## 0.1 Legislative Candidates

Brazil's electoral system puts individual candidates at the center of political choice. Indeed, the literature notes that (i) parties are weak, under-resourced, and often unable to constrain opportunistic behavior by individual legislators (**?**, **?**, **?**); (ii) open-list PR and a lack of formal mechanisms channeling resources to congressional party leaders promote candidate-centric legislative careers (**?**, **?**); and (iii) Brazilian elections tend to be candidate-centric rather than party-centric, with voters effectively responding to candidate characteristics above party labels (**?**, **?**, **?**).

Understanding the drivers of voters' choices, therefore, requires that we analyze them at the candidate level. To that end, we bring together data on candidates running for a seat in the Câmara dos Deputados in the 2006, 2010, and 2014 elections. In total, across these three elections and all 27 legislative districts, there were 15,698 candidate-years: 4,944 in 2006, 4,887 in 2010, and 5,867 in 2014. For each candidate running for office, we observe the number of votes obtained by the candidate in each municipality, along with a rich set of individual characteristics including their previous professional and political experience, level of education, and gender.[1] For 10,021 candidate-years, we are also able to obtain a measure of their policy positions using individual campaign contributions and the methodology outlined in **?**. We describe this in detail below.

<span style="color:red">revisar desde aca</span>

Figure **??** provides summary statistics of candidates' observable non-policy characteristics.[2] Overall, we find that incumbents constitute only a fraction of all candidates,but are disproportionately represented among candidates who secure a seat in the chamber. While only about half of the candidates have higher education, this figure increases to about 75% for elected candidates; women compose only about a quarter of total candidates, but an even far lower percentage of elected candidates; candidates with business or government (bureaucratic) experience make about 10% of the pool of candidates, and they represent a significantly lower proportion of elected candidates.

*NEED TO INCLUDE*

Education, professional experience, gender, and other valence attributes of candidates are clearly important to voters. Psolicy positions candidates adopt matter as well. This has

---

[1]This information is available from the *Tribunal Superior Eleitoral* (TSE).

[2]following standard practice, we refer to these non-policy attributes as valence.

been documented in the U.S. and elsewhere.[3] In Brazil, there is growing evidence of the strengthening of programmatic parties and policy position **?**. These defined policy positions can alter voting behavior, but defining the extent to which voters respond to these changes has been fraught by a set of challenges which we address in this exercise.

Estimating how voters' preferences for candidates vary with the policies endorsed by legeislative candidatess requires a measure of both elected and non-elected candidates' policy positions. Unfortunately, there are currently no available measures of both incumbents' and challengers' policy positions for Brazilian legislative elections.[4] To address this gap, we adopt the estimation approach of **?** and produce our own estimates of candidates' policy choices, using micro-level data on campaign contributions for 2004-2014.[5] The key assumption is that a contributor's marginal benefit of giving to a particular candidate is decreasing in the distance between the contributor's ideal policy and the candidate's choice. Contributors weakly allocate more money to candidates that are closer to their ideal point, which in turn allows us to rank candidates' positions in the policy space. [6]

The intuition for estimation is as follows. For each candidate $j$, she receives a vector of campaign donation from a subset of voters. Intuitively, if two political candidates receive donations from the same ssubset of voters, in the same amount, we presume them to be close to each other deologically. Correspondance analysis allowss us to rduce thees multidimensional patterns of campaign contributionss into a one-dimensional Estimation is carried out with an augmented correspondence analysis with a two-way frequency matrix, where rows

---

[3]See, e.g, **?**, **?**, and **?**.

[4]**?** and **?** estimate *incumbents*' ideal points using surveys that ask them to place themselves and all the main political parties represented in the legislature on a left-right, ten-point scale. As these scholars note, estimation of incumbents' ideal points via DW-Nominate is challenging due to widespread "vote-buying" of legislators through pork-barrel spending and cabinet allocations.

[5]In U.S. data, Bonica's estimates closely match ideal point estimates obtained using row call data.

[6]While Bonica interprets these estimates as politicians' preferred policies, we only make the assumption that these are the candidates' policy choices, which could or not correspond to their true preferences.

correspond to unique contributors and columns to candidates.[7] Each element in the matrix is the total amount of contributions made by contributor $i$ to candidate $j$, for the time span of our data. We then perform a singular value decomposition to retrieve ideal points for contributors and recipients.

To implement this approach, we use data on micro-level dyadic contributions, which include both individual and corporate donations.[8] Since corporations may donate to candidates strategically to secure access, we exclude them from our data and focus on private contributions by non-partisans and non-politicians. In total, we leverage over 650 thousand unique contributions at the federal level, and 3.8 million unique contributions at the local level.[9] Because many non-viable candidates tend to receive a small number of contributions, we are forced to drop a sizable number of candidates from the database. Nevertheless, our final sample includes 10,752 candidates across the three elections.[10]

We perform a battery of sanity checks of the external and internal validity of our candidate

---

[7]Correspondance analysis is a statistical method that decompoess multidimensional contingency matrices into a set of eigenvectors, usualy

[8]The campaign contribution data is available since the 2002 election, when the TSE mandated the disclosure of electoral campaign contributions to candidates at all levels of government.

[9]Under-the-table donations—*caixa dois*—are common, but previous research using the same data shows that officially-declared donations capture the majority of campaign contributions (**?**).

[10]The candidates for whom we are able to recover policy positions make an overwhelming fraction of all candidates seriously contending for a seat in the Câmara dos Deputados—see Figure **??** in the Appendix. In fact, only 0.02% of candidates for whom we don't have policy data were ultimately elected. Table **??** in the Appendix summarizes coverage of the final dataset by state and electoral cycle.

policy estimates. The left panel of figure **??** in the Appendix shows that there is a strong correlation between policy positions within the same party at both the local and federal level. The right panel of the same figure shows that our policy estimates are correlated with the ideology scores estimated by **?** at the party level. Figure **??**, on the other hand, shows that our estimates capture the leftward ideological shift of voters and parties in the 2000s found in Latinobarometer surveys. and here add the new robustness checks gali made for the referee.

In the next section, we use this information on candidates' valence characteristics and policy choices, along with election results, to estimate voters' preferences. The key for doing this, of course, is that voters can in principle give their vote to any candidate in the district but choose someone with particular attributes. Another alternative that is de facto available to voters is to abstain or to cast a void vote. This "outside option" is thus effectively competing with all the candidates for votes. As Figure **??** illustrates, this, in itself, is a formidable alternative. The 29% average abstention rate and 8.6% average blank vote rate in what is formally a compulsory voting system provide suggestive evidence that voters are not enthusiastic about the candidates they face.

*NEED TO INCLUDE*