
diffIV(Crime)	0.07	0.06	0.07	0.03
	(0.03)	(0.03)	(0.04)	(0.05)
diffIV(Assets)	-0.03		-0.10	
	(0.04)		(0.06)	
J p-val	0.00	0.00	0.15	0.14
IIA p-val	0.00	0.00	0.00	0.06

Notes: Gandhi and Houde (2019) specification checks. The dependent variable is vote shares. Candidate characteristics are instrumented as described in section 5.2.2 in the paper, and the differentiation IVs are entered as controls. Specifications also control for broadcast allowance and fixed effects for state, year, party, alliance, imputed characteristics, and reserved constituencies. J p-val is the p-value of the overidentification J test. IIA p-val is the p-value of the overidentification J test when the diffIV variables are used as instruments instead of controls. Robust standard errors in parentheses.

Table A.6: Specification choice: differentiation IVs and voter demographics

	Lit	Rural	Roads	Empl	SC/ST	Lit	Rural	Roads	Empl	SC/ST	SC/ST / Rural
Crime	0.02	0.05	-0.00	0.13	0.18						0.18
	(0.07)	(0.05)	(0.05)	(0.10)	(0.10)						(0.10)
Muslim						-0.12	-0.07	-0.10	-0.13	-0.05	-0.09
						(0.05)	(0.04)	(0.03)	(0.07)	(0.07)	(0.03)
J p-val	0.15	0.14	0.18	0.12	0.14	0.14	0.13	0.16	0.14	0.13	0.14
IIA p-val	0.06	0.06	0.06	0.05	0.04	0.10	0.07	0.02	0.12	0.24	0.05

Notes: Gandhi and Houde (2019) specification checks. The dependent variable is vote shares. Candidate characteristics are instrumented as described in section 5.2.2 in the paper, and the differentiation IVs are entered as controls. The first 5 columns include diffIVs for Crime, the next five include diffIVs for Muslim. The last column includes the diffIV for SC/ST × Crime and Muslim × Rural. Only the coefficients on the diffIVs are shown. Specifications control for broadcast allowance and fixed effects for state, year, party, alliance, imputed characteristics, and reserved constituencies. J p-val is the p-value of the overidentification J test. IIA p-val is the p-value of the overidentification J test when the diffIV variables are used as instruments instead of controls. Robust standard errors in parentheses.

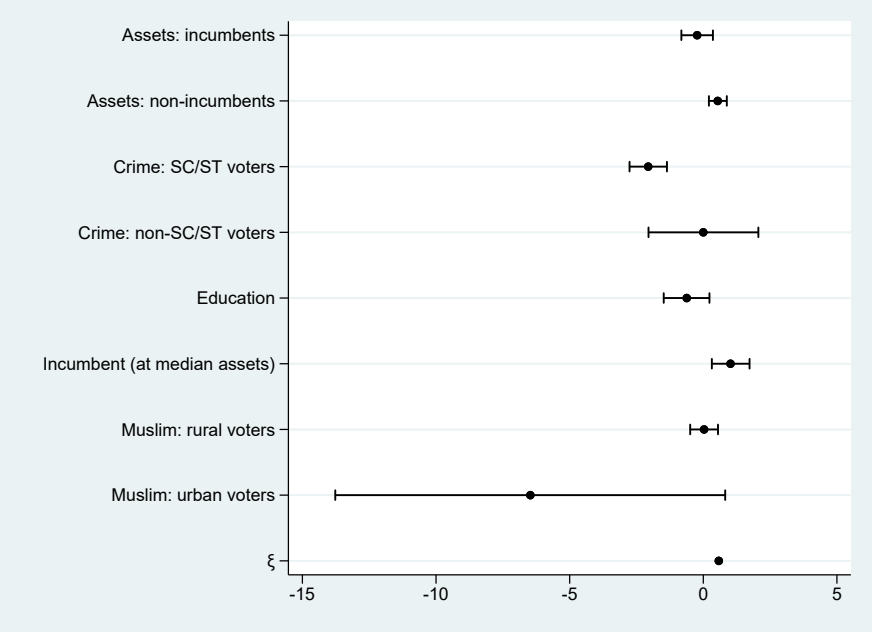


Figure A.1: Magnitude of voter preference for different candidate characteristics

Notes: Estimated effect of increasing the listed characteristic from its lowest to its highest value (if binary) or by 1 standard deviation (if continuous) on voter utility, measured in standard deviation units. The effect of Asset is shown separately for incumbent and non-incumbent candidates, and the effect of Muslim and Crime is shown separately for different voter groups. Values based on the specification in Table 1, column 1. Bands show the 95% confidence intervals.

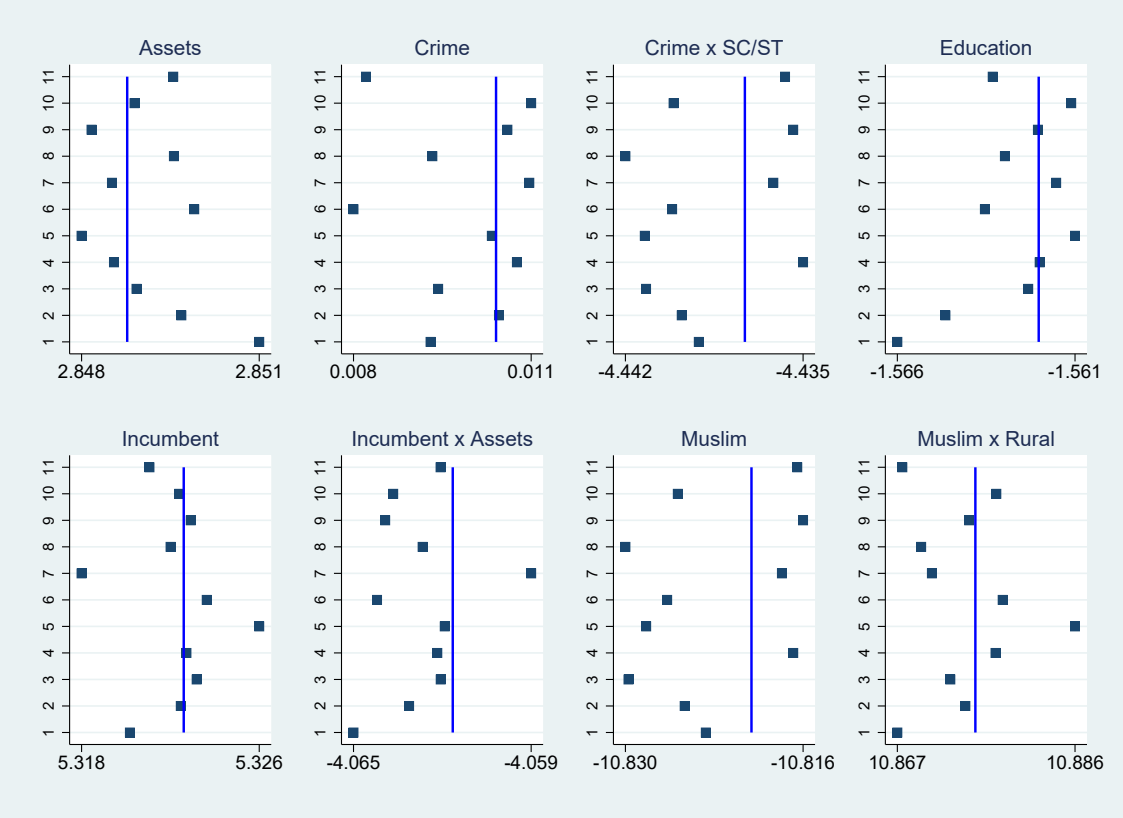


Figure A.2: Sensitivity of demand estimates to failures of the moment conditions

Notes: Andrews et al. (2017) sensitivity measures for the demand parameters listed. On each panel, the vertical line shows the point estimate in our main specification. The dots show the "true" parameter (our estimate minus its bias) assuming that a 1 standard deviation change in one of the instruments results in a 1 unit change in the econometric error ξ. The instruments listed on the Y axis are 1-2: Education instruments, 3-4: Muslim instruments, 5-6: Crime instruments, 7: Asset instrument, 8-10: Incumbent instruments, 11: Instrument for Asset $\times$ Incumbent.

Table A.7: Additional demand specifications

	(1)	(2)	(3)	(4)	(5)
Incumbent	5.46	4.46	4.12	5.60	8.74
	(1.92)	(1.96)	(1.74)	(1.80)	(4.39)
Assets	2.67	2.63	2.80	2.79	5.22
	(0.81)	(1.01)	(0.91)	(0.97)	(3.93)
Incumbent × Assets	-3.82	-3.6	-3.77	-4.02	-8.26
	(1.60)	(1.48)	(1.55)	(1.76)	(6.91)
Education	-1.32	-1.46	-1.67	-1.61	-1.47
	(0.96)	(0.96)	(0.84)	(1.35)	(1.70)
Muslim	-9.67	0.26	0.35	-11.49	-11.17
	(5.69)	(2.72)	(3.45)	(6.14)	(8.92)
Crime	0.17	0.41	0.34	-0.06	0.66
	(0.90)	(0.49)	(0.92)	(0.73)	(0.93)
Nonlinear parameters:					
Crime × SC/ST	-5.28	-4.47	-3.43	-4.56	-9.72
	(3.02)	(2.50)	(4.13)	(2.52)	(6.64)
Muslim × Rural	9.67	-0.86	-0.94	11.50	11.14
	(5.70)	(3.97)	(4.58)	(6.21)	(8.97)
v_{Crime}	-0.11		-0.17		
	(7.15)		(6.84)		
v_{Muslim}		0.06	0.02		
		(10.95)	(16.59)		
$v_{Education}$				0.07	
				(14.55)	
v_{Asset}					-2.91
					(2.53)
J	5.42	10.51	10.63	4.68	2.11
p-value	0.37	0.06	0.06	0.46	0.83
Newey-West pval	0.12	0.34	0.55	0.16	0.41

Notes: BLP estimates. Specifications include broadcast allowance and fixed effects for state, year, party, alliance, imputed characteristics, and reserved constituencies. J is the overidentification J-statistic (df=5) with its p-value. The bottom row shows the p-value of the Newey-West D-test for the null that all nonlinear parameters are jointly 0. Robust standard errors clustered by constituency in parentheses. N = 2,649.

3 Identification of model parameters

We formally establish identification of party payoffs and discuss the intuition of the identification results. As is standard in the empirical game literature (see, for example, Bajari et al. (2013)), we assume that choice probabilities are known to the researcher.[5] Then given vote shares, win probabilities and expected vote shares are also known and respectively given by $w_p^P(k,\mathbf{z}) = E_p[w_p(a_p, a_{-p}, \mathbf{z})|a_p = k]$ and $s_p^P(k,\mathbf{z}) = E_p[s_p(a_p, a_{-p}, \mathbf{z})|a_p = k]$ for choice $a_p = k$, $k = 1, ..., K$, where the expectation $E_p[w_p(a_p, a_{-p}, \mathbf{z})|a_p = k]$ is an integration over a_{-p} using player $-p$'s choice probability and $\mathbf{z}$ is the vector of constituency specific observable payoff variables.

We begin with a simple case with no costs so the only parameters to identify are (b^w, b^s). With identification in this simple case established, we then re-introduce cost parameters below. In this simple case, party p's choice probability satisfies

$$P_p(k,\mathbf{z}) = \Lambda\big(b^w \times w_p^P(k,\mathbf{z}) + b^s \times s_p^P(k,\mathbf{z})\big) = \frac{\exp\left\{b^w \times w_p^P(k,\mathbf{z}) + b^s \times s_p^P(k,\mathbf{z})\right\}}{\sum_{k'} \exp\left\{b^w \times w_p^P(k',\mathbf{z}) + b^s \times s_p^P(k',\mathbf{z})\right\}}$$

(As the argument for identification is symmetric across players, we drop the p subscript in what follows.) Since $\Lambda(\cdot)$ is strictly increasing, inverting the choice probability gives:

$$\begin{aligned} \Lambda^{-1}\big(P(k,\mathbf{z})\big) &= \ln\big(P(k,\mathbf{z})\big) - \ln\big(P(K,\mathbf{z})\big) \\ &= b^w \times (w^P(k,\mathbf{z}) - w^P(K,\mathbf{z})) + b^s \times (s^P(k,\mathbf{z}) - s^P(K,\mathbf{z})) \end{aligned} \tag{1}$$

where we have taken type K as the reference type. Define $\Delta_w^P(k,\mathbf{z}) \equiv w^P(k,\mathbf{z}) - w^P(K,\mathbf{z})$ and $\Delta_s^P(k,\mathbf{z}) \equiv s^P(k,\mathbf{z}) - s^P(K,\mathbf{z})$ and consider two constituencies, $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$. These satisfy a system of equations:

$$\begin{bmatrix} \Lambda^{-1}\big(P(k,\mathbf{z}^{(1)})\big) \\ \Lambda^{-1}\big(P(k,\mathbf{z}^{(2)})\big) \end{bmatrix} = \begin{bmatrix} \Delta_w^P(k,\mathbf{z}^{(1)}) & \Delta_s^P(k,\mathbf{z}^{(1)}) \\ \Delta_w^P(k,\mathbf{z}^{(2)}) & \Delta_s^P(k,\mathbf{z}^{(2)}) \end{bmatrix} \begin{bmatrix} b^w \\ b^s \end{bmatrix} \tag{2}$$

and the parameters (b^w, b^s) are identified under the rank condition:

$$\frac{\Delta_w^P(k,\mathbf{z}^{(1)})}{\Delta_w^P(k,\mathbf{z}^{(2)})} \neq \frac{\Delta_s^P(k,\mathbf{z}^{(1)})}{\Delta_s^P(k,\mathbf{z}^{(2)})} \tag{3}$$

Roughly, $\mathbf{z}$ should shift expected vote share independently of win probability. As vote shares can change without turning a loss into a win, this condition is easy to satisfy and b^w, b^s are identified.

[5] Consistent estimates of choice probabilities can be obtained, in principle, using a raw frequency estimator.

One of our key findings is that $b^w > 0 > b^s$. We illustrate how this can arise as well as the intuition given the identification argument. Solving the system in (2) we get

$$b^w = \frac{\Delta_s^P(k, \mathbf{z}^{(2)})\Lambda^{-1}(P(k, \mathbf{z}^{(1)})) - \Delta_s^P(k, \mathbf{z}^{(1)})\Lambda^{-1}(P(k, \mathbf{z}^{(2)}))}{\Delta_s^P(k, \mathbf{z}^{(2)})\Delta_w^P(k, \mathbf{z}^{(1)}) - \Delta_s^P(k, \mathbf{z}^{(1)})\Delta_w^P(k, \mathbf{z}^{(2)})} \tag{4}$$

and

$$b^s = \frac{\Delta_w^P(k, \mathbf{z}^{(1)})\Lambda^{-1}(P(k, \mathbf{z}^{(2)})) - \Delta_w^P(k, \mathbf{z}^{(2)})\Lambda^{-1}(P(k, \mathbf{z}^{(1)}))}{\Delta_s^P(k, \mathbf{z}^{(2)})\Delta_w^P(k, \mathbf{z}^{(1)}) - \Delta_s^P(k, \mathbf{z}^{(1)})\Delta_w^P(k, \mathbf{z}^{(2)})} \tag{5}$$

Assume without loss that the denominator of Equations (4) and (5) is positive. This would occur if, for example, $\Delta_s^P(k, \mathbf{z}^{(2)}) > \Delta_s^P(k, \mathbf{z}^{(1)})$ and $\Delta_w^P(k, \mathbf{z}^{(2)}) < \Delta_w^P(k, \mathbf{z}^{(1)})$: Type k delivers relatively higher expected vote share but lower win probability in constituency 2 relative to constituency 1. For $b^w > 0$ in this case, it is sufficient that: $P(k, \mathbf{z}^{(1)}) > P(k, \mathbf{z}^{(2)})$. In other words, $b^w > 0$ rationalizes type k being chosen more frequently in constituency 1 than in constituency 2 even though $\Delta_s^P(k, \mathbf{z}^{(2)}) > \Delta_s^P(k, \mathbf{z}^{(1)})$.

Turning to Equation (5), for $b^s < 0$ given that $\Delta_w^P(k, \mathbf{z}^{(2)}) < \Delta_w^P(k, \mathbf{z}^{(1)})$, $P(k, \mathbf{z}^{(1)})$ must be *sufficiently* larger than $P(k, \mathbf{z}^{(2)})$. So in sum, if $\Delta_s^P(k, \mathbf{z}^{(2)}) > \Delta_s^P(k, \mathbf{z}^{(1)})$ and $\Delta_w^P(k, \mathbf{z}^{(2)}) < \Delta_w^P(k, \mathbf{z}^{(1)})$ then $P(k, \mathbf{z}^{(1)}) >> P(k, \mathbf{z}^{(2)})$ is rationalized by $b^w > 0$, $b^s < 0$.

Adding costs to the payoff function is straightforward. In this case the inverted choice probability satisfies:

$$\begin{aligned} \Lambda^{-1}\big(P(k, \mathbf{z})\big) &= \ln\big(P(k, \mathbf{z})\big) - \ln\big(P(K, \mathbf{z})\big) \\ &= b^w \times (w^P(k, \mathbf{z}) - w^P(K, \mathbf{z})) + b^s \times (s^P(k, \mathbf{z}) - s^P(K, \mathbf{z})) - (c_k - c_K) \end{aligned} \tag{6}$$

Now, consider two values of $\mathbf{z}$, say $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$. Differencing (6) across these two values we eliminate the cost parameters, assumed to be independent of $\mathbf{z}$. We can do the same at another pair of value of $\mathbf{z}$, say $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(3)}$, which gives us a system of two equations and two unknown parameters (b^w, b^s):

$$\begin{bmatrix} \Delta_\Lambda^P(k, \mathbf{z}^{(1,2)}) \\ \Delta_\Lambda^P(k, \mathbf{z}^{(1,3)}) \end{bmatrix} = \begin{bmatrix} \Delta_w^P(k, \mathbf{z}^{(1,2)}) & \Delta_s^P(k, \mathbf{z}^{(1,2)}) \\ \Delta_w^P(k, \mathbf{z}^{(1,3)}) & \Delta_s^P(k, \mathbf{z}^{(1,3)}) \end{bmatrix} \begin{bmatrix} b^w \\ b^s \end{bmatrix} \tag{7}$$

where $\Delta_\Lambda^P(k, \mathbf{z}^{(1,2)}) \equiv \Lambda^{-1}\big(P(k, \mathbf{z}^{(1)})\big) - \Lambda^{-1}\big(P(k, \mathbf{z}^{(2)})\big)$ (and similarly for $\Delta_\Lambda^P(k, \mathbf{z}^{(1,3)})$), and $\Delta_w^P(k, \mathbf{z}^{(1,2)}) \equiv \Delta_w^P\big(P(k, \mathbf{z}^{(1)})\big) - \Delta_w^P\big(P(k, \mathbf{z}^{(2)})\big)$ (and similarly for $\Delta_s^P(k, \mathbf{z}^{(1,2)})$).

The parameters (b^w, b^s) are identified if, for at least one triple $(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^3)$:

$$\frac{\Delta_w^P(k, \mathbf{z}^{(1,2)})}{\Delta_w^P(k, \mathbf{z}^{(1,3)})} \neq \frac{\Delta_s^P(k, \mathbf{z}^{(1,2)})}{\Delta_s^P(k, \mathbf{z}^{(1,3)})}$$

This condition is analogous to the condition in Equation (3) in the simpler case, except we now need variation in win probability and expected vote share across more values of $\mathbf{z}$ - this is how we identify b^w, b^s separately from the cost difference $c_k - c_K$.

Solving the system in (7) we get expressions for b^w and b^s analogous to equations (4) and (5) in the simple case with no costs. With b^w and b^s identified we can obtain an expression for the cost difference $c_k - c_K$ by substituting the identified values into (6).

4 Clustering and candidate types

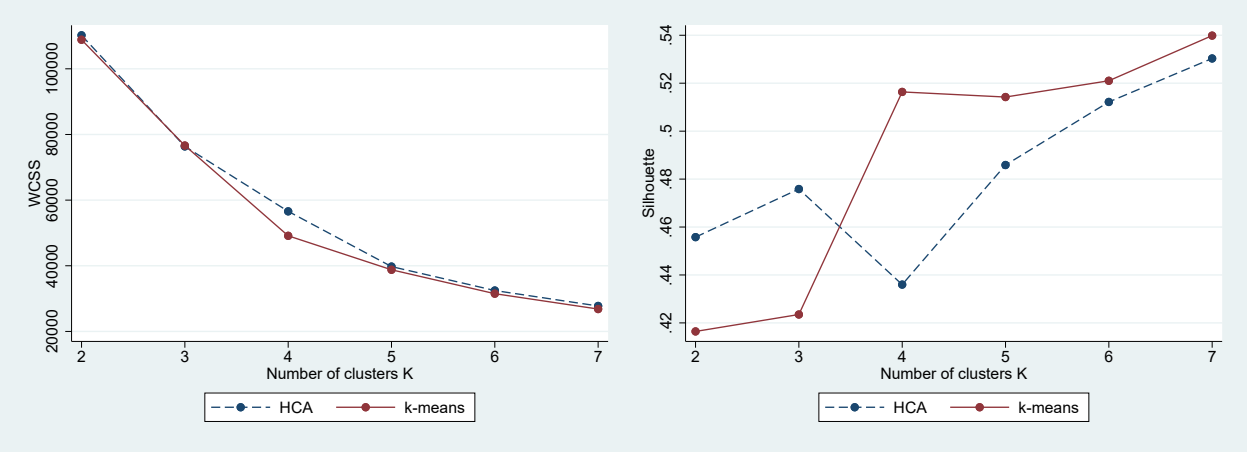


Figure A.3: WCSS and Silhouette

Notes: The Within-Cluster Sum of Squares (WCSS) measures the similarity of units within clusters. The Silhouette coefficient measures the dissimilarity of units across clusters. The graph shows the values of these measures that result from running the k-means or HCA clustering algorithm for different numbers of clusters K in our data.

We illustrate that standard methods of validation for unsupervised clustering methods support $K = 4$ using k-means in our application and compare our k-means approach to an alternative clustering algorithm, Hierarchical Clustering Analysis (HCA).

HCA begins by treating each of the N individual observation as its own "cluster," then iteratively agglomerates the clusters. In the first iteration, we take the two points that are closest in the sense that they have the lowest variance among any pair of points in the sample (the Ward merging criterion), and then combine those to form a new cluster. We now have $N-1$ clusters. We again find the two clusters that are closest and combine them, and so on. The number of clusters is determined by where we stop this iterative agglomeration process.

On the left panel of Figure A.3, we compare the within-cluster sum of squares, WCSS, over a number of possible clusters across the two methods. The k-means approach is uniformly better by this metric (i.e., produces lower WCSS scores).[6] On the right panel, we compare the Silhouette score associated with the two approaches. While HCA performs better for 2 or 3 clusters, k-means performs better from 4 clusters on. Moreover, k-means scores with 4 or more clusters dominate any HCA score below 4 clusters. Finally, note that we only ever hit a Silhouette score of 0.5 with k-means and the number of clusters at least 4. (In cluster analysis, a Silhouette score of 0.5 is often used as a rule of thumb for "good clustering" (Rousseeuw 1987).) These patterns indicate that k-means is preferable to HCA for our application.

Nevertheless, we computed the types that the HCA algorithm predicts when we set $k = 4$. These types look remarkably similar to what we obtain using k-means. Similar to k-means, HCA clusters all Muslim candidates together, and then among the remaining candidates clusters all criminals together. The one difference is that the HCA algorithm puts some uneducated individuals into Type 1 so that there is a larger gap in assets between the Type 1 and Type 2 candidates than there is using k-means.

Table A.8: Distribution of candidate types

	State elections			National elections			
	All candidates	UPA	NDA	All candidates	UPA	NDA	Incumbents
Type 1	34.74	46.78	44.15	37.59	50.00	52.53	47.03
Type 2	39.63	16.37	20.76	28.78	5.53	9.45	7.03
Type 3	11.20	11.27	6.55	12.71	11.29	5.07	8.11
Type 4	14.43	25.58	28.55	20.91	33.18	32.95	37.84
Total	100	100	100	100	100	100	100
N	36,764	3,788	4,109	6,049	434	434	185

Notes: Percentage of candidates in state and national elections assigned to each type by the k-means clustering algorithm. Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. National elections are the constituencies in the final estimation sample. The last column shows the distribution of re-nominated UPA or NDA incumbents.

5 Additional results with ideology measures

[6]Note that k-means is based on centroids (the average of all points in the cluster), while HCA is not. To generate Figure A.3, we compute the centroid of each HCA cluster analogously.

Table A.9: Voter preference parameter estimates with heterogeneity on parties

	(1)	(2)	(3)
Incumbent	3.46	6.81	3.12
	(2.23)	(2.88)	(1.82)
Assets	2.70	2.53	2.85
	(0.68)	(0.81)	(1.43)
Incumbent × Assets	-3.48	-4.19	-3.56
	(2.59)	(1.88)	(1.53)
Education	-0.57	-1.39	-1.30
	(0.78)	(1.09)	(1.25)
Muslim	0.04	-8.46	0.11
	(4.90)	(5.59)	(3.54)
Crime	0.79	0.25	0.05
	(0.48)	(0.83)	(0.85)
Ideology (Left-Right)	1.24	2.84	1.06
	(1.70)	(1.50)	(1.11)
Nonlinear parameters:			
Muslim × Rural	-0.55	8.13	-0.82
	(6.13)	(5.67)	(4.83)
Crime × SC/ST	-4.08	-5.41	0.81
	(5.48)	(2.49)	(4.08)
v_{UPA}	-0.04		
	(57.05)		
v_{NDA}	-0.01		
	(38.88)		
UPA × Roads		-5.50	
		(2.69)	
NDA × Roads		13.53	
		(21.47)	
UPA × SC/ST			-5.38
			(24.10)
NDA × SC/ST			-0.08
			(5.31)
J	18.00	6.32	14.64
p-value	0.00	0.28	0.01
Newey-West pval	0.59	0.05	0.75

Notes: BLP estimates. Column 1 includes Normally distributed random coefficients on the UPA and NDA dummies. Columns (2) and (3) allow for heterogeneity on these dummies by access to paved roads and SC/ST.

Table A.10: Additional voter preference parameter estimates with alternative ideology measures

	Left-Right	Left-Right	Left-Right	DALP	DALP	DALP
Incumbent	3.68	3.86	4.85	5.01	4.57	5.19
	(1.71)	(2.08)	(1.87)	(1.71)	(1.77)	(1.92)
Assets	3.39	3.69	3.17	2.05	2.88	2.73
	(1.11)	(0.99)	(1.14)	(0.56)	(1.24)	(0.88)
Incumbent × Assets	-4.42	-5.06	-4.31	-2.68	-4.05	-3.31
	(1.72)	(1.93)	(2.51)	(1.22)	(2.00)	(1.77)
Education	-1.54	-1.79	-1.53	0.18	-0.80	-1.27
	(1.23)	(1.19)	(1.00)	(0.72)	(1.33)	(1.18)
Muslim	-7.25	-11.84	-12.23	-6.92	-9.96	-11.90
	(5.85)	(7.38)	(6.89)	(4.62)	(6.46)	(6.84)
Crime	0.57	0.01	0.23	0.89	0.47	0.51
	(0.65)	(0.83)	(0.87)	(0.45)	(0.67)	(0.72)
Ideology	2.46	-0.13	1.84	0.50	0.14	0.46
	(1.55)	(2.45)	(2.87)	(0.20)	(0.58)	(0.27)
Nonlinear parameters:						
Muslim × Rural	7.25	11.94	12.28	6.91	9.98	11.91
	(5.96)	(7.46)	(6.92)	(4.67)	(6.51)	(6.82)
Crime × SC/ST	-4.69	-3.06	-4.69	-6.52	-5.55	-4.84
	(2.64)	(3.67)	(2.58)	(1.85)	(2.41)	(2.77)
v_{Ideology}	0.02			-0.01		
	(34.03)			(4.48)		
Ideology × Roads		3.44			0.69	
		(2.19)			(0.80)	
Ideology × SC/ST			0.84			0.51
			(4.55)			(0.32)
J	7.82	5.07	3.78	11.86	7.81	4.96
p-value	0.17	0.41	0.58	0.04	0.17	0.42
Newey-West pval	0.53	0.18	0.16	0.06	0.07	0.05

Notes: BLP estimates. Column headers show the ideology measures used. In each case, the first specification includes a Normally distributed random coefficient on Ideology, the second allows for heterogeneity by voters' access to paved roads, and the third allows for heterogeneity by voter's SC/ST status.

Table A.11: Party objective function estimates with ideology measure

	(1)	(2)		(1)	(2)
b^w	3.18	2.75	c_4^{educ}	0.46	0.56
	(0.77)	(0.85)		(0.42)	(0.40)
b^s	-8.52	-6.70	c_1^{crime}	0.92	0.98
	(1.13)	(1.17)		(0.28)	(0.27)
$c_{1,NDA}^0$	2.48	1.78	c_2^{crime}	0.77	0.82
	(0.50)	(0.45)		(0.36)	(0.36)
$c_{2,NDA}^0$	3.36	2.74	c_3^{crime}	0.46	0.53
	(0.50)	(0.45)		(0.36)	(0.35)
$c_{3,NDA}^0$	4.49	3.78	c_4^{crime}	-0.13	-0.05
	(0.56)	(0.51)		(0.29)	(0.28)
$c_{4,NDA}^0$	3.55	2.78	c_1^{asset}	0.40	0.19
	(0.50)	(0.45)		(0.33)	(0.31)
$c_{1,UPA}^0$	1.91	1.48	c_2^{asset}	0.36	0.09
	(0.41)	(0.38)		(0.41)	(0.39)
$c_{2,UPA}^0$	3.50	3.10	c_3^{asset}	0.83	0.61
	(0.49)	(0.48)		(0.41)	(0.40)
$c_{3,UPA}^0$	3.22	2.77	c_4^{asset}	0.03	-0.09
	(0.48)	(0.45)		(0.34)	(0.32)
$c_{4,UPA}^0$	2.91	-2.38	c_1^{Muslim}	0.28	0.32
	(0.43)	(0.39)		(0.40)	(0.39)
c_1^{educ}	0.28	0.33	c_2^{Muslim}	0.26	0.30
	(0.40)	(0.38)		(0.56)	(0.56)
c_2^{educ}	0.77	0.85	c_3^{Muslim}	-1.05	-1.00
	(0.51)	(0.50)		(0.45)	(0.43)
c_3^{educ}	0.54	0.58	c_4^{Muslim}	0.05	0.14
	(0.52)	(0.51)		(0.43)	(0.41)
Log likelihood				-819.31	-836.41

Notes: Column 1 uses the demand specification in column 1 of Table 6 in the paper. Column 2 uses the demand specification in column 2 of Table A.10. The cost parameters c are measured relative to the excluded category, incumbents. $c_{k,p}^0$ is party p's direct cost of choosing a type k candidate. c_k^l is the impact of characteristic l in the pool of candidates on the recruitment cost of a type k candidate. Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. Standard errors in parentheses. Number of markets: 434.

6 Criminals as strategic complements

To obtain a precise measure of strategic complementarity, write party p's choice probability associated with choosing some type a' as $P_p(a') = \frac{\exp\left\{\tilde{U}_p(a',P)\right\}}{\sum_{a=1}^{K}\exp\left\{\tilde{U}_p(a,P)\right\}}$, where $\tilde{U}_p(a,P) = E_P[b^w w_p(a)+b^s s_p(a)]+c_p(a)$ and K is 5 if the party can choose to retain the incumbent and 4 otherwise. Letting $\Delta^P(a) \equiv \tilde{U}_p(a,P)-\tilde{U}_p(4,P)$, we have $P_p(4) = 1/\left(1+\sum_{a\neq 4}\exp\left\{\Delta^P(a)\right\}\right)$. Differentiating with respect to the opponent's probability of choosing type 4, we get

$$\frac{\partial P_p(4)}{\partial P_{-p}(4)} = -\sum_{a\neq 4}\frac{\exp\left\{\Delta^P(a)\right\}\frac{\partial \Delta^P(a)}{\partial P_{-p}(4)}}{\left(1+\sum_{a\neq 4}\exp\left\{\Delta^P(a)\right\}\right)^2} = -\sum_{a\neq 4} P_p(4)P_p(a)\frac{\partial \Delta^P(a)}{\partial P_{-p}(4)}$$

We have

$$\frac{\partial \Delta^P(a)}{\partial P_{-p}(4)} = b^w\frac{\partial E_P[w_p(a)-w_p(4)]}{\partial P_{-p}(4)} + b^s\frac{\partial E_P[s_p(a)-s_p(4)]}{\partial P_{-p}(4)}.$$

where

$$\frac{\partial E_P[w_p(a)-w_p(4)]}{\partial P_{-p}(4)} = w_p(a,4)-w_p(4,4)+\sum_{t\neq 4}\left[w_p(a,t)-w_p(4,t)\right]\frac{\partial P_{-p}(t)}{\partial P_{-p}(4)}, \tag{8}$$

and similarly for $\frac{\partial E_P[s_p(a)-s_p(4)]}{\partial P_{-p}(4)}$.

Using our parameter estimates, we compute (8) for both parties in every constituency. We do this by assuming that $\frac{\partial P_{-p}(t)}{\partial P_{-p}(4)}$ in (8) is the same for all $t \neq 4$ (if we instead set $\frac{\partial P_{-p}(t)}{\partial P_{-p}(4)} = -1$ for some t we obtain very similar results). A positive derivative $\frac{\partial P_p(4)}{\partial P_{-p}(4)}$ indicates strategic complementarity between criminal candidates.

Among constituencies with no incumbent available, we find that the derivative is positive for both parties in 84%, and positive for at least one of the parties in 99%.

7 Additional robustness checks

7.1 Alternative payoff specifications

Table A.12 probes the robustness of our results, in particular $b^s < 0$, to modifications of parties' payoffs. The first two columns include nonlinearities in win probability, and we continue to find $b^s < 0$. Columns 3 and 4 replace vote share with vote margin or number of votes, respectively, as alternative measures of popularity. Both of these measures also yield a significant negative coefficient.

In Table A.13, we add an (expected) interaction term

$$\sum_{\mathbf{a}_{-p}} w_p(a_p, \mathbf{a}_{-p}) s_p(a_p, \mathbf{a}_{-p}) P(\mathbf{a}_{-p})$$

to the party payoffs. This allows for the possibility that parties value expected vote share differentially, depending on whether the share is associated with an anticipated win or a loss. According to the estimates, parties have a negative but insignificant payoff from more popular candidates who are likely to lose, and a significant negative payoff from candidates who are likely to win. This confirms that our main estimates in the paper are driven by parties' aversion to candidates who are expected to win big.

Table A.12: Alternative payoff specifications

	(1)	(2)	(3)	(4)
b^w	8.29	4.87	3.00	3.85
	(1.20)	(2.29)	(0.75)	(0.80)
b_2^w	-5.68	3.54		
	(1.17)	(5.33)		
b_3^w		-6.78		
		(3.85)		
b^s	-8.53	-8.40		
	(1.34)	(1.36)		
b^m			-6.65	
			(0.81)	
b^n				-0.73
				(0.08)
Log likelihood	-782.00	-779.93	-809.89	-796.14

Notes: Estimates of alternative payoff specifications. Cost parameters c are not shown. Column 1 adds squared win probability (the expected value of w^2) and column 2 also adds cubed win probability. Column 3 replaces expected vote share s with expected vote margin (with a corresponding parameter b^m) and column 4 with expected number of votes (with parameter b^n). Standard errors in parentheses. Number of markets: 434.

Table A.13: Party objective functions, with interaction between win probability and vote share

	(1)	(2)	(3)		(1)	(2)	(3)
b^w	7.50	4.79	4.30	c_4^{educ}		2.03	0.26
	(0.68)	(0.86)	(0.90)			(0.41)	(0.49)
b^s	-12.70	-2.98	-2.09	c_1^{crime}		1.71	0.95
	(1.58)	(1.94)	(2.20)			(0.33)	(0.34)
$b^{w,s}$	2.01	-7.51	-8.79	c_2^{crime}		1.82	0.74
	(1.29)	(1.59)	(1.81)			(0.38)	(0.41)
$c_{1,NDA}^0$			3.02	c_3^{crime}		1.72	0.48
			(0.60)			(0.37)	(0.41
$c_{2,NDA}^0$			4.06	c_4^{crime}		0.94	-0.10
			(0.60)			(0.33)	(0.35)
$c_{3,NDA}^0$			5.09	c_1^{asset}		2.04	0.65
			(0.64)			(0.33)	(0.39)
$c_{4,NDA}^0$			4.14	c_2^{asset}		2.93	0.71
			(0.61)			(0.35)	(0.45)
$c_{1,UPA}^0$			1.80	c_3^{asset}		3.20	1.13
			(0.49)			(0.37)	(0.46)
$c_{2,UPA}^0$			3.57	c_4^{asset}		2.27	0.38
			(0.54)			(0.33)	(0.40)
$c_{3,UPA}^0$			3.19	c_1^{Muslim}		0.10	0.06
			(0.54)			(0.44)	(0.44)
$c_{4,UPA}^0$			2.84	c_2^{Muslim}		-0.17	0.13
			(0.50)			(0.52)	(0.60)
c_1^{educ}		1.71	0.05	c_3^{Muslim}		-1.51	-1.28
		(0.33)	(0.48)			(0.46)	(0.48)
c_2^{educ}		2.68	0.59	c_4^{Muslim}		-0.31	-0.08
		(0.45)	(0.58)			(0.44)	(0.46)
c_3^{educ}		2.80	0.32				
		(0.44)	(0.58)				
Log likelihood					-1187.01	-859.85	-781.58

Notes: $b^{w,s}$ denotes the parameter of the interaction between w and s. Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. Number of markets: 434.

Table A.14: Party objective function estimates: Constituency-level heterogeneity

	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
b^w	4.36	4.40	4.27	4.40	c_2^{educ}	0.39	0.69	0.68	0.55
	(0.84)	(0.85)	(0.83)	(0.86)		(0.54)	(0.55)	(0.55)	(0.56)
b^s	-11.31	-12.15	-11.43	-12.59	c_3^{educ}	0.26	0.43	-0.46	0.43
	(1.22)	(1.27)	(1.23)	(1.31)		(0.55)	(0.55)	(0.55)	(0.56)
c_1^*	1.10	0.62	5.81		c_4^{educ}	0.13	0.31	0.31	0.28
	(0.47)	(0.44)	(1.87)			(0.45)	(0.45)	(0.45)	(0.46)
c_2^*	1.99	0.01	4.18		c_1^{crime}	0.89	0.95	0.83	0.97
	(0.56)	(0.52)	(2.29)			(0.32)	(0.31)	(0.31)	(0.32)
c_3^*	0.99	0.85	3.34		c_2^{crime}	0.72	0.72	0.69	0.73
	(0.55)	(0.53)	(2.21)			(0.39)	(0.39)	(0.39)	(0.40)
c_4^*	1.50	1.58	5.17		c_3^{crime}	0.41	0.46	0.40	0.49
	(0.48)	(0.47)	(1.93)			(0.39)	(0.38)	(0.38)	(0.39)
$c_{1,NDA}^0$	3.16	3.30	-1.58	4.40	c_4^{crime}	-0.19	-0.08	-0.22	-0.04
	(0.55)	(0.56)	(1.59)	(1.35)		(0.32)	(0.32)	(0.32)	(0.33)
$c_{2,NDA}^0$	3.45	3.91	0.20	5.86	c_1^{asset}	-0.14	0.68	0.92	0.91
	(0.53)	(0.56)	(1.95)	(1.51)		(0.46)	(0.36)	(0.38)	(0.38)
$c_{3,NDA}^0$	5.07	5.07	2.30	6.29	c_2^{asset}	-0.57	0.62	0.84	0.94
	(0.60)	(0.61)	(1.91)	(1.59)		(0.53)	(0.43)	(0.45)	(0.45)
$c_{4,NDA}^0$	4.23	4.15	0.11	5.02	c_3^{asset}	0.37	1.18	1.23	1.34
	(0.56)	(0.56)	(1.64)	(1.44)		(0.54)	(0.44)	(0.46)	(0.46)
$c_{1,UPA}^0$	2.34	2.24	-2.42	3.32	c_4^{asset}	-0.79	0.32	0.43	0.56
	(0.46)	(0.48)	(1.57)	(1.37)		(0.47)	(0.37)	(0.38)	(0.39)
$c_{2,UPA}^0$	3.50	3.63	0.17	5.55	c_1^{Muslim}	0.19	0.32	0.34	0.35
	(0.52)	(0.56)	(1.96)	(1.56)		(0.44)	(0.44)	(0.43)	(0.45)
$c_{3,UPA}^0$	3.54	3.31	0.83	4.54	c_2^{Muslim}	0.16	0.22	0.27	0.31
	(0.52)	(0.54)	(1.90)	(1.60)		(0.59)	(0.59)	(0.59)	(0.61)
$c_{4,UPA}^0$	3.40	3.05	-0.77	3.90	c_3^{Muslim}	-1.14	-0.99	-0.99	-0.91
	(0.47)	(0.49)	(1.62)	(1.45)		(0.48)	(0.48)	(0.48)	(0.49)
c_1^{educ}	0.01	0.21	0.12	0.17	c_4^{Muslim}	-0.08	0.18	0.06	0.24
	(0.44)	(0.43)	(0.44)	(0.44)		(0.46)	(0.46)	(0.45)	(0.47)
Log likelihood						-786.17	-778.51	-789.12	-774.30

Notes: Estimates of party objective functions allowing for additional heterogeneity. Columns 1-3 interact candidate types with the following variables, respectively: 1: indicator for election year 2014; 2: indicator for reserved constituencies; 3: rural population. c_k^* is the parameter on the interaction for type k. In column 4, we include interactions with fixed effects for constituency clusters following Bonhomme et al. (2022) (the parameters on these are not shown). Types 1-4 are the educated, uneducated, Muslim, and criminal types, respectively. Number of markets: 434.

7.2 Constituency-level heterogeneity

In Table A.14 we study the robustness of our payoff function estimates to various types of observable and unobservable heterogeneity. In column 1 we consider the possibility that selection decisions are systematically different across election years 2009 and 2014. We augment the fixed costs to $c_k^0 + c_k^*$, where c_k^* is an additional direct cost in 2014. The estimates suggest that parties had a higher cost of replacing incumbents, and higher costs of selecting Type 2 and 4 candidates in 2014 than in 2009. However, there is little change in other estimates relative to our main estimates in the paper. In particular, the estimates of b^w and b^s are only marginally different, and so too are the rankings of the direct costs c_k^0.

In column 2, we allow costs to depend on constituency reservation status. The parameters c_k^* now denote an additional direct cost in reserved constituencies. The only estimate that is significantly different from 0 (at the 5% level) is c_4^*: there is a higher cost of selecting the criminal type in a reserved constituency. Again, our results are qualitatively unaffected by allowing for this type of heterogeneity.

In column 3 we allow for the possibility that selection decisions may depend on the rural/urban split in the constituency. The parameters c_k^* now denote an additional direct cost from a higher share of rural population. The estimates of b^w and b^s are again only marginally different from the estimates in the paper. The cost estimates indicate that the cost of replacing an incumbent is larger in rural constituencies.

Finally, we consider the possibility that constituency level *unobservables* may confound our estimates. Including a full set of constituency dummies in the model is not an option (interacting these with type choices would result in 651 additional parameters). Instead, we follow the approach of Bonhomme et al. (2022), who propose to reduce the dimensionality of the fixed effects by clustering units based on observable characteristics into groups in a first step. The idea is that constituencies in the same cluster likely have similar *unobservables*. We use kmeans as proposed by Bonhomme et al. (2022) to cluster constituencies using the full set of constituency level observables in the data. The algorithm identifies four constituency types, and we include dummies for these types (interacted with candidate types) in column 4 of Table A.14. Controlling for constituency level unobserved heterogeneity yields similar b^w and b^s estimates to our main specification, and causes little qualitative change in the cost parameters.

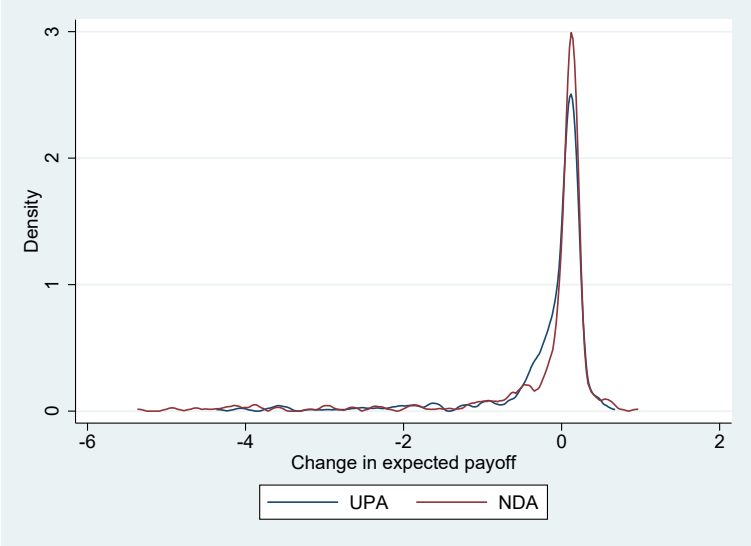


Figure A.4: Distribution of the change in parties' payoffs with a criminal ban

References

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *The Quarterly Journal of Economics*, 132 (4), 1553–1592. 8, 11

Asher, Sam, Tobias Lunt, Ryu Matsuura, and Paul Novosad (2021): "Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in India using the SHRUG open data platform," *The World Bank Economic Review*, 35 (4), 845–871. 3

Bajari, Patrick, Han Hong, and Denis Nekipelov (2013): "Game theory and econometrics: A survey of some recent research," . 13

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2022): "Discretizing unobserved heterogeneity," *Econometrica*, 90 (2), 625–643. 23, 24

Chhibber, Pradeep K and Rahul Verma (2018): *Ideology and identity: The changing party systems of India*, Oxford University Press. 7

Gandhi, Amit and Jean-François Houde (2019): "Measuring substitution patterns in differentiated-products industries," *NBER Working paper w26375*. 7, 9

Kitschelt, Herbert and Daniel M. Smidt (2013): "Democratic Accountability and Linkages Project (DALP)," https://sites.duke.edu/democracylinkage/. 7

Olea, José Luis Montiel and Carolin Pflueger (2013): "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 31 (3), 358–369. 8

Rousseeuw, P.J. (1987): "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis," *Computational and Applied Mathematics*, 20, 53–65. 16