

Bose, Paul

**Conference Paper**

## Political (self-)selection and competition: Evidence from U.S. Congressional elections

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Bose, Paul (2021) : Political (self-)selection and competition: Evidence from U.S. Congressional elections, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242377>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Political (self-)selection and competition: Evidence from U.S. Congressional elections \*

Paul Bose<sup>†</sup>

This version: October 2020  
see [most recent version](#)

## Abstract

How does competition affect the entry and selection of politicians? I use data on U.S. Congressional primary and general elections for the years 1998-2014 to study this question. I measure quality using previous legislative experience and the novel “identity match” quantifying how well candidates demographically represent their district. To identify causal effects, I rely on variation in competition caused by demographic changes resulting from decennial redistricting. Difference-in-difference estimates reveal differences between the electorally dominant and weak party. They show that experienced candidates avoid competition in primary elections in the strong party. As opposed to this, experienced candidates and candidates with a good identity match run relatively more frequently in primary elections in the weak party as competition increases. The effects of competition and entry overall cancel each other out so that there are no effects on the quality and identity match of the eventual winner of the general election.

**JEL:** *D72; D78; J45; P16*

**Keywords:** *Political competition; Political selection; Quality of politicians; Descriptive representation; Redistricting*

---

\*I am grateful to Dana Sisak and Olivier Marie for continuous guidance and support throughout the project. I would like to thank Sacha Kapoor, Bas ter Weel, James M. Snyder, Benoit Crutzen and Michal Šoltés for many helpful comments and suggestions. I am indebted to Daniel Jones and Randall Walsh for graciously sharing data and code with me. Finally, this paper has benefited greatly from discussions with participants at the Spring Meeting of Young Economists in Brussels (2019), EEA Rotterdam (2020) and seminars at the Erasmus University Rotterdam and the PhD Economic Virtual Seminar.

<sup>†</sup>Erasmus University Rotterdam, Burg. Oudlaan 50, 3062 PA Rotterdam, [bose@ese.eur.nl](mailto:bose@ese.eur.nl)

# 1 Introduction

How does competition affect political selection? This question is of particular importance for two reasons. First, politicians' identity and ability play a large role in determining policy outcomes. When women are elected to office for instance, spending increases on public goods that are arguable valued by women (Chattopadhyay and Duflo, 2004; Baskaran and Hessami, 2019). Similar results hold for officials belonging to ethnic or racial minorities (Pande, 2003) and more educated politicians are more likely to select policies that are more growth friendly (Besley et al., 2011). Second, political competition likely functions to improve the quality of elected politicians (Stigler, 1972; Wittman, 1989; de Paola and Scoppa, 2011). This could be driven by voters, who are better able to select on quality in competitive elections or by parties who anticipate voters' choices and select better politicians to run (Galasso and Nannicini, 2011). A third dimension of political selection and competition has largely remained unstudied: How do politicians' themselves react to competition? That is, how does competition affect the entry of high quality candidates? There are many reasons for the lack of evidence regarding this question. Most importantly, the intention to run for office is often unobserved in particular if the political parties play a large role in choosing who to field for the next elections, such that the pool of potential candidates from which the parties and later voters choose remains unseen.

In this paper I make use of data on more than 5,000 candidates in primary elections to the U.S. House of Representatives for the years 1998-2014. Primary elections offer the advantage of allowing researchers to observe not only the eventual party nominee that competes in the general election against nominees from other parties, but also the field of candidates she is chosen from.

A second difficulty in studying the effect of competition on the quality of the candidate pool is that candidates in different parties are likely to respond differently to it (Dal Bó and Finan, 2018). Indeed even in a two party system, a candidate in the electorally strong party (the one which is projected to win) will view an increase in competition as a decreased chance of winning the general election, while one in the weaker party will benefit from the increase in competitiveness. This makes it necessary to measure ex ante contestability to classify parties into electorally strong and weak ones.

Finally, electoral returns are directly related to candidates' characteristics and hence equilibrium outcomes. The analysis therefore requires exogenous variation in electoral competition. To solve both these issues, I predict the ex-ante vote shares of parties based on demographic district characteristics. I define strong and weak parties and make use of changes to district demography caused by decennial redistricting to measure exogenous changes to the predicted vote share of the strong party (*StrongPartyShare* – the party with an ex ante predicted share of more than 50%) (Jones and Walsh, 2018, first exploit this variation to district demography in a different context).<sup>1</sup> Testing the validity and

---

<sup>1</sup> Jones and Walsh (2018) study the influence that district partisanship has on the voting behavior of

relevance of this approach, I find that changes to district-level demographic characteristics are strongly predictive of future vote shares and hence a good measure of competition, but changes in my sample are unrelated to previous electoral returns and are thus exogenous to electoral competition prior to redistricting.

To measure candidates' quality or "valence", I exploit two characteristics – previous experience and a novel measure "identity match". Previous experience is widely used in the literature on elections to the U.S. Congress (Jacobson and Kernell, 1981; Jacobson, 1989; Carson and Roberts, 2005) to proxy for job-specific human capital (Hirano and Snyder, 2014). It hence likely matters for on-the-job performance once a candidate is elected. I show that candidates with previous experience (co-)sponsor more bills and record fewer absences in roll call votes as compared to those without and furthermore perform better in both primary and general elections. Identity match measures the how likely voters in a district are to identify with a given candidate based on shared demographic identity. I use information on candidates' ethnicity or race, gender and a proxy for their income to estimate an "ethnicity match", a "gender match" and an "income match" for each candidate and district.<sup>2</sup> Since identity matters both for policy choices (Pande, 2003; Chattopadhyay and Duflo, 2004; Baskaran and Hessami, 2019) as well as executive tasks (Butler and Broockman, 2011; Broockman, 2013; Lowande et al., 2019) all three measures for the identity match proxy for the choices made in office beyond party lines. I show that indeed candidates with better ethnicity and gender matches are more likely to vote against a majority of their party in roll call votes, consistent with voting behavior more favorable to the districts constituents rather than the party.<sup>3</sup> The electoral returns to a good identity match are less clear however. In primary elections of the strong party candidates benefit from shared identity in terms of gender and ethnicity or race while candidates with a better identity match (in terms of ethnicity or race and income) receive fewer votes in the weak party. I argue that the difference in the returns to shared identity is most likely due to the different set of voters in the strong as well as the weak party's primary election. A candidate with a shared identity with many voters in the district is likely to share characteristics with voters in the strong party, but less likely to share characteristics with voters in the weak party. In the smaller party, a good identity match should then be interpreted as a "negative quality". Similar to cues on gender and ethnicity (Mcdermott, 1998; Conroy-Krutz, 2013; Casey, 2015) shared identity matters more in low

---

representatives in the House of Representatives. In particular they show that as districts become more dominated by Democrats, roll call voting becomes more in line with Democratic policy, both because a Democrat is more likely to be elected, but also because Republican representatives in such a district adjust their voting to fit with their constituents.

<sup>2</sup> In particular I define the ethnicity match to be the share of voters with the same race or ethnicity as the candidate, the gender match to be the share of constituents in the labor force with the same gender as the candidate and the income match to be 1 minus twice the share of voters between the income group of the candidate and the income of the median voter.

<sup>3</sup> Similar results do not hold for representatives who are better representative of the median voter in terms of income however.

information primary elections than in general elections when voters are more informed on candidates policy preferences. Nonetheless, candidates who share their race or ethnicity with more voters in the district also perform better in general elections.

Using a difference-in-difference approach in which changes in *StrongPartyShare* measure treatment intensity, I show that experienced candidates in the electorally strong party avoid competition and enter relatively more frequently as the predicted vote share of the strong party increases. A 1 percentage point increase in *StrongPartyShare* is associated with an increase of average experience in the strong party of 1.95%. As opposed to this, experienced candidates in the weak party enter more frequently as the general election becomes more competitive. Here a 1 percentage point increase in the predicted vote share of the strong party leads to a decrease in the average experience of candidates in the weak party by 5.21%. Results for the average identity match in the candidate pool differ substantially. While there is no effect of competition on the average identity match in the strong party in any dimension, in the weak party candidates with higher ethnicity and income matches (i.e. “worse” candidates from the perspective its voters) enter relatively more often as competition increases.<sup>4</sup>

I argue that the findings for both experience and identity match are consistent with theoretical models of candidate entry in which candidates differ in their outside option depending on quality (Dal Bó and Finan, 2018). Experienced candidates are likely to have higher opportunity costs and thus need to be compensated with an increased chance of getting elected to office. Therefore increased competition should reduce average experience in the strong but not in the weak party. No such logic applies to the identity match, which is driven only by differing district characteristics.<sup>5</sup> On this dimension there is no reason to expect differing outside options amongst candidates. Intuitively then, “better” candidates (those with a poor match to the district) find it beneficial to enter in the weak party in any case, however as the election becomes more competitive, weaker candidates (with a better match) become willing to enter thus driving down average valence.

Regarding the quality of the elected officials, I show that the negative effect of competition on the entry of experienced candidates in the strong party combined with the lower chance of these candidates to win the general election cancels out the positive effect of competition on the candidate pool in the weak party. Thus in systems in which candidates themselves are fully deciding when and where to run, there are no aggregate positive nor negative effects of competition after taking into account endogenous responses of the candidates. This is good news in particular with regard to recent trends of decreasing

---

<sup>4</sup> Note that the results are also consistent with an alternative explanation: As the district becomes more competitive, the smaller party voters in the weaker party are more representative of the district and hence a better identity match would be “less negative”. Unfortunately the data does not allow me to cleanly disentangle the two mechanisms from one another.

<sup>5</sup> In all regressions involving any measure for identity match I control for candidates gender and race or ethnicity and income, so that differences in the match are driven only by differences within gender-ethnicity-income clusters.

competition in elections in the United States (Abramowitz et al., 2006).

My study most closely relates to the literature on political selection (see e.g. Besley, 2005, for an early survey) and in particular its relation to competition (see Dal Bó and Finan, 2018, for a recent review of the literature). Galasso and Nannicini (2011) show that elected politicians in the Italian parliament are of higher quality (have on average 2 more years of schooling) if the lagged vote margin in their district was below 5%. De Paola and Scoppa (2011) show that similar results also hold for Italian local elections using an instrument for competition based on early termination of municipal councils. In a cross country study, Besley and Reynal-Querol (2011) show that Democracies, which display a relatively higher degree of competition, select more educated leaders than autocracies. Electoral competition is furthermore associated with lower rates of absenteeism (Galasso and Nannicini, 2011; Bernecker, 2014), higher intrinsic motivation (Gavoille and Vershelde, 2017) and perceived competence (Atkinson et al., 2009), the selection more growth friendly policies (Besley et al., 2010) and fewer political dynasties, which are often associated with less qualified candidates (Dal Bo et al., 2009). Different from these studies, my focus does not lie on elected politicians but rather on the candidate pool. That is, I focus not on the aggregate effect of competition on the selection of voters, political parties and candidate entry, but my setting allows me to study candidate entry and competition while abstracting from the other drivers of political selection.

I am aware of three other studies that investigate the effect of competition on the candidate pool. Dal Bó et al. (2017) make use of Swedish administrative data on elected and non-elected politicians and document a positive correlation between competition and the quality of the candidate pool. Since parties in Sweden must submit a list of candidates before the election, party selection may still play some role. In the setting I am studying, parties typically play little role so that I am able to identify the role of competition only on candidates' decision to run. Hirano and Snyder (2014) also study primary elections to the U.S. House of Representatives. Focusing on open-seat elections (e.g. after incumbent retirement), they show that primary elections are especially effective at selecting experienced politicians in non-competitive districts in the majority party. Hirano and Snyder (2014) also investigate mechanisms behind this finding. They argue that the improved selection in primary elections of the majority party is due to increased entry of experienced candidates either due to better incentives or due to a higher number of potential experienced candidates as well as increased ability of voters to select better candidates due to improved media coverage. My study differs from the work of Hirano and Snyder (2014) in three aspects. First, my identification strategy based on redistricting induced changes to competition allows me to consider all elections (open and incumbent-held seats) to the U.S. House of Representatives. I can thus show that competition matters more broadly for candidate entry even when facing an incumbent. Second and more importantly, I discuss the effect of competition on both skill (experience) and on the novel measure "identity match", which broadly proxies for policy preferences. Finally, I show that the effect of competition cancels out with regards to the eventually elected politicians. Carson et al. (2006) also consider

how competition changes due to redistricting affect the entry of “high quality” candidates. However their study focuses on nineteenth-century politicians in general elections whereas I discuss entry in primary elections in the more recent political environment.

More broadly, my study relates to research on the role of incentives in political selection. Beyond competition, scholars have mostly investigated the role of financial and non-financial returns to holding office and its effect on the quality of elected officials (Ferraz and Finan, 2008; Kotakorpi and Poutvaara, 2011; Gagliarducci and Nannicini, 2013; Dal Bó et al., 2017; Cerina and Deidda, 2017). Again my work differs from this line of research by analyzing the effect of incentive changes on the *entry* of high quality politicians rather than on elected officials.

Competition can function on multiple levels in affecting the quality and identity of elected politicians. Voters might be better able to select on these dimensions in more competitive elections, parties might choose different candidates and politicians themselves might choose whether or not to run for office. Precisely understanding the role of competition in each of these levels may aide us in designing institutional frameworks, selection procedures and even elections themselves so as to select better leaders. My study contributes to this understanding by shedding some light on how competition affects the self-selection of politicians into running for office.

## 2 Identification and empirical specification

Identifying an effect of competition on entry of high quality candidates requires exogenous variation in the level of competition. Better candidates are likely to affect election outcomes and hence district competitiveness. A simple cross-district comparison using electoral returns to measure competitiveness hence suffers from potential reverse-causality. I make use of a continuous-treatment difference-in-difference approach. The treatment variable measures changes to partisan composition. The measure was first developed by Jones and Walsh (2018) and is constructed using demographic data and predicting the share of voters that support the electorally strong party (*StrongPartyShare*) – the party with ex ante more than 50% predicted votes. I briefly describe the construction of the *StrongPartyShare* in the next section (see Jones and Walsh, 2018, for a more detailed description).

In the context of Congressional primary elections, it is furthermore far from straightforward to measure candidates’ quality. I make use of two valence characteristics that I discuss in detail in the next section – legislative experience and an identity match. Candidates’ legislative experience is used to proxy for job-specific human capital and hence performance in office (see e.g. Jacobson and Kernell, 1981; Hirano and Snyder, 2014). The identity match measures how well candidates demographically represent their district. Because identity affects policy decisions (Chattopadhyay and Duflo, 2004; Besley et al., 2011), voters are likely to prefer candidates that better represent their social group. Thus for both aspects of valence high values are likely to be preferred.

I exploit variation in *StrongPartyShare* from regular redistricting. Redistricting is done once every 10 years following the U.S. census. There are two redistricting cycles in my observation period (2000 and 2010). For each cycle, I make use of 2 elections before and 2 elections after redistricting. I compare candidate entry in each district and cycle between these elections. Importantly changes to *StrongPartyShare* only happen at the time redistricting. These changes are solely due to redistricting, as I use the data from the same census for before and after redistricting, such that the only change is from the redrawing of district lines.

Since the effect of competition is likely to depend on electoral strength I estimate the baseline regressions separate for the strong and weak party. Denoting by  $p \in \{W, S\}$  the electorally weak (strong respectively) party, I estimate the specification

$$V_{idct}^p = \alpha_t^p + \beta \text{StrongPartyShare}_{dct} + X_{idct}^p + \eta_{dc}^p + \varepsilon_{idct}^p \quad (1)$$

where  $V_{ict}^p$  captures candidate valence (i.e. experience or identity match) for candidate  $i$  in district  $d$  in redistricting cycle  $c$  at time  $t$  relative to redistricting for party  $p \in \{W, S\}$ .  $X_{idct}^p$  is a set of control variables,  $\eta_{dc}^p$  denotes district-redistricting-cycle fixed effect and  $\alpha_t^p$  time fixed effects. The coefficient of interest  $\beta$  measures the impact of changes to the share of the strong party due to redistricting on the valence of a candidate in the primary election. Given multiple candidates in the primary election, it can be interpreted as the effect on the average quality of the candidates.

The main threat to this identification strategy is that changes in partisan composition are not random since redistricting is done by state legislatures and the rules governing the procedure leave ample room to redraw districts to favor a certain party or candidate.<sup>6</sup> Partisan redistricting is a particular concern if parties respond to trends in candidate valence in their primary elections prior to redistricting by redrawing the electoral map in such a way so as to encourage or discourage particular candidates from standing for the election. That is, if trends in candidate valence in primary elections prior to redistricting cause larger changes in partisan composition of districts.

I deal with this threat in several ways, first I verify that the typical common trend assumptions hold. I do so by including an interaction between time fixed effects and changes to *StrongPartyShare*. If e.g. larger changes were indeed caused by low levels of candidate valence prior to redistricting, the estimate for the interaction at  $t = -1$  should reveal this. Second, I restrict the estimation sample to include only primary elections of parties that did not have full control over the redistricting procedure.<sup>7</sup> In a robustness check I also focus on only those states in which a split legislature, independent commission

---

<sup>6</sup> Redistricting with partisan interests in mind is termed gerrymandering. Gerrymandering has received a lot of scholarly attention (Friedman and Holden, 2008; McCarty et al., 2009; Ansolabehere and Snyder, 2012) and has been shown to have a significant effect on electoral competition (Lindgren and Southwell, 2013; Carson et al., 2014).

<sup>7</sup> Parties do not have full control if the other party controls the procedure, if the legislature is split or if an independent commission or court ruling determines the redrawn lines.



or court ruling determined the new district lines. Third, in all specifications, I control for previous election outcomes in the same district, to alleviate the threat that redistricting might have been done to e.g. motivate the winner of the previous election to run. Finally, I additionally present all results with and without including incumbent candidates to ensure that the results are not driven by incumbent favoring redistricting, done to sway incumbents to retire or stand for an other term.

### 3 Data

I make use of data on Congressional primary and general elections for the years 1998-2014 on the candidate level. The original sources are the “America Votes” book series<sup>8</sup>, the Clerk of the U.S. House of Representatives<sup>9</sup>, the Federal Election Commission, the U.S. census<sup>10</sup> as well as many online resources in order to hand-collect information on candidates’ legislative experience and characteristics<sup>11</sup>. I furthermore collect data on partisan composition in U.S. states and redistricting control at the time of redistricting (Levitt, 2018). My data contains information on candidate name, gender, legislative experience, filing address as well as performance in the primary and potentially general election. I furthermore add estimated information on gender (if this was missing) and race/ethnicity based on candidates’ first and last names.<sup>12</sup> Finally I make use of the postal address candidates report to the Federal Election Commission to broadly proxy their income with ZIP code level median income.

#### 3.1 Measuring competition: *StrongPartyShare*

Given the potential reverse causality between electoral competition and candidates’ valence stemming from higher electoral returns for better candidates, it is important to measure competition without the use of electoral returns. To do so, I follow closely Jones and Walsh (2018) to predict the share of voters for each party using demographic characteristics of a district and exploiting the changes due to redistricting. The construction of *StrongPartyShare* involves 5 steps. First, using county level voter registration data (Leip,

---

<sup>8</sup> The data for primary elections between 1998 and 2010 have been digitized by Pettigrew et al. (2016), for 2012 and 2014, I digitized the results using the same original source (Cook, 2013, 2015).

<sup>9</sup> The data for general election returns was digitized by the MIT Election Data and Science Lab (2017).

<sup>10</sup> Data from the U.S. census is made available through NHGIS (Mason et al., 2019).

<sup>11</sup> Pettigrew et al. (2016) collect data for candidate experience using Google searches as well as LexisNexis for the years 2000-2010. I extend the available years using data from the candidate database ourcampaigns.com, the CNN election archive for 1998, 2002 and 2004, the newspaper archive Factiva and LexisNexis and Google searches.

<sup>12</sup> I predict race/ethnicity based on name using the python package ethnicolr by Suriyan Laohaprapanon and Gaurav Sood which relies on census data as well as data on first and last names from Wikipedia which was collected by Ambekar et al. (2009). For gender prediction I make use of the Social Security Agency’s list of first names in the United States.

2013) and demographics, I estimate coefficients for the share of registered Democrats.<sup>13</sup> Next, I use historical boundaries of congressional districts (Lewis et al., 2013) to map census block groups into the district lines in order to generate demographic information for each congressional district pre- and post-redistricting. In a third step, in order to match pre- and post-redistricting districts, I calculate the population overlap between these districts.<sup>14</sup> I then use district level demographic data from the second step as well as the coefficients from the first step to calculate district level predicted shares of Democratic voters both before and after redistricting. In a final step I calculate the share of the ex-ante strong (predicted vote share of more than 50% prior to redistricting) party *StrongPartyShare* using the predicted share of Democratic voters.<sup>15</sup> Importantly, the procedure relies only on demographic changes due to redistricting and not on changes over time, as block groups and demographic information within a redistricting cycle remain constant. That is for the years 1998-2004 I use census data from the year 2000 and from 2010 for the redistricting cycle 2008-2014.

[Figure 1 about here.]

Figure 1 summarizes the geographic distribution of changes to *StrongPartyShare*. Note that districts without any changes function as “control” districts. Redistricting intensity does not seem to follow a particular geographic pattern. Furthermore, there does not seem to be any clear evidence of redistricting bias in favor of the strong party.

I discuss the relation between *StrongPartyShare* and election returns in section 4, however note that since the procedure does not make use of any election data, it does not suffer from the typical reverse causality problem. Furthermore, different from Jones and Walsh (2018) I do not make use of any demographic information on race or ethnicity and gender in the construction of *StrongPartyShare* to ensure that this does not threaten the construction of the identity match measure based on these characteristics.<sup>16</sup>

---

<sup>13</sup> The demographic variables used to predict the share of Democrats are: population density, the share of urban population, those under 18 and over 65, the share of single male households, single female households, households with a married couple with children, households with a married couple without children, single male head households with children, single female head households with children, single male head households without children, single female head households without children and non-family households with a male head of household. Furthermore the share of houses which are vacant and renter owned. I estimate 4 separate models for each region in the U.S.. The estimated coefficients are presented in the appendix in table A1.

<sup>14</sup> Districts are only matched if they are both a best match for the pre-redistricting and the post-redistricting district.

<sup>15</sup> *StrongPartyShare* is the predicted share of Democrats if Democrats are the strong party. If Republicans are the ex ante strong party it is 1 less the predicted share of Democrats.

<sup>16</sup> In the appendix I show that the results do not differ if I include race and ethnicity as well as gender information into the construction of *StrongPartyShare*.

### 3.2 Measuring valence: Experience

The prediction of *StrongPartyShare* enables a clear measurement of the treatment. Analyzing how changes in competition affect political selection, i.e. the entry of high quality candidates, furthermore requires reliably available information on candidates valence. I make use of previous legislative experience in order to proxy for job specific human capital.

Previous experience in elected office has frequently been used to account for the quality of candidates in studies of Members of U.S. Congress (Jacobson and Kernell, 1981; Jacobson, 1989; Hirano and Snyder, 2014). Indeed multiple studies have found that experienced candidates tend to do well at the ballot (Carson and Roberts, 2005; Buttice and Stone, 2012). Hirano and Snyder (2014) argue that this advantage is most likely due to accumulated job-specific human capital as opposed to advantages due to experience in running elections. They show that candidates who have previously held executive posts are more likely to win primary elections for executive positions, while legislators are most likely to win primary elections for legislative office. Indeed elected officials with previous legislative experience are not only more likely to win their primary elections, but furthermore differ in their performance in office. Figure 2 shows the number of bills sponsored or co-sponsored by experienced and inexperienced representatives over the course of their Congressional career between 1973 and 2017. Throughout the careers of the politicians, experienced candidates tend to be more active in sponsoring and co-sponsoring bills in Congress. Similar results hold also for vote participation rates in roll call votes which proxies for absenteeism (see figure A1 in the appendix).<sup>17</sup>

[Figure 2 about here.]

Importantly, information on previous experience is readily available to voters since local and national newspapers frequently report on candidates for the U.S. House in both primary and general elections. Given the human capital advantage of experienced candidates, we might expect these candidates to perform better in both primary and general elections as well as in office. It is hence important to understand how such high valence candidates react to incentives and in particular to competition when deciding on whether or not to run for elected office.

Figure 3 summarizes the distribution of experience in my data. It presents the average experience of candidates in the weak and in the strong party's primary elections as well as how they gathered this legislative experience. Most experienced candidates are incumbents, but a substantial amount also gathered experience through state and city offices. Overall candidates in the strong party tend to be more experienced but this is mostly explained by a larger share of incumbents.

[Figure 3 about here.]

---

<sup>17</sup> There are many reasons for absenteeism. Possible explanations include more visits in the representative's district, other legislative activities but also slacking off.

### 3.3 Measuring valence: Identity match

To understand how competition affects entry of high valence candidates across different dimensions, I also propose a novel measurement of valence – “identity match”. Using the available information about candidates in the primary elections, I define the identity match along three characteristics. First, I use information on candidates’ race or ethnicity and calculate the share of the population in the district that shares the same racial or ethnic background as the candidate.<sup>18</sup> I define this share of the electorate as a candidate’s “ethnicity match”. Second, I define as “gender match” the share of adults in the labor force with the same gender as the candidate.<sup>19</sup> Finally, using a proxy for candidates’ income (median household income at the ZIP code level of their postal address), I define as “income match” 1 minus twice the proportion of voters between the median household income and the candidate’s income.<sup>20</sup>

All three forms of the identity match are likely to matter in two ways – representatives’ identity matters for policy making and voters prefer candidates with shared identity when making their vote choice. First, there is a large body of literature suggesting that identity matters for policy making. When female politicians are elected for example, spending on fields arguably valued by women tends to increase (Chattopadhyay and Duflo, 2004; Baskaran and Hessami, 2019).<sup>21</sup> Furthermore, politicians with a minority background are more likely to be intrinsically motivated to advance the interests of their ethnic group (Pande, 2003; Minta, 2009; Butler and Broockman, 2011; Broockman, 2013). Lowande et al. (2019) find that women and ethnic minority representatives are more likely to interact with the U.S. bureaucracy on behalf of voters that share their identity.<sup>22</sup> Regarding income, recent evidence suggests that wealth also matters in voting. In particular, wealthier representatives are more likely to vote conservatively mostly aligned with wealthier voters (Eggers and Klačnjaja, 2019). Second, voters generally care about electing a representative with a shared identity. Multiple studies have found that women prefer female candidates (Smith and Fox, 2001; Sanbonmatsu, 2002; Brians, 2005). Furthermore so called “eth-

---

<sup>18</sup> Since information on candidate gender and race/ethnicity often is unavailable, I make use of name based prediction for gender and race or ethnicity. This is a well established method in particular in the fields of health and population research. A far from exhaustive list of studies includes studies by Mateos (2007); Elliott et al. (2009) and Kandt and Longley (2018).

<sup>19</sup> Hence for women, this takes the value *share of women in district*  $\times$  *female labor force participation*, for men *share of men in district*  $\times$  *male labor force participation*.

<sup>20</sup> This definition captures the logic that a candidate is less representative if there are more voters between the candidate and the median voter. In order to ensure that the income match measure increases in how representative a candidate is and to scale it between 0 and 1, I subtract this proportion twice from 1.

<sup>21</sup> Similar results hold also for male politicians, though the effect of gender may not always extend to elected executives such as mayors (Ferreira and Gyourko, 2014).

<sup>22</sup> Similar results also hold for other aspects of representatives’ identity such as age (Curry and Haydon, 2018), veteran status (Lowande et al., 2019) and socialization (Washington, 2008).

nic voting” is well documented though it is not always distinguishable from partisanship (Wolfinger, 1965; Eifert et al., 2010; Adida, 2015; Houle et al., 2019). Ben-Bassat and Dahan (2012) link group identity to turnout and find that voters who share a common social identity with candidates are more likely to vote. A link between representatives’ income and voters’ preferences is less well established. Thompson et al. (2019) show that representatives in the United States of the mid to late 20th century are mostly economically unrepresentative of their voters. They attribute this to both to voters preferring individuals with more success in the private sector and to entry barriers for poorer candidates. One might nonetheless expect that voters care about income, however if voters prefer wealthier candidates, being economically representative might actually hurt candidates.

It is worth pointing out that in the construction of the gender match, I make use of female and male labor force participation rather than only gender composition of districts. There are two reasons for this. First, focusing on the share of the population in the labor force that shares the gender of candidates reduces the size of the group with shared identity which likely increases identification (Ben-Bassat and Dahan, 2012). Indeed in particular educated women, which are more likely to be in the labor force, are prone to vote for female candidates (Smith and Fox, 2001). Second, higher female labor force participation might increase demand effects for female candidates to run for office, so that more men might also prefer to vote for women (Iversen and Rosenbluth, 2008).

[Figure 4 about here.]

Figure 4 provides some evidence on which dimensions of shared identity can proxy for important aspects of candidates voting decisions in Congress. The panels shows the residual share of roll call votes in which representatives voted with the majority of their party after controlling for gender, race or ethnicity and income. The top panel shows that a representative’s support for the majority position of her party decreases with ethnicity match. A similar result holds for the gender match as the panel in the center shows. No such effect exists for the income match (bottom panel). Overall there is clear evidence to suggest that representatives with better identity matches are less likely to vote with the majority of their party. Such voting behavior is consistent with representatives being more likely to vote against party lines in favor of their constituents.

Figures A2, A3 and A4 in the appendix provide distributional information for the three dimensions of the identity match considered here. Figures A5, A6 and A7 furthermore show that the candidate pool is overall more white, male and richer than the average constituent. The latter suggests that inline with the findings of Thompson et al. (2019) the decision to run for office may be subject to financial barriers. Nonetheless, it is important to understand what role self-selection with relation to competition plays to create this relatively unrepresentative picture of constituents.

### 3.4 Estimation sample and baseline covariates

As argued before, my identification strategy relies on variation in *StrongPartyShare* stemming from redistricting. Since the changes might be non-random and strategically related to entrance of high valence candidates if a party has full control over the redistricting procedure, I restrict my estimation sample to only primary elections in districts in which the party holding the election did not have full control over the redistricting procedure. This means that it must have made some compromises (as in the case of a split legislature or when a bi-partisan commission redraws the lines) or did not have any influence over the procedure at all (e.g. when the other party controls the procedure, the redistricting is done by a non-partisan commission or courts mandated the new lines). For all results involving the general elections, I focus on elections in which neither party controlled the redistricting procedure. I furthermore discard all races in states in which primary elections do not follow the typical two-party structure (but e.g. as in California, a so-called chaos primary in which all candidates face off against each other).<sup>23</sup> I similarly exclude elections in states in which candidates were chosen by means of a party convention or caucus as opposed to a primary election.<sup>24</sup> Next, I exclude primary elections in Maine for the years before 2008, since state law required that redistricting takes place in the third year following the census, such that changes to the district composition appear later than in the other states. Finally, I exclude all elections taking place in Louisiana, where primary elections and general elections are not separated. This leaves me with a total of 5,453 primary election candidates in 1470 general (and corresponding primary) elections between 1998 and 2014.

[Table 1 about here.]

Table 1 summarizes the covariates used in the following analysis as well as some information about the elections in the sample. Columns (1) and (2) report the mean as well as standard deviation of the respective variable on the left. Overall there are between 15 and 18% female candidates and even fewer non-white candidates (between 7 and 8%). The average age (for those candidates for whom age data is available) is 54.3 in the strong party and 52.2 in the electorally weak party. Candidates in both parties tend to file for election in similarly wealthy ZIP codes, which can be used as a rough proxy for candidates' income. Column (3) reports the coefficient of the estimation of equation (1) with the variable on the left as outcome. Column (4) shows the respective standard error. If the changes of *StrongPartyShare* are systematically related to any of the covariates (e.g. if redistricting was done to increase the chance of open seats) the estimate in column (3) should show this. Columns (3) and (4) can thus be interpreted as a “balancing test”. The table does not suggest any systematic relationship between the covariates and the treatment. Despite this

---

<sup>23</sup> I exclude California and Washington from 2012 on-wards, after which year both states adopted a chaos primary.

<sup>24</sup> This amounts to the exclusion of most elections in Connecticut, Virginia and Utah.

finding, I always control for lagged election outcomes to ensure that potential redistricting in order to motivate a previous winner to run does not affect the results. Furthermore, in all regressions involving a form of identity match I control for candidates' gender, race or ethnicity and proxied income in order to ensure that any potential effect does not stem solely from differences in those candidate characteristics.

## 4 Results

The aim of this paper is to understand how high valence candidates react to changes in competitiveness with their decision on whether or not to enter into electoral races for office. I study two dimensions of valence, experience and identity match.

### 4.1 *StrongPartyShare* and electoral returns

In order to credibly answer this question using *StrongPartyShare* to measure competition changes, two things need to be verified. First, it is important that *StrongPartyShare* does not systematically depend on any pre-existing conditions in the district and state. Second, *StrongPartyShare* should be related to actual election outcomes and hence be relevant to electoral competition.

[Figure 5 about here.]

Figure 5 shows the distribution of changes to *StrongPartyShare* depending on the values of *StrongPartyShare* immediately before redistricting took place. Importantly, despite a small trend to 0.5, whereby a 1 percentage point increase in *StrongPartyShare* immediately before redistricting is associated with a 0.044 percentage point decrease in changes to *StrongPartyShare*, no clear pattern emerges. Similar results hold for the relation between the strong party's share of seats in the state house and senate at the time of redistricting and changes in *StrongPartyShare*, as figures A8 and A9 in the appendix show.

[Table 2 about here.]

To ensure that *StrongPartyShare* reliably measures competition, I estimate equation (1) using electoral returns in the general election as outcome measure. Table 2 shows the corresponding regression analysis. The results show that a 1 percentage point increase in *StrongPartyShare* results in an increase in the vote share of the strong party of about 0.6 percentage points. Changes in *StrongPartyShare* are thus highly predictive for changes to electoral returns. Figure A10 in the appendix furthermore verifies that the common trend assumption holds in this case.

## 4.2 Electoral returns to valence

Before focusing on the effect of competition on entry of high valence candidates, I discuss the effect of valence on electoral returns. That is, do high valence candidates profit from their experience or good identity matches in primary or general elections?

To answer these question I run candidate-level OLS models for the electoral returns of a candidate in primary and general elections respectively. The main explanatory variables of interest are a candidate's previous legislative experience and identity match in three dimensions – ethnicity, gender and income match. I always control for gender, race or ethnicity, income as well as for the competition in election. In particular, I include fixed effects for every observable combination of the number of experienced and inexperienced competitors. That is the variation I exploit stems from comparing candidates that face the same number of experienced and inexperienced candidates in their primary election. I furthermore include election-year fixed effects and controls for incumbent status as well as the type of primary election. At a maximum I also include state fixed effects.

[Table 3 about here.]

Table 3 presents the estimates for candidates' vote share in primary elections as well as for the probability of winning the election separately for the weak and the strong party. As expected, voters seem to value experience. Candidates in the strong party that have previously been serving in elected office receive on average 11.1-11.5 percentage points more votes than those without. In the weak party the effect is slightly smaller at about 10.7-10.8 percentage points but remains highly significant at the 1% level. The advantage in terms of vote share also translates into higher probabilities of winning the primary election as the bottom panel of the table shows. Here the advantage lies between 18.1 and 21.8 percentage points. Thus experienced candidates are about 39%-43% more likely to win their primary elections as compared to a mean probability of between 47% and 51% for candidates with no prior experience.

The results for candidates' identity match reveal a different picture. While it seems to be an advantage to be of a the same gender and race or ethnicity as a large share of the electorate in the advantaged party's primary elections, candidates with better identity matches in terms of race or ethnicity receive smaller vote shares in the weak party. For every 1% of voters who share the race or ethnicity of a candidate in the strong party, this candidate receives on average between 0.02 and 0.05 percentage points more votes and is between 0.04 and 0.07 percentage points more likely to win the primary election. In the weak party a 1 percentage point increase in the ethnicity match reduces the vote share by about 0.02 to 0.03 percentage points. Candidates with a better match in terms of income also receive fewer votes in the weak party but there is no effect in the strong party.

To interpret this finding, it is useful to remember why the identity match might be an electoral advantage – voters prefer to vote for candidates that are similar to themselves. Furthermore voters in primary elections belong mostly to a candidate's own party whose



demographic characteristics may thus differ substantially from those of all voters in the district. Indeed race- or ethnicity-based voting often aligns with voting for a certain party<sup>25</sup> (Michelson, 2005) and women are more likely to vote for a Democratic candidate than men (Cascio and Shenhav, 2020). This means that if a candidate is similar to voters in the district, she is also likely to share characteristics with voters in the strong party, but might not share any with those supporting the weak party.<sup>26</sup> In the absence of strategic voting, the results from table 3 are fully consistent with this interpretation, as in such a case being more similar to the district might indeed be a disadvantage in the weak party.<sup>27</sup>

The results in table 3 include uncontested as well as contested primary elections, because advantages in experience and shared identity might not only increase a candidate’s vote share in contested elections but also dissuade other candidates from entering the election. The estimates in table 3 combine the two effects. In table A2 in the appendix I show that the effect of both shared identity and experience remain qualitatively unchanged in contested elections.

[Table 4 about here.]

In the general election, there are no reasons to expect any differences between strong and weak party as they face the same electorate. Table 4 presents the coefficients for the corresponding analysis of the electoral returns of primary winners in the general election. In this analysis I additionally control for the predicted vote share of the candidate’s party (*StrongPartyShare* for the strong party and  $1 - \textit{StrongPartyShare}$  in the weak party). As in the primary electorate, voters prefer candidates with legislative experience also in the general election. In particular, the vote share of experienced candidates increases on average by 7.8 percentage points from an average of 49.2%. This corresponds to an increase

---

<sup>25</sup> Cuban-Americans for example are much more likely to vote Republican whereas Mexican-Americans are likely to vote for Democrats.

<sup>26</sup> Empirically testing this hypothesis with the available data is difficult. One way would be to exploit variation in primary election laws. In states that make use of open primaries in which theoretically any voter is allowed to participate in a party primary election we would expect to see smaller effects of high identity matches (fewer voters for the candidates in the strong party and more for those in the weak party) due to cross-over voting. However empirical evidence suggests that even in presidential primary elections, cross-over voting is relatively low (Hedlund, 1977). Similar results hold in experimental settings (Blackwell and Calcagno, 2019). I nonetheless estimate the OLS models separately for open and closed primary elections. The results (not shown here) go in the expected direction however are never significant and should not be over-interpreted due to the limited data availability and likely low share of cross-over voting.

<sup>27</sup> Under strategic voting, voters in the weak party might recognize that candidates with higher levels of the identity match would perform better with the full electorate and hence select them more frequently. However experimental evidence suggests that there is very little strategic voting in party primary elections (Cherry and Kroll, 2003). Furthermore, as I show below, good identity matches only come with moderate advantages in the general election, possibly because it matters mostly in low information elections.

of about 16% over candidates without prior experience. Again these advantages in vote shares translate into substantial benefits in terms of the chance to win the election.

The estimates for shared identity are marginally significant in the dimensions ethnicity or race and income in at least some specifications. The sign for shared ethnicity consistently suggests a small advantage for candidates who are more similar to their district, as would be expected. For the income match, table 4 reports negative estimates. This is consistent with broad evidence that voters tend to prefer wealthier candidates in the elections over those more economically representative (Thompson et al., 2019). Overall, the reduction of the effect size is not very surprising given that racial and gender cues are most often used in low information elections, such as primaries (Mcdermott, 1998; Conroy-Krutz, 2013; Casey, 2015). In general elections on the other hand, voters often observe detailed policy proposals as well as endorsements and other indicators for candidates' positions and qualities.

As for primary elections, I include all contested and uncontested elections in the estimation sample, however results remain substantially unchanged when focusing only on contested elections as table A3 in the appendix shows.

Overall, the estimations show that candidates with previous legislative experience benefit substantially from their experience both in primary and general elections in either party. On the other hand, similar identity only seems to benefit candidates in the strong party whereas it might even be a disadvantage for politicians in the weak party. The latter is driven by the primary disadvantage of these candidates.

### 4.3 Competition and entry of experienced candidates

Having established that experience as well as shared identity matters both for voters' choices as well as for policy choices of politicians, I now focus on the question of how competition affects the entry of these candidates. That is, how does increased or decreased competition motivate or dissuade potential high valence candidates from entering? As detailed in section 2 I measure competition using predicted vote share of the strong party, which by definition has an ex ante predicted vote share of more than 50%, such that the estimates for an increase of *StrongPartyShare* should always be interpreted as the effect of a decrease in competition.

[Table 5 about here.]

Table 5 reports the results for the estimation of eq. (1) using a dummy for candidates' experience as outcome variable. Throughout all specifications, I report estimation results with and without control variables. In all estimations, I include the lagged election result as well as district-redistricting-cycle or district-redistricting-cycle-party fixed effects. The first two columns in the table presents results for all primary elections in both the strong and the weak party to give an idea of aggregate effects. Columns 3 and 4 include only primary elections in the strong party and columns 5 and 6 only those in the weak party. The top panel of the table includes all candidates. The bottom panel excludes incumbents from

the analysis to account for potential strategic redistricting aimed at incumbents decision to rerun for office. Furthermore incumbents are more likely to run than other experienced candidates independent of predicted competition and hence estimates might be lower when including incumbents.

The results indicate that experienced candidates in the strong party enter relatively more frequently than inexperienced ones as *StrongPartyShare* increases. Even when including incumbents in the analysis, a 1 percentage point increase in *StrongPartyShare* increases the probability that a candidate in the party’s primary election is experienced by about 1 percentage point. Based on the average share of experienced candidates in the strong party, this is an increase of about 1.95%. Hence candidates in the strong party tend to shy away from competition in the general election but instead rally relatively more frequently as competition decreases. In the weak party the opposite holds. The average experience of candidates decreases by 1.6 percentage points for every percent increase in *StrongPartyShare*. Given the much lower average, this effect size amounts to a decrease of 5.21%. Hence experienced candidates in the weak party tend to be more willing to enter into the primary elections as competition increases as compared to inexperienced candidates. When focusing on non-incumbents, the results are comparable overall, though the size of the effect increases substantially suggesting that incumbents might be less affected by changes to competition.

[Figure 6 about here.]

Figure 6 graphically presents a test of the typical difference-in-difference common trend assumption for average experience of the candidate pool. The figure shows the estimated coefficients for  $\Delta StrongPartyShare \times I_{[T=t]}$ , that is the change in *StrongPartyShare* interacted with time relative to redistricting, and otherwise using the specification of eq. (1) with  $t = -3$  as base category. The top panel uses all available data on candidates in the estimation sample. The lower panel shows the estimates when focusing only on non-incumbent candidates. Black dots refer to the estimates for the strong party, while blue ones refer to the weak party’s primary elections. The lines indicate 95% confidence intervals. The figure highlights two important aspects. First, there does not seem to be any pre-trend in terms of average quality in the candidate pool, neither in the strong nor in the weak party. This is reassuring and suggests that the sample restrictions sufficiently control for any possible strategic redistricting with regard to the candidate pool. Second, both for the full sample as well as for non-incumbents effects after redistricting are relatively stable suggesting that indeed the immediate change to competition is the driving factor.<sup>28</sup> All estimates are at least marginally significant (at  $p < 0.1$ ). The only exception to this is the estimate for  $t = 1$  in the full sample and the strong party, possibly because of noisy estimates due to low sample size as compared to the weak party.

---

<sup>28</sup> The slight increases in the estimates at  $t = 3$  might be explained by candidates learning better about the changes in competitiveness once they have observed an election.

Overall the results strongly suggest that experienced candidates in the strong party enter less frequently compared to inexperienced ones as districts become more competitive. In the weak party the opposite holds. A possible explanation for this behavior could be based on other income possibilities. Experienced candidates are likely to have better outside options than inexperienced ones. Indeed politicians often benefit from their position in the job market after they retire from public office.<sup>29</sup> Furthermore candidates with prior legislative experience often need to give up their elected office to run an effective campaign. Hence in order for them to enter into the race for the House of Representative more frequently, they would need to be compensated with an increased chance of winning the general election. Such a response to competition would be fully compatible with the reported results.<sup>30</sup>

#### 4.4 Competition and entry of candidates with shared identity

How do candidates with similar identity respond to competition? Importantly, since this trait differs from experience in the electoral returns to candidates, results on the decision to enter the race might be vastly different. In particular, while shared identity is an advantage in the strong party, it can be seen as a disadvantage in the weak party as the results above suggest.

[Table 6 about here.]

Tables 6, 7 and 8 present results for the estimation of eq. (1) with the three dimensions of candidates' identity match as outcome measure. As in table 5, the first two columns present aggregate effects for all parties.

Table 6 shows that overall the average ethnicity match decreases with a decrease in competition. This effect is driven by the weak party in which the average ethnicity match is decreasing by between 0.644 and 0.655 percentage points for every percentage point

<sup>29</sup> Evidence on the benefits from being elected are still mixed. Neither Kotakorpi et al. (2017) nor Berg (2018) find any evidence of increased wealth among elected politicians when compared to those barely loosing in Scandinavian countries. Eggers and Hainmueller (2009) on the other hand document sizable wage benefits as well as better outside job opportunities to British MPs between 1950-1970. Querubin and Snyder (2009) find similar evidence for U.S. Congressmen between 1850 and 1880.

<sup>30</sup> Testing whether the change in competition is indeed driven by outside options is tricky given that outside options are not observable. Furthermore, income data before entering or after leaving office is rarely available. It is thus not only difficult to verify that legislative experience changes the outside options of potential candidates, but also whether they drive the entry effect. In table A4 in the appendix, I test whether the entry effect is less pronounced for candidates above 66, the official retirement age in the United States. Indeed if outside options are a driving factor, we would expect the effect to be most pronounced amongst younger candidates with many career options. Table A4 shows that the entry effect reverses for old candidates and is driven by younger ones. This would be consistent with an explanation based on differing outside options. However given that age data is only available for a non-random sub sample of candidates, one should take these results with some reservations.

increase in *StrongPartyShare* depending on the specification. There does not seem to be a significant difference between the effect of competition in the entry decision of candidates with shared race or ethnicity and those without in the strong party, as the columns 3 and 4 indicate.

[Figure 7 about here.]

Figure 7 show a test for the common-trend assumption using the race or ethnicity dimension of the identity match as outcome. The figure shows estimates for eq. (1) with interactions between the change in *StrongPartyShare* and time relative to redistricting. The top panel includes all candidates. In the bottom panel I only use non-incumbents in the analysis. The figure shows that there are no differences in terms of trend prior to the redistricting procedure (highlighted by the vertical red line). Furthermore, there is no effect in the strong party. In the weak party, the effect of changes in *StrongPartyShare* immediately materializes after the redistricting and is remarkably stable over time.

[Table 7 about here.]

[Figure 8 about here.]

Table 7 and figure 8 perform the corresponding analysis for the gender dimension of shared identity. There does not seem to be any competition effect. That is neither women nor men in districts with many female (male) labor force participants seem to be entering more frequently when elections become less competitive as compared to those in districts with few participants with shared gender. While this result is somewhat surprising given the well documented difference in attitude towards competition between the genders (e.g. Niederle and Vesterlund, 2007), note that in the regressions for the gender match, I control for candidates' gender, such that any remaining effect is only explained by *shared* identity in terms of gender and labor force participation rather than candidates' gender alone.

[Table 8 about here.]

Analysing the third dimension of shared identity, table 8 presents the results for the income match. As for the race or ethnicity dimension, there does not seem to be any effect for entry in the strong party, but candidates in the weak party are on average better matches for their constituents as competition increases. A 1 percentage point decrease in *StrongPartyShare* leads to roughly a 0.76% increase in the income match. As figure 9 shows, the effect materializes immediately after redistricting and remains relatively stable in the next election. There is no pre-trend in terms of income match.

[Figure 9 about here.]

Similar to experience, results suggest that candidates which score high in the identity match in the weak party enter relatively more frequently as the competition increases. How can we interpret these findings? There are two important differences to the findings about candidates' experience. First, different from experience, shared identity in the weak party does not come with electoral advantages but indeed disadvantages. We should thus interpret shared identity in the weak party as “negative valence”. This would mean, that at least from the perspective of the primary voters in the weak party, “better” candidates enter more frequently as competition decreases. The effect would hence be reversed as compared to candidates with experience. Second, there is no reason to suspect that outside options differ between candidates with high and low values of identity match.<sup>31</sup> This means that for candidates with high values of the identity match, entering is less profitable than for those with low values. However as competition increases their chance of winning office increases. This suggests that even for potential candidates with high values who previously did not choose to enter, it becomes relatively more attractive to enter, thus driving down the the average estimates of shared identity of the candidate pool.<sup>32</sup>

The results for entry of experienced candidates in both parties and candidates with shared identity in the weak party are indeed fully consistent with theoretical considerations regarding the effect of competition on the entry of high valence candidates. Dal Bó and Finan (2018) develop a probabilistic voting model with valence and endogenous candidate entry. They consider two different cost structures. The authors show that when outside options (or costs of running for office) increase sufficiently in valence, candidates in the electorally dominant party shy away from competition. The opposite holds for the smaller party. This conclusion fits well with the results for candidate experience. When costs do not systematically differ between candidates with different valence levels, effects reverse and average valence levels of candidates in the strong party increase with competition while they decrease in the weak party. Again this fits well for the results regarding shared identity, once we think of this trait as a form of “negative valence” in the weak party.

## 4.5 Robustness checks

Despite the evidence that there is no pre-trend for the districts that experience larger changes in *StrongPartyShare*, a concern remains that parties might choose to redistrict in a way so as to influence the primary elections in the opposing party. In order to address

---

<sup>31</sup> While outside options are likely to differ between female and male candidates as well as between candidates with different racial or ethnic background and obviously relate to income, all regressions control for these characteristics, such that the identity match solely measures the shared identity of candidates and their district.

<sup>32</sup> It is worth noting that the results are also consistent with an alternative mechanism based on how “negative” shared identity is in the weak party. If the district becomes more competitive, it is indeed less likely to be a bad trait because there is a larger share of voters who identify with the weak party making it more representative. If this is the case, we would also expect to see more entry of these candidates because of increased chances of winning office.

this issue, in the appendix, I further restrict the sample to only primary elections in which no party solely controls the redistricting procedure. Results are qualitatively the same as in the main analysis for both experience (table A5) and identity match (tables A6, A7 and A8).<sup>33</sup> This result suggests that parties do not use redistricting in order to influence the candidate pool of the opposing party and the identification is valid in this regard.

A second concern relates to the measures for shared identity, which by definition changes with redistricting also for candidates who have previously ran. The concern is that the effect of competition on the entry of candidates with high levels of shared identity is only driven by this with-in candidate variation and might hence not be meaningful in terms of entry of candidates but rather how redistricting affects the identity match of existing candidates. To alleviate this concern, tables A9, A10, A11 and A12 in the appendix present the estimation of eq. (1) when restricting the sample only to first time candidates for experience and identity match respectively. Since in this case every candidate is at a maximum included once, the estimated results cannot be driven by within-candidate variation. As the tables show, all estimates are qualitatively the same.

A final concern relates to the construction of *StrongPartyShare*, which is not based on any gender and race or ethnicity information. In the absence of this information the identity match might proxy for any role these variables play. For this reason, I also predict a second version of *StrongPartyShare* using both gender and race or ethnicity information on the district level and estimate eq. (1) using this version of *StrongPartyShare*. The estimates, presented in tables A13, A14 and A15 in the appendix, are remarkably similar to the one's in the benchmark regressions, thus alleviating this concern.

## 4.6 Competition and elected representatives

Experienced candidates avoid competition in the strong party more than their inexperienced competitors, whereas they enter more frequently than those without experience in the weak party as general elections become more competitive. As opposed to this “better” candidates in terms of identity tend to enter relatively less frequently in the weak party as competition increases. From the perspective of the strong party as well as the weak party this is not necessarily optimal. Indeed parties would benefit from choosing their best candidates for the most competitive districts as they often do in election systems without primary elections (see e.g. de Paola and Scoppa, 2011; Galasso and Nannicini, 2011). From the voter's or policy maker's perspective however, this does not immediately follow. Voters should care about the quality of their elected representatives rather than the candidates. This leads to the question, how does competition affect the experience and identity match of the winner of the general election?

[Table 9 about here.]

---

<sup>33</sup> Unfortunately further restricting the sample to include only non-partisan commission or court mandated redistricting, does not leave enough observations for meaningful results.

Table 9 presents the corresponding regression results. For columns 1 and 2, the outcome variable is the legislative experience of the elected representative. In columns 3 to 8 I use the three different measures for identity match as outcomes. For this analysis I consider only elections in which neither party controlled the redistricting procedure by itself. The results indicate that there is no effect of competition on either the experience of the winner or on the identity match. Intuitively the result can be explained by two counteracting forces. First, as competition decreases, the average experience of the candidates in the weak party decreases, which should lower the expected experience of the elected representative. On the other hand, candidates' experience in the strong party increases. Furthermore, the selected candidate from the strong party becomes more likely to be elected as representative due to the decrease in competition. Jointly the two forces seem to be canceling each other out. A similar logic can explain the null-effect for shared identity in which case the null-effect for the strong party seems to cancel out the weak party's negative effect.

This is overall reassuring, as it suggests that even in non-competitive general elections high valence candidates are selected. This is mainly due a shift of competition to the primary election.<sup>34</sup> Taking a broader view, the results also suggest that redistricting (at least as long as no party solely controls the procedure) does not negatively impact the quality of elected officials.

## 5 Discussion and conclusion

In this paper I analyze how high valence politicians react to changes in competition as compared to low valence ones. In particular, I consider the average experience as well as shared identity of the candidates that stand for the primary elections for the U.S. House of Representatives. I show that experience always benefits candidates in the election, however shared identity is only an advantage in the strong party and a disadvantage in the weak party. A possible explanation lies in the different set of voters in the primary elections. I then use district-level demographic changes due to redistricting and predicted vote shares to measure changes in competition in each district. My difference-in-difference results suggest that experienced candidates in the electorally strong party enter less frequently than inexperienced ones as competition increases. The opposite holds in the weak party. Regarding shared identity, I show that there is no effect of competition on the entry of candidates in the strong party with high values for identity match as compared to those with low values, but candidates with this trait enter more frequently in weak party when general elections become more competitive. The results are consistent with models of candidate entry in which differing outside options drive the entry decision.

Regarding the effects for the eventually elected representatives, I show that competition

---

<sup>34</sup> This insight resembles earlier results from Hirano and Snyder (2014) who show that primary elections in open seat races select high valence candidates. My analysis suggests that their result holds in an even broader context.



does not positively nor negatively affect the quality of the selected politician. A possible reason for this lies in the selection in the primary elections as well as entry responses of the potential candidates. While the general elections become less competitive, competition in the primary elections becomes even more fierce which ultimately cancels out the negative effect from the reduced general election competition. Thus competition moves to a different stage of the contest. This insight expands the literature on political selection in an important dimension – the entry of candidates. Once candidates themselves decide when and where to run, there does not seem to be a positive general election competition effect on political selection.

There are two main shortcomings of the analysis. First, my identification strategy relies on changes due to redistricting that happens only once every 10 years. It is thus relatively restricted in picking up short term variations in candidates' average valence as well as competition changes within a decade. Second, given the context of primary elections to the U.S. House, there is only limited data available on candidates in particular in the first redistricting cycle (1998-2008). Future studies could make use of better data on candidates' background for more recent years as well as exploit more short term variations in competition to better understand what drives the entry patterns and what role parties play in motivating potential candidates to run.

## References

- Abramowitz, A. I., B. Alexander, and M. Gunning (2006). Incumbency, redistricting, and the decline of competition in U.S. House elections. *Journal of Politics* 68(1), 75–88.
- Adida, C. L. (2015). Do African Voters Favor Coethnics? Evidence from a Survey Experiment in Benin. *Journal of Experimental Political Science* 2(1), 1–11.
- Ambekar, A., C. Ward, J. Mohammed, S. Male, and S. Skiena (2009). Name-Ethnicity Classification from Open Sources. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '09, New York, NY, USA, pp. 49–58. Association for Computing Machinery.
- Ansolabehere, S. and J. M. Snyder (2012). The Effects of Redistricting on Incumbents. *Election Law Journal: Rules, Politics, and Policy* 11(4), 490–502.
- Atkinson, M. D., R. D. Enos, and S. J. Hill (2009). Candidate Faces and Election Outcomes: Is the Face–Vote Correlation Caused by Candidate Selection? *Quarterly Journal of Political Science* 4(3), 229–249.
- Baskaran, T. and Z. Hessami (2019). Competitively Elected Women as Policy Makers. CESifo Working Paper No. 8005.
- Ben-Bassat, A. and M. Dahan (2012). Social identity and voting behavior. *Public Choice* 151(1-2), 193–214.
- Berg, H. (2018). Is it Worth it? On the Returns to Holding Political Office. CESifo Working Paper No. 7406.
- Bernecker, A. (2014). Do Politicians Shirk When Reelection is Certain? Evidence from the German Parliament. *European Journal of Political Economy* 36, 55–70.
- Besley, T. (2005). Political Selection. *Journal of Economic Perspectives* 19(3), 43–60.
- Besley, T., J. G. Montalvo, M. Reynal, and M. Reynal-Querol (2011). Do Educated Leader Matter? *Economic Journal* 121(August), F205–F227.
- Besley, T., T. Persson, and D. Sturm (2010). Political Competition, Policy and Growth: Theory and Evidence from the US. *Review of Economic Studies* 77(4), pp. 1329–1352.
- Besley, T. and M. Reynal-Querol (2011). Do Democracies Select More Educated Leaders? *American Political Science Review* 105(3), 552–566.
- Blackwell, C. and P. T. Calcagno (2019). Party-crashers or wallflowers? The lack of strategic voting in experimental primaries. *Applied Economics Letters* 26(19), 1641–1648.

- Brians, C. L. (2005). Women for women? Gender and party bias in voting for female candidates. *American Politics Research* 33(3), 357–375.
- Broockman, D. E. (2013). Black politicians are more intrinsically motivated to advance blacks’ interests: A field experiment manipulating political incentives. *American Journal of Political Science* 57(3), 521–536.
- Butler, D. M. and D. E. Broockman (2011). Do politicians racially discriminate against constituents? A field experiment on state legislators. *American Journal of Political Science* 55(3), 463–477.
- Buttice, M. K. and W. J. Stone (2012). Candidates matter: Policy and quality differences in congressional elections. *Journal of Politics* 74(3), 870–887.
- Carson, J. L., M. H. Crespín, and R. D. Williamson (2014). Reevaluating the Effects of Redistricting on Electoral Competition, 1972–2012. *State Politics and Policy Quarterly* 14(2), 165–177.
- Carson, J. L., E. J. Engstrom, and J. M. Roberts (2006). Redistricting, Candidate Entry, and the Politics of Nineteenth-century U.S. House Elections. *American Journal of Political Science* 50(2), 283–293.
- Carson, J. L. and J. M. Roberts (2005). Strategic Politicians and U.S. House Elections, 1874 – 1914. *Journal of Politics* 67(2), 474–496.
- Cascio, E. U. and N. Shenhav (2020). A Century of the American Woman Voter: Sex Gaps in Political Participation, Preferences, and Partisanship Since Women’s Enfranchisement. NBER Working Paper No. 26709.
- Casey, K. (2015). Crossing party lines: The effects of information on redistributive politics. *American Economic Review* 105(8), 2410–2448.
- Cerina, F. and L. G. Deidda (2017). Rewards from Public Office and the Selection of Politicians by Parties. *European Journal of Political Economy* 47(February 2016), 1–18.
- Chattopadhyay, R. and E. Duflo (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica* 72(5), 1409–1443.
- Cherry, T. L. and S. Kroll (2003). Crashing the party: An experimental investigation of strategic voting in primary elections. *Public Choice* 114(3-4), 387–420.
- Conroy-Krutz, J. (2013). Information and ethnic politics in Africa. *British Journal of Political Science* 43(2), 345–373.
- Cook, R. (2013). *America Votes 30*. Thousand Oaks, CA: SAGE Publications, Inc.

- Cook, R. (2015). *America Votes 31: 2013-2014, Election Returns by State*. Thousand Oaks, CA: CQ Press.
- Curry, J. M. and M. R. Haydon (2018). Lawmaker Age, Issue Salience, and Senior Representation in Congress. *American Politics Research* 46(4), 567–595.
- Dal Bo, E., P. Dal Bo, and J. Snyder (2009). Political Dynasties. *Review of Economic Studies* 76(1), 115–142.
- Dal Bó, E. and F. Finan (2018). Progress and Perspectives in the Study of Political Selection. *Annual Review of Economics* 10(1), 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(4), 1877–1914.
- de Paola, M. and V. Scoppa (2011). Political Competition and Politician Quality: Evidence from Italian Municipalities. *Public Choice* 148(3-4), 547–559.
- Eggers, A. C. and J. Hainmueller (2009). MPs for Sale? Returns to Office in Postwar British Politics. *American Political Science Review* 103(04), 513–533.
- Eggers, A. C. and M. Klačnja (2019). Wealth, Voting, and Fundraising in the U.S. Congress. Working Paper.
- Eifert, B., E. Miguel, and D. N. Posner (2010). Political Competition and Ethnic Identification. *American Journal of Political Science* 54(2), 494–510.
- Elliott, M. N., P. A. Morrison, A. Fremont, D. F. McCaffrey, P. Pantoja, and N. Lurie (2009). Using the Census Bureau’s surname list to improve estimates of race/ ethnicity and associated disparities. *Health Services and Outcomes Research Methodology* 9(2), 69–83.
- Ferraz, C. and F. Finan (2008). Exposing Corrupt Politicians: The Effects of Brazil’s Publicly Released Audits on Electoral Outcomes. *Quarterly Journal of Economics* 123(2), 703–745.
- Ferreira, F. and J. Gyourko (2014). Does gender matter for political leadership? The case of U.S. mayors. *Journal of Public Economics* 112, 24–39.
- Friedman, J. N. and R. T. Holden (2008). Optimal Gerrymandering: Sometimes Pack, but Never Crack. *American Economic Review* 98(1), 113–144.
- Gagliarducci, S. and T. Nannicini (2013). Do Better Paid Politicians Perform Better? Disentangling Incentives From Selection. *Journal of the European Economic Association* 11(2), 369–398.

- Galasso, V. and T. Nannicini (2011). Competing on Good Politicians. *American Political Science Review* 105(1), 79–99.
- Gavoille, N. and M. Verschele (2017). Electoral Competition and Political Selection: An Analysis of the Activity of French Deputies, 1958–2012. *European Economic Review* 92(September 2015), 180–195.
- Hedlund, R. D. (1977). Cross-Over Voting in a 1976 Open Presidential Primary. *Public Opinion Quarterly* 41(4), 498.
- Hirano, S. and J. M. Snyder (2014). Primary Elections and the Quality of Elected Officials. *Quarterly Journal of Political Science* 9(4), 473–500.
- Houle, C., P. D. Kenny, and C. Park (2019). The structure of ethnic inequality and ethnic voting. *Journal of Politics* 81(1), 187–200.
- Iversen, T. and F. Rosenbluth (2008). Work and Power: The Connection Between Female Labor Force Participation and Female Political Representation. *Annual Review of Political Science* 11(1), 479–495.
- Jacobson, G. C. (1989). Strategic Politicians and the Dynamics of U.S. House Elections, 1946–86. *American Political Science Review* 83(3), 773–793.
- Jacobson, G. C. and S. Kernell (1981). *Strategy and Choice in Congressional Elections*. Yale: Yale University Press.
- Jones, D. B. and R. Walsh (2018). How do Voters Matter? Evidence from US Congressional Redistricting. *Journal of Public Economics* 158, 25–47.
- Kandt, J. and P. A. Longley (2018). Ethnicity estimation using family naming practices. *PLoS ONE* 13(8), 1–24.
- Kotakorpi, K. and P. Poutvaara (2011). Pay for Politicians and Candidate Selection: An Empirical Analysis. *Journal of Public Economics* 95(7-8), 877–885.
- Kotakorpi, K., P. Poutvaara, and M. Tervio (2017). Returns to Office in National and Local Politics: A Bootstrap Method and Evidence from Finland. *Journal of Law, Economics, and Organization* 33(3), 413–442.
- Leip, D. (2013). Dave Leip’s Atlas of U.S. Presidential Elections.
- Levitt, J. (2018). All About Redistricting.
- Lewis, J., B. DeVine, L. Pitcher, and K. Martis (2013). Digital Boundary Definitions of United States Congressional Districts, 1789–2012.

- Lindgren, E. and P. Southwell (2013). The Effect of Redistricting Commissions on Electoral Competitiveness in U.S. House Elections, 2002-2010. *Journal of Politics and Law* 6(2), 13–18.
- Lowande, K., M. Ritchie, and E. Lauterbach (2019). Descriptive and Substantive Representation in Congress: Evidence from 80,000 Congressional Inquiries. *American Journal of Political Science* 63(3), 644–659.
- Mason, S., J. Schroeder, D. Van Riper, and S. Ruggles (2019). IPUMS National Historical Geographic Information System: Version 14 [Database].
- Mateos, P. (2007). Classification Methods and their Potential in Population Studies. *Population, Space and Place* 13(May), 243–263.
- McCarty, N., K. T. Poole, and H. Rosenthal (2009). Does Gerrymandering Cause Polarization? *American Journal of Political Science* 53(3), 666–680.
- Mcdermott, M. L. (1998). Race and Gender Cues in Low-Information Elections. *Political Research Quarterly* 51(4), 895–918.
- Michelson, M. R. (2005). Does ethnicity trump party? Competing vote cues and latino voting behavior. *Journal of Political Marketing* 4(4), 1–25.
- Minta, M. D. (2009). Legislative oversight and the substantive representation of black and Latino interests in congress. *Legislative Studies Quarterly* 34(2), 193–218.
- MIT Election Data and Science Lab (2017). U.S. House 1976–2016.
- Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122(3), 1067–1101.
- Pande, R. (2003). Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India. *American Economic Review* 93(4), 1132–1151.
- Pettigrew, S., K. Owen, and E. Wanless (2016). U.S. House Primary Election Results (1956-2010).
- Querubin, P. and J. Snyder (2009). The Returns to U . S . Congressional Seats in the Mid-19th Century. In E. Aragonés, C. Bevia, H. Llavador, and N. Schofiel (Eds.), *The Political Economy of Democracy*, pp. 225–240. Madrid: Fund. BBKA: Fund. BBKA.
- Sanbonmatsu, K. (2002). Gender Stereotypes and Vote Choice. *American Journal of Political Science* 46(1), 20–34.

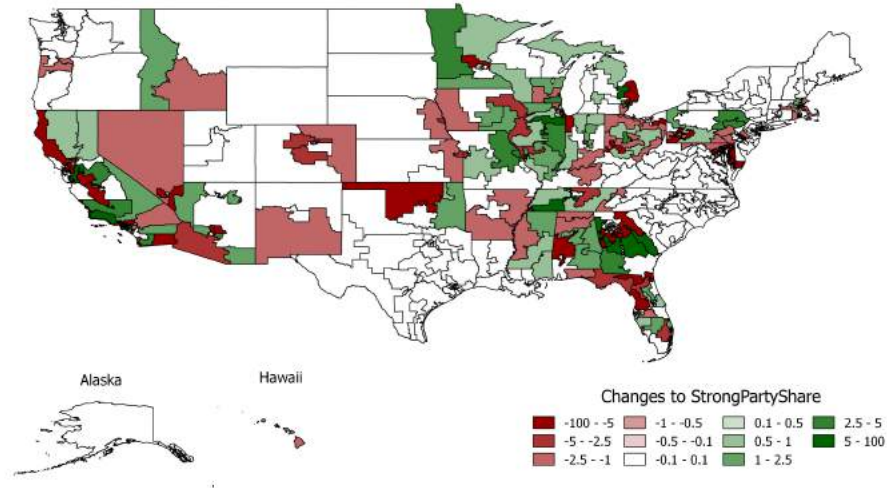
- Smith, E. R. and R. L. Fox (2001). The electoral fortunes of women candidates for congress. *Political Research Quarterly* 54(1), 205–221.
- Stigler, G. J. (1972). Economic Competition and Political Competition. *Public Choice* 13, 91–106.
- Thompson, D. M., J. J. Feigenbaum, A. B. Hall, and J. Yoder (2019). Who becomes a Member of Congress? Evidence from de-anonymized census data. NBER Working Paper No. 26156.
- Washington, E. L. (2008). Female socialization: How daughters affect their legislator fathers' voting on women's issues. *American Economic Review* 98(1), 311–332.
- Wittman, D. A. (1989). Why Democracies Produce Efficient Results. *Journal of Political Economy* 97(6), 1395–1424.
- Wolfinger, R. E. (1965). The Development and Persistence of Ethnic Voting. *American Political Science Review* 59(4), 896–908.

## List of Figures

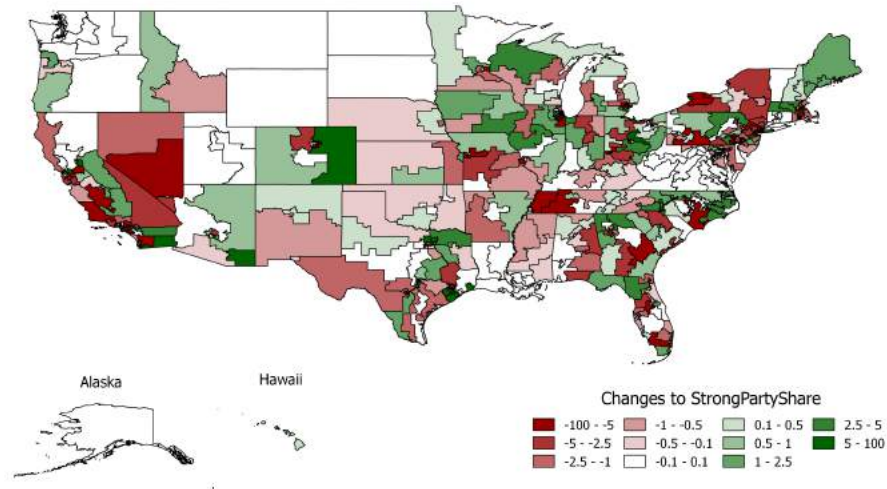
1	Change in <i>StrongPartyShare</i> due to redistricting . . . . .	32
2	Number of bills (co-)sponsored by Representative . . . . .	33
3	Share of experienced candidates by party . . . . .	34
4	Party unity support in U.S. House of Representatives . . . . .	35
5	Pre-redistricting <i>StrongPartyShare</i> and change in <i>StrongPartyShare</i> . . . . .	36
6	Experienced candidates and change in <i>StrongPartyShare</i> : Common trend . . . . .	37
7	Candidate ethnicity match and change in <i>StrongPartyShare</i> : Common trend . . . . .	38
8	Candidate gender match and change in <i>StrongPartyShare</i> : Common trend . . . . .	39
9	Candidate income match and change in <i>StrongPartyShare</i> : Common trend . . . . .	40
A1	Vote participation in U.S. House of Representatives . . . . .	51
A2	Distribution ethnicity match . . . . .	52
A3	Distribution gender match . . . . .	53
A4	Distribution income match . . . . .	54
A5	Racial or ethnic representativeness of candidate pool . . . . .	55
A6	Gender representativeness of candidate pool . . . . .	56
A7	Economic representativeness of candidate pool . . . . .	57
A8	Seat distribution in state house and change in <i>StrongPartyShare</i> . . . . .	58
A9	Seat distribution in state house and change in <i>StrongPartyShare</i> . . . . .	59
A10	Vote share of candidate in strong party and change in <i>StrongPartyShare</i> : Common trend . . . . .	60



Figure 1: Change in *StrongPartyShare* due to redistricting  
 Distribution of changes to *StrongPartyShare* for redistricting cycle 2000

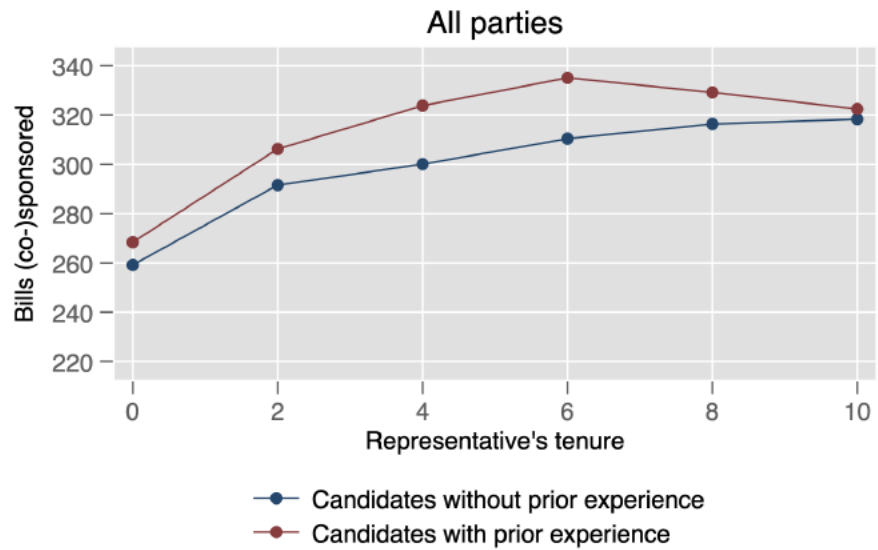


Distribution of changes to *StrongPartyShare* for redistricting cycle 2010



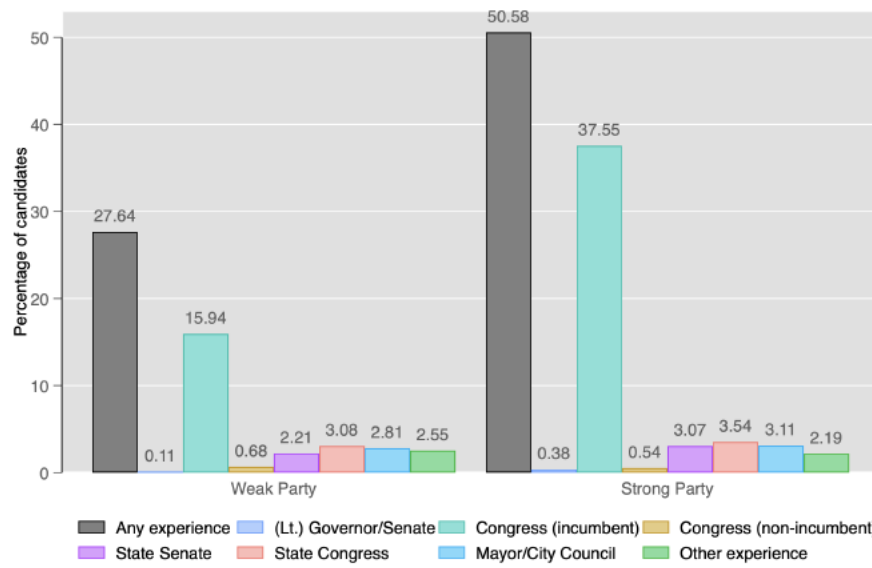
*Notes:* The top part of the figure shows the change in *StrongPartyShare* due to redistricting in 2000 across the entire United States using the Congressional map for the 108th Congress (post redistricting). The bottom panel shows the change in *StrongPartyShare* due to redistricting in 2010 using the Congressional map for the 113th Congress (post redistricting).

Figure 2: Number of bills (co-)sponsored by Representative



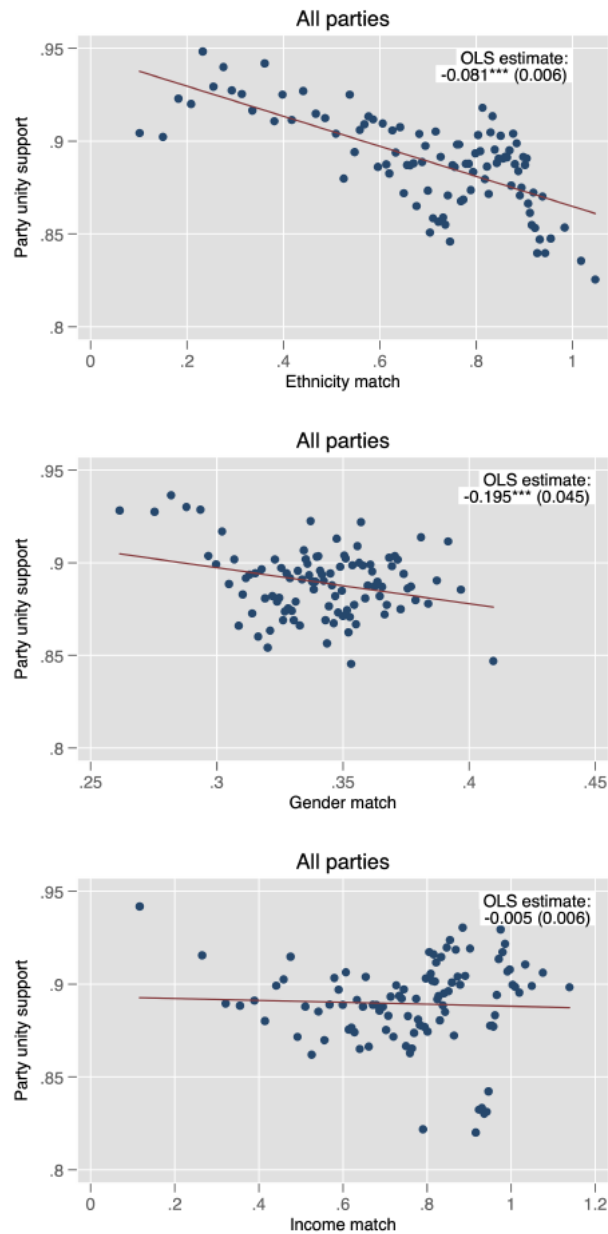
*Notes:* The figure shows the number of bills sponsored or cosponsored between 1973 and 2017 for all elected representatives in the U.S. House of Representatives that have been elected after 1973 by previous legislative experience and tenure. The data on number of bills sponsored and co-sponsored stems from the data of the Library of Congress. Data on previous legislative experience stems from CQ Press's congressional member profiles.

Figure 3: Share of experienced candidates by party



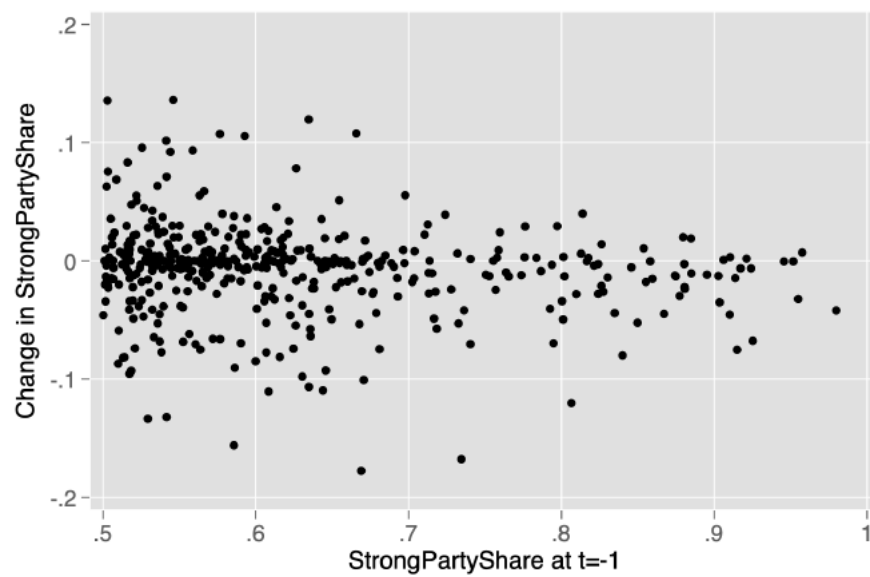
*Notes:* The figure shows the share of candidates in the weak and in the strong party's primary elections in the estimation sample. It also shows the type of previous legislative experience when this information is available.

Figure 4: Party unity support in U.S. House of Representatives



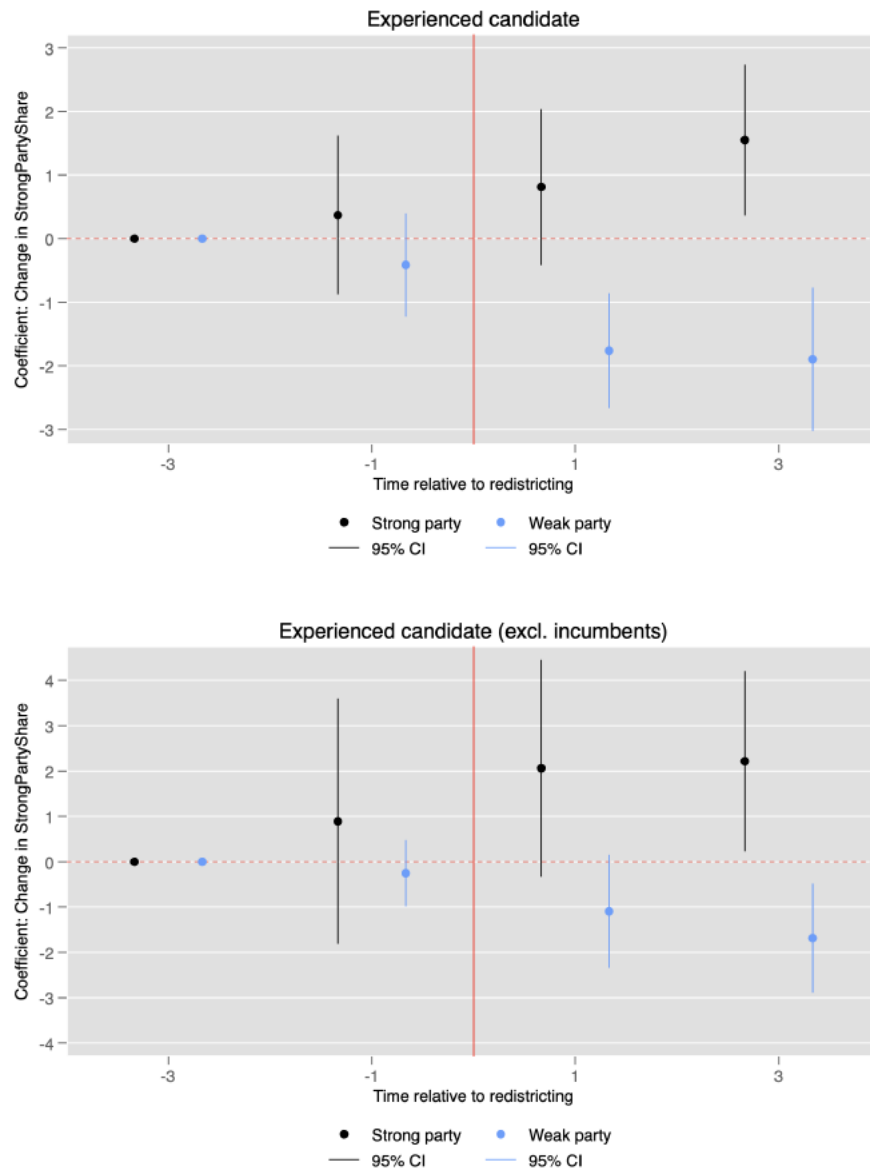
*Notes:* The figure shows the binned share of roll call votes in which a representative voted with the majority of his party opposing the other party for all roll call votes between 1988 and 2015 after controlling for gender, race/ethnicity and income. The x-axis represents a candidate's identity match in the indicated dimension at the time of the election. The top graph corresponds to the ethnicity match – the share of constituents with the same race or ethnicity as the candidate. The middle one to the gender match – the share of labor force participants with the same gender as the candidate. Lastly the bottom graph shows how roll call voting relates to the income match of the candidate - a measure that decreases in the share of voters that lie between the candidate and the median voter on the income distribution. The data on party unity support and representative characteristics stems from CQ Press's congressional member profiles. Information on candidate income is proxied by ZIP code information from FEC data and ZIP code level household income data. I use census data on district and ZIP code characteristics from the IPUMS database (Mason et al., 2019).

Figure 5: Pre-redistricting *StrongPartyShare* and change in *StrongPartyShare*



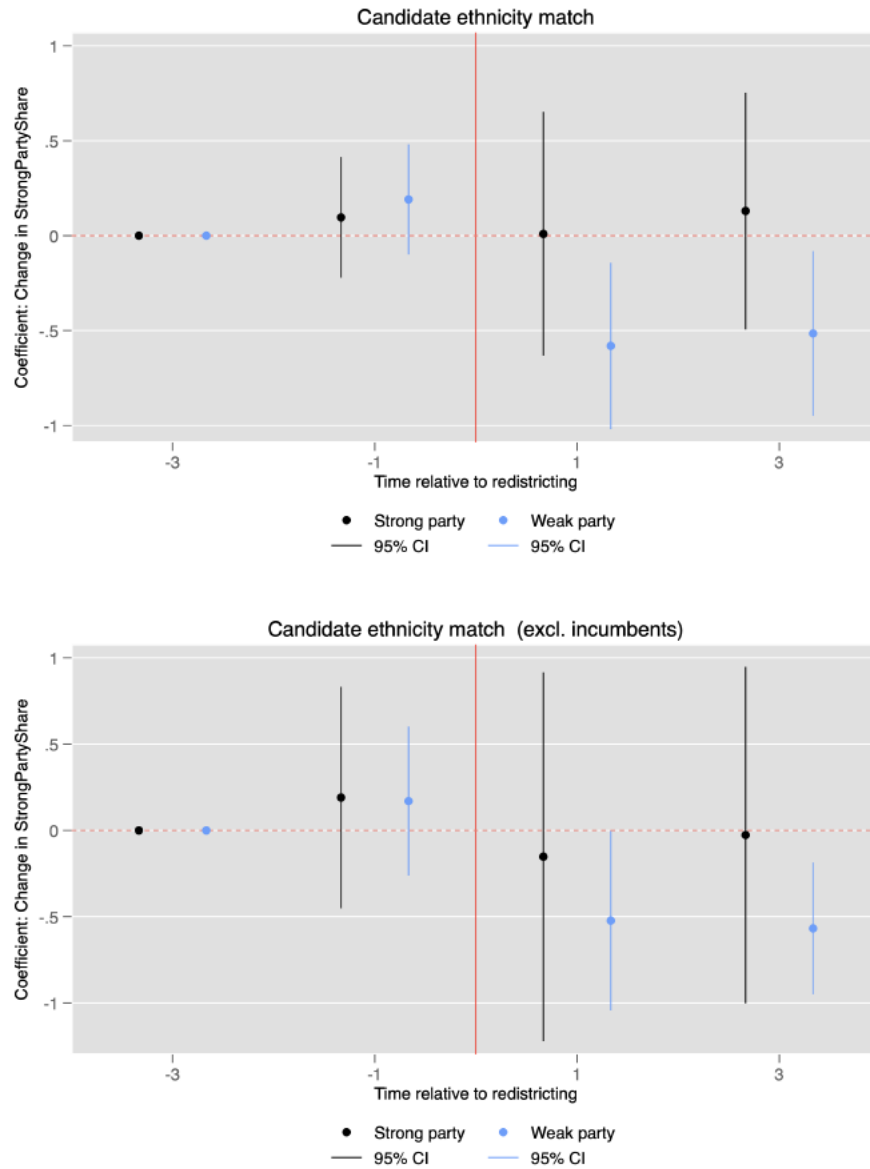
*Notes:* The figure shows the change in *StrongPartyShare* due to redistricting and its relation to the *StrongPartyShare* immediately before redistricting in the same district. The sample of districts is restricted to only include cases where the strong party does not have sole control over the redistricting procedure.

Figure 6: Experienced candidates and change in *StrongPartyShare*: Common trend



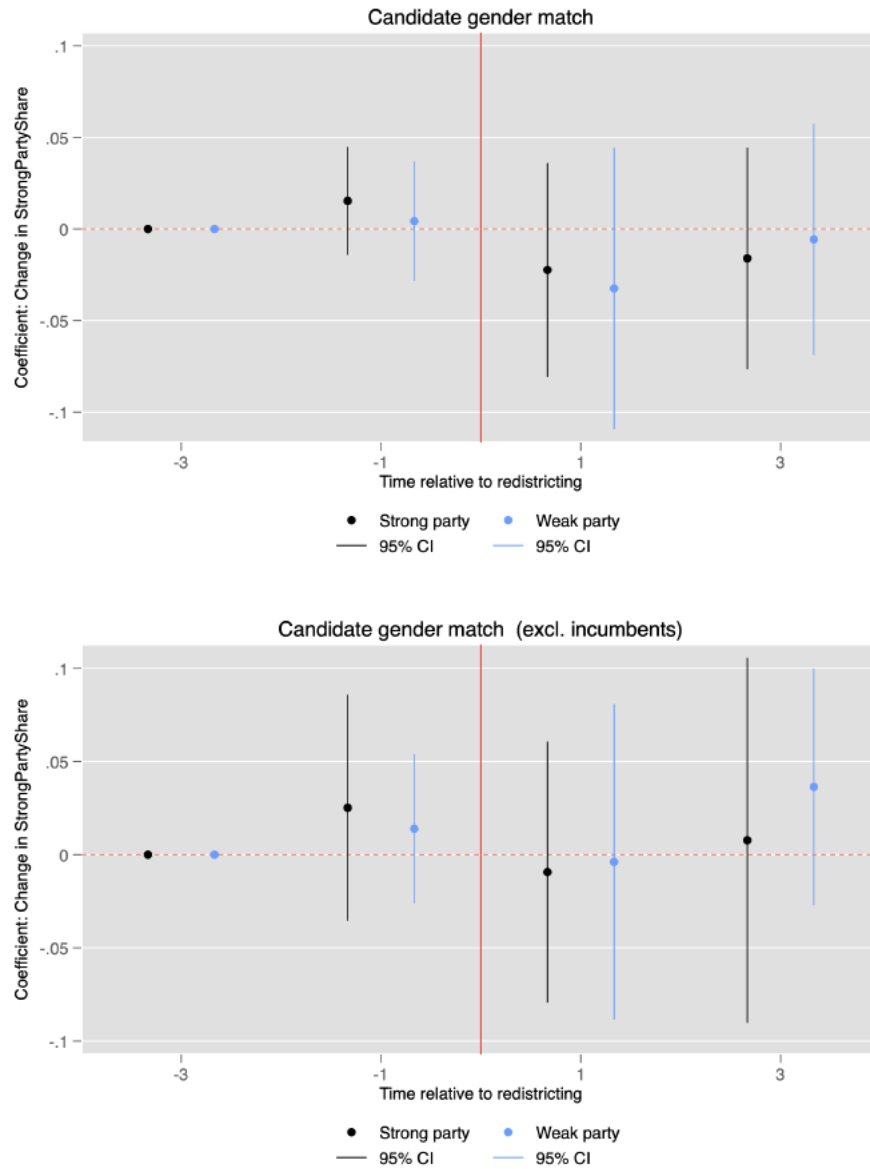
*Notes:* The figure plots coefficients and 95% confidence intervals for regressions of the change in *StrongPartyShare* interacted with time relative to redistricting on the candidate experience. To top panel includes all candidate, while incumbents are excluded in the bottom panel. Estimates in black refer to estimates for the strong party. Estimates for the weak party are in blue. The vertical red line indicates the timing of redistricting. The sample only includes primary elections of the parties that did not solely control the redistricting procedure in the respective state and redistricting cycle.

Figure 7: Candidate ethnicity match and change in *StrongPartyShare*: Common trend



Notes: The figure plots coefficients and 95% confidence intervals for regressions of the change in *StrongPartyShare* interacted with time relative to redistricting on the ethnicity match measure. The top panel includes all candidate, while incumbents are excluded in the bottom panel. Estimates in black refer to estimates for the strong party. Estimates for the weak party are in blue. The vertical red line indicates the timing of redistricting. The sample only includes primary elections of the parties that did not solely control the redistricting procedure in the respective state and redistricting cycle.

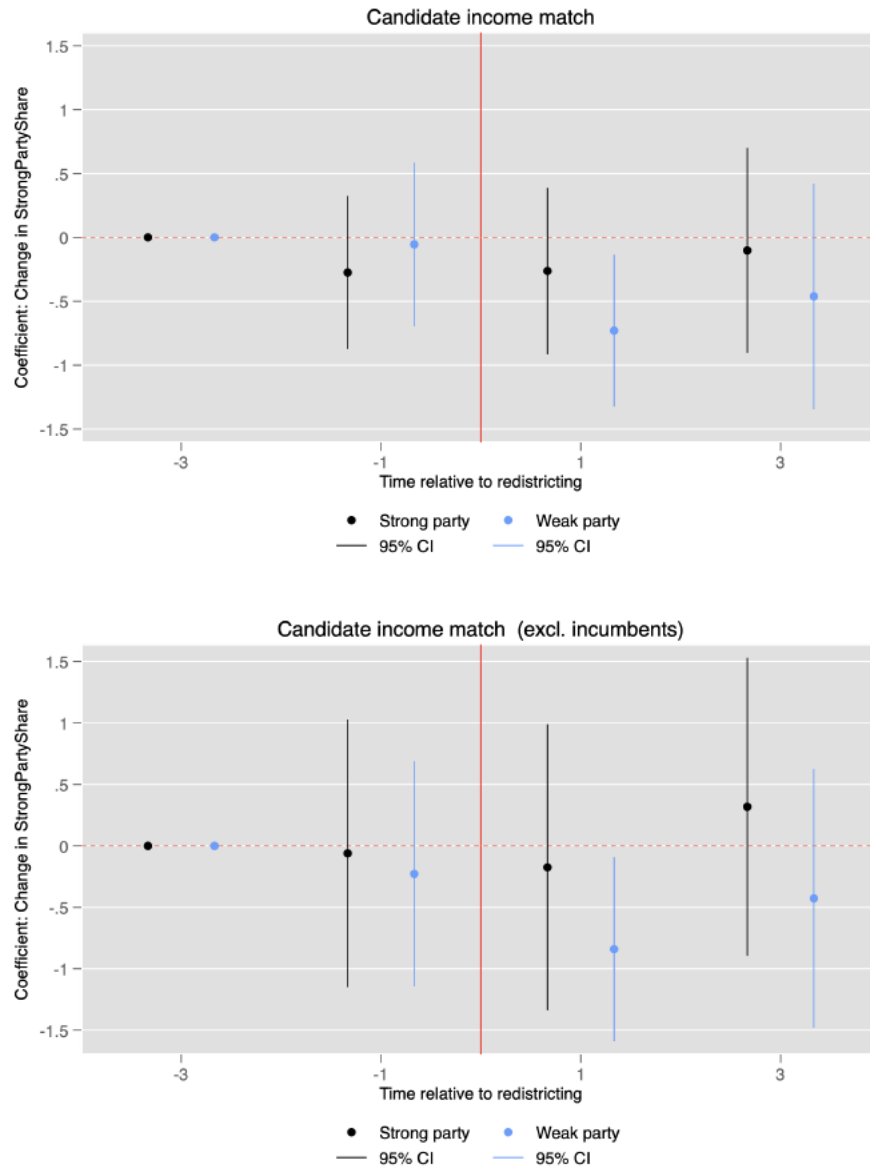
Figure 8: Candidate gender match and change in *StrongPartyShare*: Common trend



*Notes:* The figure plots coefficients and 95% confidence intervals for regressions of the change in *StrongPartyShare* interacted with time relative to redistricting on the candidate gender match measure. To top panel includes all candidate, while incumbents are excluded in the bottom panel. Estimates in black refer to estimates for the strong party. Estimates for the weak party are in blue. The vertical red line indicates the timing of redistricting. The sample only includes primary elections of the parties that did not solely control the redistricting procedure in the respective state and redistricting cycle.



Figure 9: Candidate income match and change in *StrongPartyShare*: Common trend



*Notes:* The figure plots coefficients and 95% confidence intervals for regressions of the change in *StrongPartyShare* interacted with time relative to redistricting on the candidate income match measure. To top panel includes all candidate, while incumbents are excluded in the bottom panel. Estimates in black refer to estimates for the strong party. Estimates for the weak party are in blue. The vertical red line indicates the timing of redistricting. The sample only includes primary elections of the parties that did not solely control the redistricting procedure in the respective state and redistricting cycle.

## List of Tables

1	Summary of covariates . . . . .	42
2	<i>StrongPartyShare</i> and vote share . . . . .	43
3	Experience and identity match: Results in the primary elections . . . . .	44
4	Experience and identity match: Results in the general election . . . . .	45
5	Decreased competition and experience . . . . .	46
6	Decreased competition and ethnicity match . . . . .	47
7	Decreased competition and gender match . . . . .	48
8	Decreased competition and income match . . . . .	49
9	Valence of winner of general election . . . . .	50
A1	Estimates for fractional logit on county level voter registration . . . . .	61
A2	Experience and identity match: Results in contested primary elections . . . . .	62
A3	Experience and identity match: Results in contested general elections . . . . .	63
A4	Entry of experienced candidates and age . . . . .	64
A5	Decreased competition and experience: No partisan redistricting . . . . .	65
A6	Decreased competition and ethnicity match: No partisan redistricting . . . . .	66
A7	Decreased competition and gender match: No partisan redistricting . . . . .	67
A8	Decreased competition and income match: No partisan redistricting . . . . .	68
A9	Decreased competition and experience: First time candidates . . . . .	69
A10	Decreased competition and ethnicity match: First time candidates . . . . .	70
A11	Decreased competition and gender match: First time candidates . . . . .	71
A12	Decreased competition and income match: First time candidates . . . . .	72
A13	Decreased competition and ethnicity match ( <i>StrongPartyShare</i> includes information on gender and race or ethnicity) . . . . .	73
A14	Decreased competition and gender match ( <i>StrongPartyShare</i> includes information on gender and race or ethnicity) . . . . .	74
A15	Decreased competition and income match ( <i>StrongPartyShare</i> includes information on gender and race or ethnicity) . . . . .	75
A16	Decreased competition and experience: No open seats in cycle . . . . .	76
A17	Decreased competition and ethnicity match: No open seats in cycle . . . . .	77
A18	Decreased competition and gender match: No open seats in cycle . . . . .	78
A19	Decreased competition and income match: No open seats in cycle . . . . .	79
A20	Decreased competition and experience: Heterogeneity in pre-redistricting <i>StrongPartyShare</i> . . . . .	80
A21	Decreased competition and ethnicity match: Heterogeneity in pre-redistricting <i>StrongPartyShare</i> . . . . .	81
A22	Decreased competition and gender match: Heterogeneity in pre-redistricting <i>StrongPartyShare</i> . . . . .	82
A23	Decreased competition and income match: Heterogeneity in pre-redistricting <i>StrongPartyShare</i> . . . . .	83

Table 1: Summary of covariates

	(1)	(2)	(3)	(4)
	Mean	SD	$\beta_{SPS}$	$SE_{SPS}$
<u>Candidates in strong party</u>				
Female	0.18	0.38	0.42	0.55
Minority	0.08	0.27	-0.32	0.30
Age	54.3	11.1	1.9	16.3
ZIP avg. income	27769	16777	-22938	12966
Candidates			2802	
<u>Candidates in weak party</u>				
Female	0.15	0.36	0.04	0.34
Minority	0.07	0.25	0.09	0.27
Age	52.2	10.9	-2.3	15.6
ZIP avg. income	28305	13813	-1112	22623
Candidates			2651	
<u>Elections</u>				
Lagged elections won by strong party	0.70	0.46	-0.12	0.40
Seats held by strong party	0.69	0.46	-0.02	0.71
Open seats	0.09	0.29	0.81	0.52
Races with two incumbents	0.01	0.09	0.03	0.19
Elections			1470	

*Notes:* The table presents summary values for the baseline covariates as well as estimates for regressions of *StrongPartyShare* on the covariate. The specification is as in eq. (1) with the covariate as outcome variable. \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors are clustered at the state-cycle level.

Table 2: *StrongPartyShare* and vote share

<b><i>Vote share of strong party</i></b>		
Mean	0.575	0.575
<i>StrongPartyShare</i>	0.616*** (0.121)	0.601*** (0.085)
Last election result	✓	✓
Year FE's	✓	✓
District-cycle FE's	✓	✓
Additional controls		✓
N	921	921
R-squared	0.865	0.872

*Notes:* \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents is included. The sample only includes elections in states and redistricting cycles in which none of the parties solely controlled the redistricting procedure. Singleton groups are excluded from the estimation.

Table 3: Experience and identity match: Results in the primary elections

	Strong party		Weak party	
<i>Candidate vote share</i>				
Mean	0.571	0.571	0.576	0.576
Experience	0.115*** (0.012)	0.111*** (0.013)	0.108*** (0.011)	0.107*** (0.012)
Ethnicity match	0.024** (0.011)	0.049*** (0.012)	-0.019** (0.008)	-0.027** (0.012)
Gender match	0.223* (0.124)	0.309** (0.149)	-0.060 (0.078)	-0.055 (0.096)
Income match	-0.001 (0.015)	-0.001 (0.016)	-0.026* (0.015)	-0.027* (0.016)
Gender, ethnicity, income	✓	✓	✓	✓
Competition controls	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
State FE's		✓		✓
N	2265	2264	2381	2381
R-squared	0.875	0.878	0.859	0.860
<i>Primary winner</i>				
Mean	0.564	0.564	0.575	0.575
Experience	0.187*** (0.027)	0.181*** (0.028)	0.218*** (0.026)	0.218*** (0.027)
Ethnicity match	0.040* (0.020)	0.068*** (0.022)	-0.028 (0.019)	-0.029 (0.027)
Gender match	0.432* (0.247)	0.585** (0.288)	-0.090 (0.183)	-0.028 (0.237)
Income match	0.005 (0.035)	0.006 (0.036)	-0.015 (0.040)	-0.017 (0.042)
Gender, ethnicity, income	✓	✓	✓	✓
Competition controls	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
State FE's		✓		✓
N	2265	2264	2381	2381
R-squared	0.639	0.642	0.520	0.522

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Competition controls are dummy variables for every observed combination of experienced and inexperienced candidates in the primary election. Additional control variables include indicator variables for incumbents and five dummy variables for primary type. Singleton groups are excluded from the estimation.

Table 4: Experience and identity match: Results in the general election

	All parties	
<i>Candidate vote share</i>		
Mean	0.492	0.492
Experience	0.078*** (0.009)	0.078*** (0.009)
Ethnicity match	0.027* (0.016)	0.031** (0.014)
Gender match	-0.028 (0.145)	-0.136 (0.088)
Income match	-0.021** (0.009)	-0.013 (0.009)
Gender, ethnicity, income	✓	✓
Competition controls	✓	✓
Additional controls	✓	✓
Year FE's	✓	✓
State FE's		✓
N	1515	1515
R-squared	0.787	0.806
<i>Election winner</i>		
Mean	0.502	0.502
Experience	0.203*** (0.024)	0.199*** (0.024)
Ethnicity match	0.016 (0.016)	0.040* (0.020)
Gender match	-0.099 (0.155)	0.086 (0.215)
Income match	-0.053 (0.037)	-0.050 (0.041)
Gender, ethnicity, income	✓	✓
Competition controls	✓	✓
Additional controls	✓	✓
Year FE's	✓	✓
State FE's		✓
N	1519	1519
R-squared	0.753	0.756

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Competition controls are dummy variables for experienced and inexperienced competitors in the general election. Additional control variables include the predicted vote share of the candidate's party and indicator variables for incumbents. The sample only includes elections in states and redistricting cycles in which none of the parties solely controlled the redistricting procedure. Singleton groups are excluded from the estimation.

Table 5: Decreased competition and experience

	<u>All parties</u>		<u>Strong party</u>		<u>Weak party</u>	
<b><i>Experienced candidate</i></b>						
Mean	0.414	0.414	0.512	0.512	0.310	0.310
<i>StrongPartyShare</i>	-0.536*	-0.591**	0.975**	1.003**	-1.715***	-1.617***
	(0.320)	(0.287)	(0.471)	(0.430)	(0.372)	(0.344)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	5453	5453	2802	2802	2651	2651
R-squared	0.438	0.443	0.393	0.410	0.448	0.456
<b><i>Experienced candidate (excluding incumbents)</i></b>						
Mean	0.210	0.210	0.248	0.248	0.181	0.181
<i>StrongPartyShare</i>	-0.589	-0.461	1.925**	1.715**	-1.418***	-1.212**
	(0.387)	(0.443)	(0.879)	(0.805)	(0.375)	(0.479)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3986	3986	1765	1765	2221	2221
R-squared	0.290	0.312	0.297	0.322	0.284	0.305

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table 6: Decreased competition and ethnicity match

	All parties		Strong party		Weak party	
<b><i>Candidate ethnicity match</i></b>						
Mean	0.668	0.668	0.666	0.666	0.670	0.670
<i>StrongPartyShare</i>	-0.323** (0.151)	-0.329** (0.148)	0.045 (0.301)	0.023 (0.300)	-0.644*** (0.184)	-0.655*** (0.175)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	4598	4598	2245	2245	2353	2353
R-squared	0.865	0.865	0.864	0.865	0.869	0.870
<b><i>Candidate ethnicity match (excluding incumbents)</i></b>						
Mean	0.666	0.666	0.683	0.683	0.653	0.653
<i>StrongPartyShare</i>	-0.446** (0.178)	-0.448*** (0.167)	-0.018 (0.502)	-0.179 (0.452)	-0.629*** (0.194)	-0.635*** (0.180)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3471	3471	1444	1444	2027	2027
R-squared	0.858	0.859	0.849	0.851	0.869	0.870

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.



Table 7: Decreased competition and gender match

	All parties		Strong party		Weak party	
<b><i>Candidate gender match</i></b>						
Mean	0.340	0.340	0.339	0.339	0.342	0.342
<i>StrongPartyShare</i>	-0.023	-0.023	-0.026	-0.026	-0.022	-0.024
	(0.025)	(0.025)	(0.029)	(0.029)	(0.038)	(0.038)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	4598	4598	2245	2245	2353	2353
R-squared	0.940	0.941	0.936	0.936	0.947	0.947
<b><i>Candidate gender match (excluding incumbents)</i></b>						
Mean	0.341	0.341	0.340	0.340	0.342	0.342
<i>StrongPartyShare</i>	0.004	0.005	-0.015	-0.013	0.010	0.005
	(0.034)	(0.033)	(0.040)	(0.036)	(0.042)	(0.041)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3471	3471	1444	1444	2027	2027
R-squared	0.942	0.942	0.943	0.944	0.943	0.943

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table 8: Decreased competition and income match

	All parties		Strong party		Weak party	
<b><i>Candidate income match</i></b>						
Mean	0.768	0.768	0.766	0.766	0.769	0.769
<i>StrongPartyShare</i>	-0.323	-0.323	-0.069	-0.059	-0.538**	-0.592**
	(0.202)	(0.205)	(0.359)	(0.360)	(0.246)	(0.254)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	4598	4598	2245	2245	2353	2353
R-squared	0.412	0.414	0.433	0.437	0.396	0.398
<b><i>Candidate income match (excluding incumbents)</i></b>						
Mean	0.766	0.766	0.766	0.766	0.765	0.765
<i>StrongPartyShare</i>	-0.322	-0.314	0.046	0.067	-0.488*	-0.542**
	(0.242)	(0.240)	(0.590)	(0.561)	(0.260)	(0.257)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3471	3471	1444	1444	2027	2027
R-squared	0.379	0.381	0.392	0.398	0.373	0.376

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

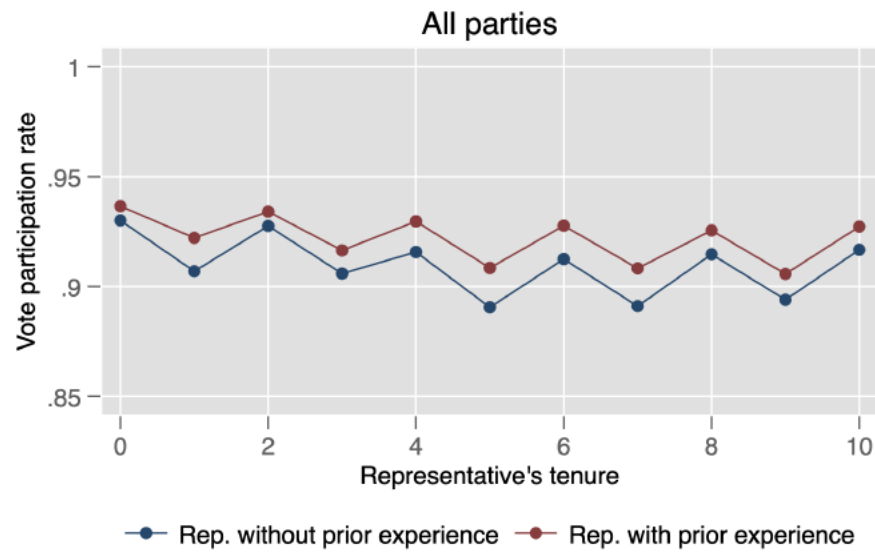
Table 9: Valence of winner of general election

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Experienced candidate		Ethnicity match		Gender match		Income match	
Mean	0.960	0.960	0.698	0.698	0.342	0.342	0.767	0.767
<i>StrongPartyShare</i>	0.035	0.147	-0.287	-0.306	-0.034	-0.036	0.225	0.149
	(0.506)	(0.524)	(0.251)	(0.255)	(0.043)	(0.044)	(0.607)	(0.628)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls		✓		✓		✓		✓
N	944	944	742	742	742	742	742	742
R-squared	0.342	0.385	0.978	0.978	0.977	0.978	0.752	0.759

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. The sample only includes elections in states and redistricting cycles in which none of the parties solely controlled the redistricting procedure. Singleton groups are excluded from the estimation.

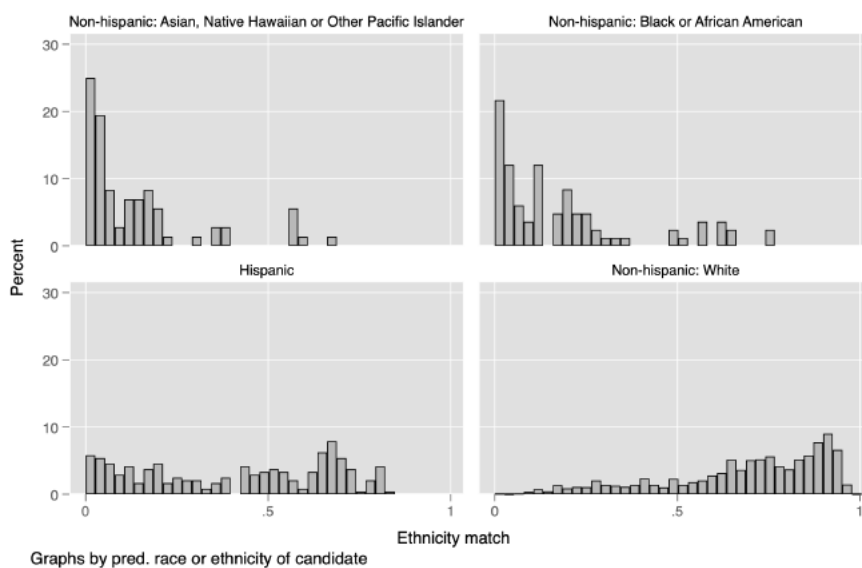
## A Appendix

Figure A1: Vote participation in U.S. House of Representatives



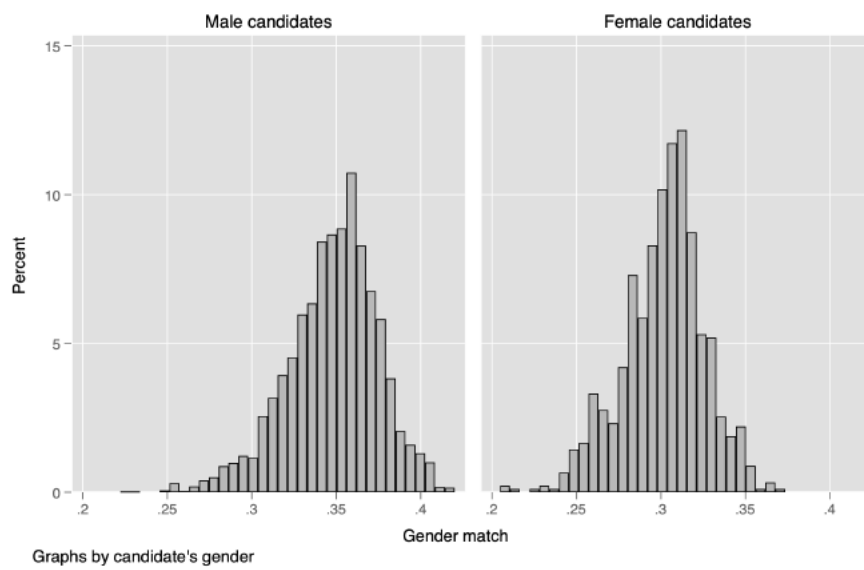
*Notes:* The figure shows vote participation rates in roll call votes between 1950 and 2017 for all elected representatives in the U.S. House of Representatives that have been elected after 1950 by previous legislative experience and tenure. The data on vote participation rates and previous legislative experience stems from CQ Press's congressional member profiles.

Figure A2: Distribution ethnicity match



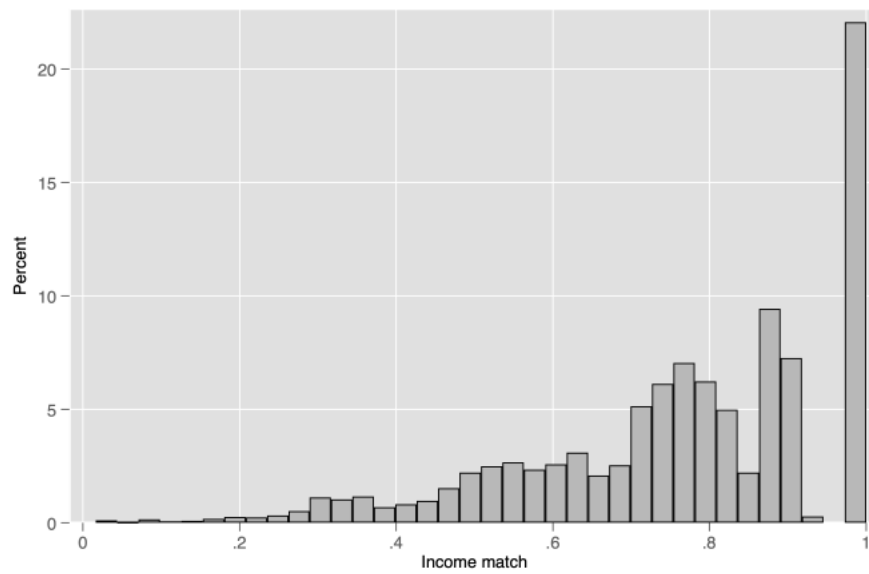
*Notes:* The figure shows the distribution of the measure “ethnicity match” by candidates race or ethnicity. The measure is defined as the share of constituents with the same race or ethnicity and in the same district as the candidate.

Figure A3: Distribution gender match



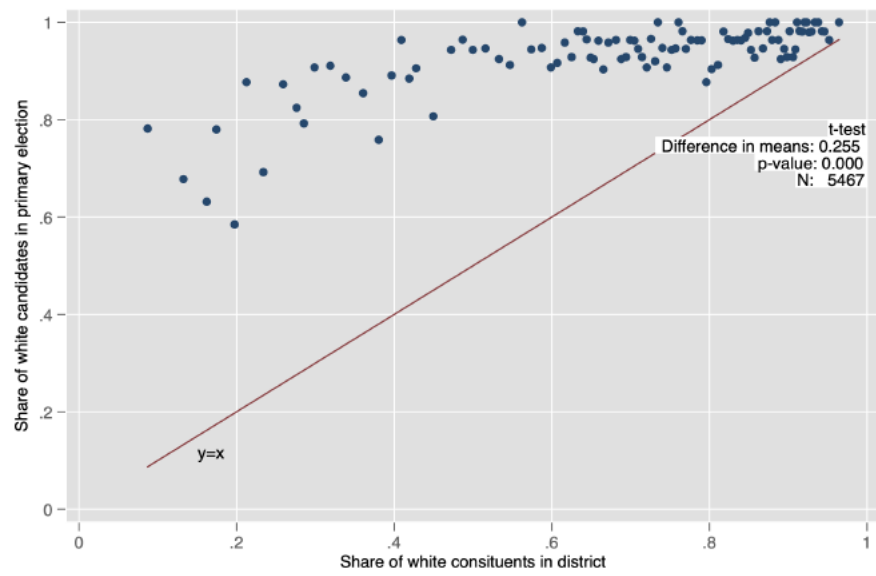
*Notes:* The figure shows the distribution of the measure “gender match” by candidates gender. The measure is defined as the share of constituents in the labor force with the same gender and in the same district as the candidate.

Figure A4: Distribution income match



*Notes:* The figure shows the distribution of the measure “income match”. The measure is defined as 1 minus twice the share of constituents between the proxied candidate’s household income and the income of the median household in the same district as the candidate.

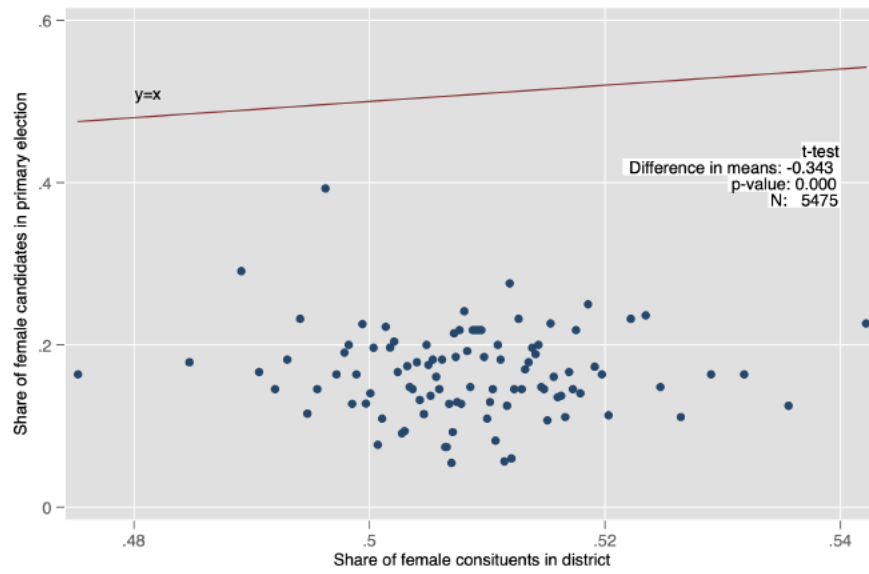
Figure A5: Racial or ethnic representativeness of candidate pool



*Notes:* The figure shows binned averages of the share of white candidates in primary elections against the share of white constituents in the district. The red line indicates equal shares.

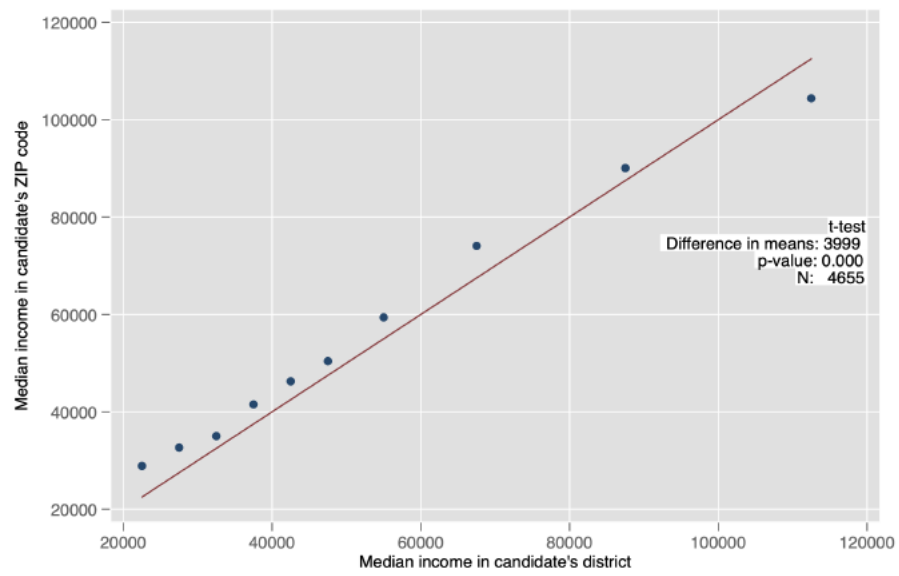


Figure A6: Gender representativeness of candidate pool



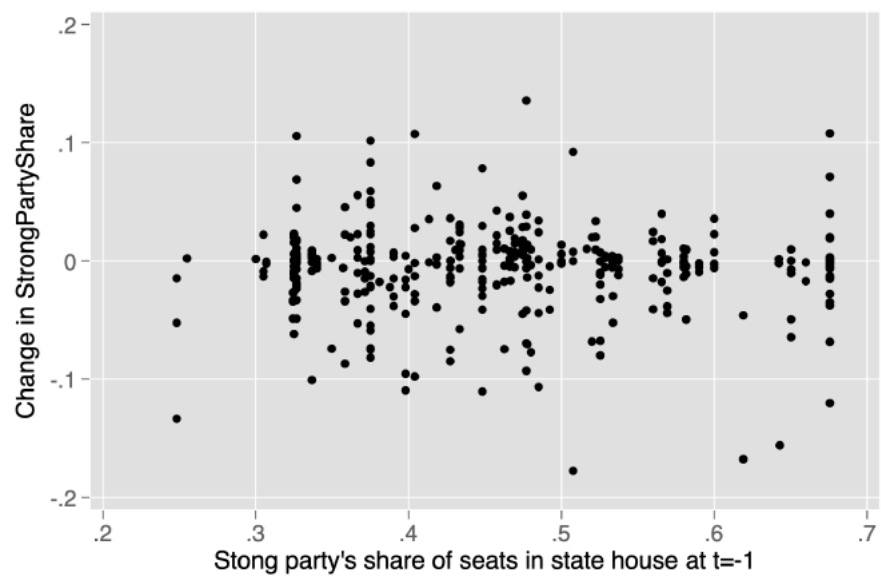
*Notes:* The figure shows binned averages of the share of female candidates in primary elections against the share of female constituents in the district. The red line indicates equal shares.

Figure A7: Economic representativeness of candidate pool



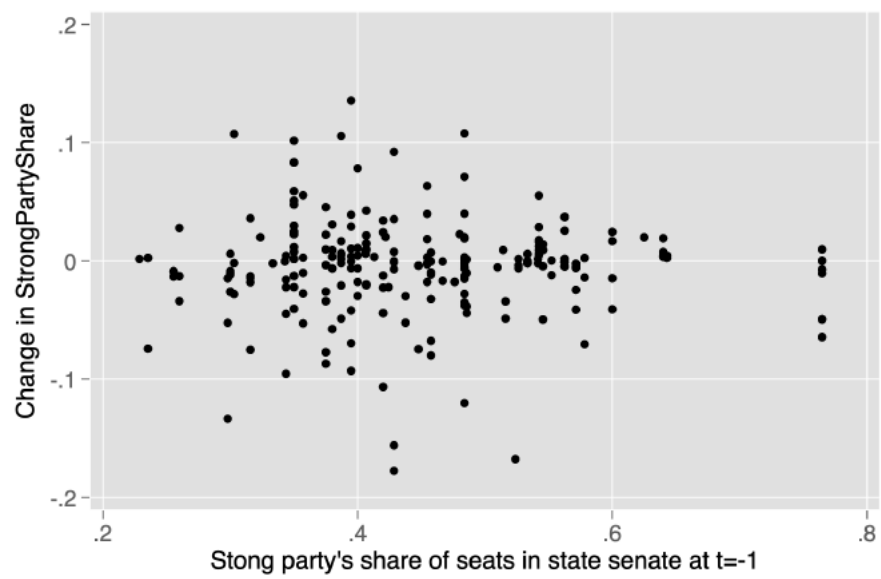
Notes: The figure shows binned averages of the median household income in the ZIP codes of candidates in primary elections against the median household income in their district. The red line indicates equal means.

Figure A8: Seat distribution in state house and change in *StrongPartyShare*



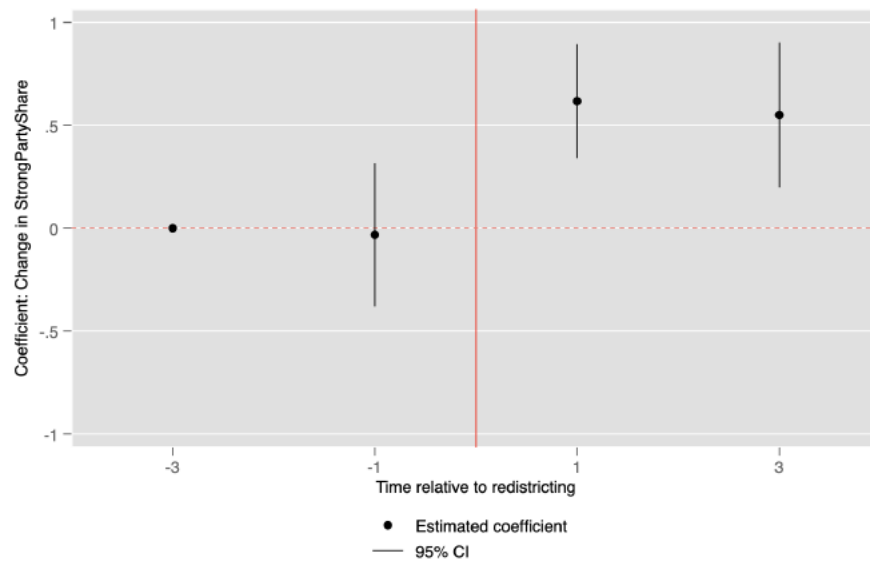
*Notes:* The figure shows the change in *StrongPartyShare* due to redistricting and its relation to distribution of seats in the state house immediately before redistricting. The sample of districts is restricted to only include cases where the strong party does not have sole control over the redistricting procedure.

Figure A9: Seat distribution in state house and change in *StrongPartyShare*



*Notes:* The figure shows the change in *StrongPartyShare* due to redistricting and its relation to distribution of seats in the state senate immediately before redistricting. The sample of districts is restricted to only include cases where the strong party does not have sole control over the redistricting procedure.

Figure A10: Vote share of candidate in strong party and change in *StrongPartyShare*: Common trend



*Notes:* The figure plots coefficients and 95% confidence intervals for regressions of the change in *StrongPartyShare* interacted with time relative to redistricting on the the general election vote share of the candidate in the strong party. The vertical red line indicates the timing of redistricting. The sample only includes elections when neither party fully controlled the redistricting procedure in the respective state and redistricting cycle.

Table A1: Estimates for fractional logit on county level voter registration

	(1)	(2)	(3)	(4)
	% of Dem.	% of Dem.	% of Dem.	% of Dem.
Population density	-0.000 (0.000)	0.001 (0.000)	0.000 (0.000)	0.000** (0.000)
% urban population	0.174 (0.237)	0.005 (0.131)	-0.597*** (0.147)	-0.049 (0.191)
% population under 18	2.955** (1.225)	-3.594*** (0.926)	-1.157 (0.891)	2.447*** (0.717)
% population over 65	-11.896** (4.633)	1.495 (1.964)	0.180 (2.086)	-1.158 (1.666)
% single male hh.	-22.269** (9.627)	-11.327* (6.210)	-8.590 (8.610)	-23.174*** (8.195)
% single female hh.	-5.301 (12.257)	-16.697** (6.860)	-12.432 (10.772)	-18.168* (9.748)
% married hh. with children	-22.402** (9.803)	-18.400*** (5.885)	-14.478 (9.256)	-23.129*** (8.189)
% married hh. without children	-14.326 (10.481)	-17.358*** (6.525)	-13.747 (9.404)	-27.796*** (8.977)
% single male hh. with children	-55.087*** (12.841)	-19.214** (8.080)	1.669 (12.746)	-18.640 (13.044)
% single male hh. without children	43.732*** (15.003)	-4.097 (9.189)	-12.967 (14.392)	-9.975 (17.288)
% single female hh. with children	-12.768 (10.779)	-3.267 (6.427)	-8.510 (10.476)	-19.366** (9.000)
% single female hh. without children	-20.588 (14.346)	-6.533 (7.697)	0.833 (11.302)	-4.906 (11.573)
% non-family hh. with a male head of hh.	-32.022* (18.768)	-2.985 (10.981)	-18.900 (17.361)	-31.205** (14.360)
% vacant houses	0.876 (0.652)	-3.272*** (0.679)	-0.911* (0.506)	1.023** (0.482)
% renter owned houses	1.339 (0.914)	-6.832*** (0.883)	-0.693 (0.952)	-1.473 (0.949)
Constant	14.745 (9.902)	18.519*** (6.121)	13.180 (9.300)	20.605** (8.672)
N	203	363	510	246

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Fractional logit estimates. Robust standard errors in parentheses. Columns 1-4 refer to 4 different regions. The sample are counties for which data on voter registration is available in Leip (2013).

Table A2: Experience and identity match: Results in contested primary elections

	Strong party		Weak party	
<i>Candidate vote share</i>				
Mean	0.347	0.347	0.353	0.353
Experience	0.138*** (0.014)	0.134*** (0.014)	0.154*** (0.014)	0.152*** (0.015)
Ethnicity match	0.058*** (0.015)	0.083*** (0.016)	-0.022** (0.010)	-0.054*** (0.017)
Gender match	0.124 (0.150)	0.320* (0.171)	-0.064 (0.111)	0.008 (0.156)
Income match	-0.007 (0.022)	-0.009 (0.023)	-0.039* (0.022)	-0.038 (0.024)
Gender, ethnicity, income	✓	✓	✓	✓
Competition controls	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
State FE's		✓		✓
N	1489	1489	1559	1559
R-squared	0.749	0.755	0.570	0.574
<i>Primary winner</i>				
Mean	0.338	0.338	0.352	0.352
Experience	0.222*** (0.033)	0.219*** (0.035)	0.316*** (0.031)	0.313*** (0.032)
Ethnicity match	0.086*** (0.024)	0.114*** (0.028)	-0.030 (0.031)	-0.064 (0.050)
Gender match	0.342 (0.282)	0.678** (0.317)	-0.099 (0.275)	0.076 (0.407)
Income match	-0.004 (0.059)	-0.005 (0.062)	-0.023 (0.060)	-0.022 (0.064)
Gender, ethnicity, income	✓	✓	✓	✓
Competition controls	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
State FE's		✓		✓
N	1489	1489	1559	1559
R-squared	0.495	0.500	0.271	0.274

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Competition controls are dummy variables for every observed combination of experienced and inexperienced candidates in the primary election. Additional control variables include indicator variables for incumbents and five dummy variables for primary type. The sample is restricted to only include primary elections with at least 2 candidates. Singleton groups are excluded from the estimation.

Table A3: Experience and identity match: Results in contested general elections

	All parties	
<i>Candidate vote share</i>		
Mean	0.474	0.474
Experience	0.078*** (0.008)	0.078*** (0.009)
Ethnicity match	0.017 (0.013)	0.020* (0.011)
Gender match	0.004 (0.138)	-0.095 (0.087)
Income match	-0.021** (0.008)	-0.015* (0.009)
Gender, ethnicity, income	✓	✓
Competition controls	✓	✓
Additional controls	✓	✓
Year FE's	✓	✓
State FE's		✓
N	1436	1436
R-squared	0.771	0.790
<i>Election winner</i>		
Mean	0.478	0.478
Experience	0.206*** (0.024)	0.202*** (0.024)
Ethnicity match	0.011 (0.017)	0.032 (0.021)
Gender match	-0.102 (0.164)	0.062 (0.235)
Income match	-0.054 (0.038)	-0.051 (0.042)
Gender, ethnicity, income	✓	✓
Competition controls	✓	✓
Additional controls	✓	✓
Year FE's	✓	✓
State FE's		✓
N	1440	1440
R-squared	0.743	0.745

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Competition controls are dummy variables for experienced and inexperienced competitors in the general election. Additional control variables include the predicted vote share of the candidate's party and indicator variables for incumbents. The sample only includes elections in states and redistricting cycles in which none of the parties solely controlled the redistricting procedure. The sample is further restricted to only include elections with 2 candidates. Singleton groups are excluded from the estimation.



Table A4: Entry of experienced candidates and age

<i>Experienced candidate</i>	Strong party			Weak party		
	All ages	≤ 45	≥ 66	All ages	≤ 45	≥ 66
Mean	0.757	0.644	0.919	0.526	0.393	0.787
<i>StrongPartyShare</i>	0.429 (0.379)	2.075 (1.752)	-0.855 (1.273)	-2.066** (0.853)	-2.622 (2.446)	0.620 (1.186)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓
N	1663	289	198	1214	272	94
R-squared	0.541	0.748	0.685	0.653	0.745	0.956

*Notes:* \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Sample size is significantly reduced because age data is not available for all candidates. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A5: Decreased competition and experience: No partisan redistricting

	<u>All parties</u>		<u>Strong party</u>		<u>Weak party</u>	
<b><i>Experienced candidate</i></b>						
Mean	0.445	0.445	0.535	0.535	0.347	0.347
<i>StrongPartyShare</i>	-0.804*	-0.676*	0.905	1.180*	-1.783***	-1.737***
	(0.422)	(0.353)	(0.802)	(0.675)	(0.560)	(0.457)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3129	3129	1623	1623	1506	1506
R-squared	0.407	0.413	0.351	0.374	0.438	0.446
<b><i>Experienced candidate (excluding incumbents)</i></b>						
Mean	0.238	0.238	0.278	0.278	0.206	0.206
<i>StrongPartyShare</i>	-0.875*	-0.807	2.524*	1.910	-1.603***	-1.405**
	(0.465)	(0.649)	(1.278)	(1.161)	(0.536)	(0.691)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2239	2239	1008	1008	1231	1231
R-squared	0.294	0.318	0.294	0.323	0.296	0.324

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A6: Decreased competition and ethnicity match: No partisan redistricting

	All parties		Strong party		Weak party	
<b><i>Candidate ethnicity match</i></b>						
Mean	0.680	0.680	0.685	0.685	0.675	0.675
<i>StrongPartyShare</i>	-0.137	-0.177	0.308	0.256	-0.629***	-0.659***
	(0.270)	(0.258)	(0.417)	(0.387)	(0.196)	(0.167)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2670	2670	1379	1379	1291	1291
R-squared	0.878	0.879	0.853	0.854	0.920	0.920
<b><i>Candidate ethnicity match (excluding incumbents)</i></b>						
Mean	0.677	0.677	0.705	0.705	0.655	0.655
<i>StrongPartyShare</i>	-0.428	-0.470*	0.443	0.221	-0.755***	-0.741***
	(0.302)	(0.263)	(0.529)	(0.530)	(0.248)	(0.222)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	1945	1945	859	859	1086	1086
R-squared	0.878	0.878	0.836	0.838	0.913	0.914

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A7: Decreased competition and gender match: No partisan redistricting

	All parties		Strong party		Weak party	
<b><i>Candidate gender match</i></b>						
Mean	0.343	0.343	0.341	0.341	0.344	0.344
<i>StrongPartyShare</i>	0.028	0.027	-0.001	-0.004	0.044	0.044
	(0.037)	(0.039)	(0.032)	(0.031)	(0.051)	(0.051)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2670	2670	1379	1379	1291	1291
R-squared	0.949	0.949	0.951	0.952	0.947	0.948
<b><i>Candidate gender match (excluding incumbents)</i></b>						
Mean	0.343	0.343	0.342	0.342	0.344	0.344
<i>StrongPartyShare</i>	0.057	0.055	-0.005	-0.012	0.071	0.061
	(0.052)	(0.053)	(0.058)	(0.054)	(0.055)	(0.053)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	1945	1945	859	859	1086	1086
R-squared	0.947	0.947	0.956	0.956	0.941	0.942

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A8: Decreased competition and income match: No partisan redistricting

	All parties		Strong party		Weak party	
<b><i>Candidate income match</i></b>						
Mean	0.771	0.771	0.764	0.764	0.778	0.778
<i>StrongPartyShare</i>	-0.221	-0.233	0.007	0.052	-0.461	-0.515*
	(0.262)	(0.256)	(0.270)	(0.295)	(0.320)	(0.305)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2670	2670	1379	1379	1291	1291
R-squared	0.418	0.419	0.431	0.435	0.408	0.410
<b><i>Candidate income match (excluding incumbents)</i></b>						
Mean	0.771	0.771	0.769	0.769	0.773	0.773
<i>StrongPartyShare</i>	-0.433	-0.420	0.034	0.348	-0.607*	-0.617*
	(0.370)	(0.363)	(0.835)	(0.949)	(0.353)	(0.325)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	1945	1945	859	859	1086	1086
R-squared	0.381	0.383	0.380	0.384	0.388	0.389

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A9: Decreased competition and experience: First time candidates

	<u>All parties</u>		<u>Strong party</u>		<u>Weak party</u>	
<b><i>Experienced candidate</i></b>						
Mean	0.225	0.225	0.270	0.270	0.189	0.189
<i>StrongPartyShare</i>	-0.459 (0.524)	-0.371 (0.560)	1.954* (0.979)	1.431 (0.970)	-1.218** (0.494)	-1.014* (0.577)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3008	3008	1326	1326	1682	1682
R-squared	0.310	0.331	0.310	0.335	0.307	0.326
<b><i>Experienced candidate (excluding incumbents)</i></b>						
Mean	0.220	0.220	0.263	0.263	0.185	0.185
<i>StrongPartyShare</i>	-0.438 (0.514)	-0.343 (0.554)	2.268** (1.010)	1.813* (0.971)	-1.258** (0.500)	-1.081* (0.602)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2984	2984	1311	1311	1673	1673
R-squared	0.312	0.335	0.316	0.342	0.307	0.330

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A10: Decreased competition and ethnicity match: First time candidates

	All parties		Strong party		Weak party	
<b><i>Candidate ethnicity match</i></b>						
Mean	0.670	0.670	0.681	0.681	0.663	0.663
<i>StrongPartyShare</i>	-0.518**	-0.539***	-0.225	-0.342	-0.615***	-0.667***
	(0.201)	(0.194)	(0.534)	(0.474)	(0.214)	(0.211)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2614	2614	1077	1077	1537	1537
R-squared	0.854	0.855	0.853	0.855	0.863	0.865
<b><i>Candidate ethnicity match (excluding incumbents)</i></b>						
Mean	0.671	0.671	0.683	0.683	0.663	0.663
<i>StrongPartyShare</i>	-0.525**	-0.549***	-0.235	-0.361	-0.621***	-0.672***
	(0.203)	(0.196)	(0.542)	(0.477)	(0.215)	(0.212)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2596	2596	1063	1063	1533	1533
R-squared	0.853	0.854	0.851	0.853	0.863	0.864

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A11: Decreased competition and gender match: First time candidates

	All parties		Strong party		Weak party	
<b><i>Candidate gender match</i></b>						
Mean	0.341	0.341	0.340	0.340	0.342	0.342
<i>StrongPartyShare</i>	-0.001 (0.032)	-0.001 (0.031)	-0.034 (0.046)	-0.032 (0.044)	0.008 (0.042)	0.005 (0.040)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2614	2614	1077	1077	1537	1537
R-squared	0.944	0.944	0.943	0.943	0.946	0.946
<b><i>Candidate gender match (excluding incumbents)</i></b>						
Mean	0.341	0.341	0.340	0.340	0.342	0.342
<i>StrongPartyShare</i>	-0.001 (0.032)	0.000 (0.031)	-0.033 (0.044)	-0.033 (0.040)	0.008 (0.042)	0.005 (0.040)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2596	2596	1063	1063	1533	1533
R-squared	0.945	0.945	0.945	0.945	0.946	0.946

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.



Table A12: Decreased competition and income match: First time candidates

	All parties		Strong party		Weak party	
<b><i>Candidate income match</i></b>						
Mean	0.766	0.766	0.766	0.766	0.767	0.767
<i>StrongPartyShare</i>	-0.333	-0.298	-0.167	-0.051	-0.433	-0.444
	(0.280)	(0.283)	(0.704)	(0.676)	(0.415)	(0.404)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2610	2610	1073	1073	1537	1537
R-squared	0.397	0.400	0.426	0.432	0.382	0.384
<b><i>Candidate income match (excluding incumbents)</i></b>						
Mean	0.766	0.766	0.764	0.764	0.767	0.767
<i>StrongPartyShare</i>	-0.308	-0.271	-0.087	0.034	-0.430	-0.443
	(0.281)	(0.284)	(0.724)	(0.696)	(0.414)	(0.403)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2592	2592	1059	1059	1533	1533
R-squared	0.397	0.400	0.426	0.431	0.382	0.384

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A13: Decreased competition and ethnicity match (*StrongPartyShare* includes information on gender and race or ethnicity)

	All parties		Strong party		Weak party	
<b><i>Candidate ethnicity match</i></b>						
Mean	0.668	0.668	0.666	0.666	0.670	0.670
<i>StrongPartyShare</i>	-0.323** (0.151)	-0.329** (0.148)	0.045 (0.301)	0.023 (0.300)	-0.644*** (0.184)	-0.655*** (0.175)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	4598	4598	2245	2245	2353	2353
R-squared	0.865	0.865	0.864	0.865	0.869	0.870
<b><i>Candidate ethnicity match (excluding incumbents)</i></b>						
Mean	0.666	0.666	0.683	0.683	0.653	0.653
<i>StrongPartyShare</i>	-0.446** (0.178)	-0.448*** (0.167)	-0.018 (0.502)	-0.179 (0.452)	-0.629*** (0.194)	-0.635*** (0.180)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3471	3471	1444	1444	2027	2027
R-squared	0.858	0.859	0.849	0.851	0.869	0.870

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A14: Decreased competition and gender match (*StrongPartyShare* includes information on gender and race or ethnicity)

	All parties		Strong party		Weak party	
<b><i>Candidate gender match</i></b>						
Mean	0.340	0.340	0.339	0.339	0.342	0.342
<i>StrongPartyShare</i>	-0.023 (0.025)	-0.023 (0.025)	-0.026 (0.029)	-0.026 (0.029)	-0.022 (0.038)	-0.024 (0.038)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	4598	4598	2245	2245	2353	2353
R-squared	0.940	0.941	0.936	0.936	0.947	0.947
<b><i>Candidate gender match (excluding incumbents)</i></b>						
Mean	0.341	0.341	0.340	0.340	0.342	0.342
<i>StrongPartyShare</i>	0.004 (0.034)	0.005 (0.033)	-0.015 (0.040)	-0.013 (0.036)	0.010 (0.042)	0.005 (0.041)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3471	3471	1444	1444	2027	2027
R-squared	0.942	0.942	0.943	0.944	0.943	0.943

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A15: Decreased competition and income match (*StrongPartyShare* includes information on gender and race or ethnicity)

	<u>All parties</u>		<u>Strong party</u>		<u>Weak party</u>	
<b><i>Candidate income match</i></b>						
Mean	0.768	0.768	0.766	0.766	0.769	0.769
<i>StrongPartyShare</i>	-0.322 (0.202)	-0.322 (0.205)	-0.069 (0.359)	-0.059 (0.360)	-0.536** (0.246)	-0.590** (0.254)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	4598	4598	2245	2245	2353	2353
R-squared	0.413	0.415	0.433	0.437	0.397	0.399
<b><i>Candidate income match (excluding incumbents)</i></b>						
Mean	0.766	0.766	0.766	0.766	0.765	0.765
<i>StrongPartyShare</i>	-0.322 (0.242)	-0.314 (0.240)	0.046 (0.590)	0.067 (0.561)	-0.488* (0.260)	-0.542** (0.257)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	3471	3471	1444	1444	2027	2027
R-squared	0.379	0.381	0.392	0.398	0.373	0.376

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A16: Decreased competition and experience: No open seats in cycle

	<u>All parties</u>		<u>Strong party</u>		<u>Weak party</u>	
<b><i>Experienced candidate</i></b>						
Mean	0.413	0.413	0.531	0.531	0.282	0.282
<i>StrongPartyShare</i>	0.181 (0.351)	0.157 (0.346)	1.165* (0.648)	1.140* (0.618)	-0.721 (0.491)	-0.807 (0.505)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2891	2891	1516	1516	1375	1375
R-squared	0.516	0.518	0.446	0.454	0.540	0.542
<b><i>Experienced candidate (excluding incumbents)</i></b>						
Mean	0.155	0.155	0.190	0.190	0.129	0.129
<i>StrongPartyShare</i>	-0.144 (0.606)	-0.104 (0.601)	2.804** (1.395)	3.032** (1.401)	-1.331** (0.590)	-1.203* (0.604)
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	1966	1966	842	842	1124	1124
R-squared	0.292	0.295	0.286	0.292	0.305	0.309

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party, a dummy variable for elections with two incumbents and five dummy variables for primary type. Singleton groups are excluded from the estimation.

Table A17: Decreased competition and ethnicity match: No open seats in cycle

	All parties		Strong party		Weak party	
<b><i>Candidate ethnicity match</i></b>						
Mean	0.645	0.645	0.646	0.646	0.644	0.644
<i>StrongPartyShare</i>	-0.375*	-0.406**	-0.239	-0.231	-0.526**	-0.587**
	(0.204)	(0.200)	(0.390)	(0.408)	(0.240)	(0.224)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2447	2447	1164	1164	1283	1283
R-squared	0.865	0.866	0.870	0.870	0.871	0.872
<b><i>Candidate ethnicity match (excluding incumbents)</i></b>						
Mean	0.636	0.636	0.665	0.665	0.617	0.617
<i>StrongPartyShare</i>	-0.553*	-0.583**	-0.697	-0.761	-0.523	-0.605*
	(0.289)	(0.271)	(0.557)	(0.529)	(0.360)	(0.315)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	1737	1737	654	654	1083	1083
R-squared	0.861	0.862	0.854	0.856	0.872	0.874

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A18: Decreased competition and gender match: No open seats in cycle

	All parties		Strong party		Weak party	
<b><i>Candidate gender match</i></b>						
Mean	0.340	0.340	0.337	0.337	0.344	0.344
<i>StrongPartyShare</i>	-0.026 (0.028)	-0.029 (0.027)	-0.012 (0.041)	-0.014 (0.041)	-0.039 (0.030)	-0.043 (0.031)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2447	2447	1164	1164	1283	1283
R-squared	0.940	0.940	0.938	0.938	0.943	0.944
<b><i>Candidate gender match (excluding incumbents)</i></b>						
Mean	0.342	0.342	0.338	0.338	0.344	0.344
<i>StrongPartyShare</i>	-0.002 (0.033)	-0.003 (0.033)	0.017 (0.038)	0.019 (0.039)	-0.013 (0.040)	-0.019 (0.040)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	1737	1737	654	654	1083	1083
R-squared	0.942	0.942	0.952	0.952	0.937	0.937

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.

Table A19: Decreased competition and income match: No open seats in cycle

	All parties		Strong party		Weak party	
<b><i>Candidate income match</i></b>						
Mean	0.772	0.772	0.778	0.778	0.767	0.767
<i>StrongPartyShare</i>	-0.600*	-0.657*	-0.419	-0.402	-0.783**	-0.840***
	(0.337)	(0.338)	(0.584)	(0.587)	(0.311)	(0.298)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	2447	2447	1164	1164	1283	1283
R-squared	0.449	0.454	0.460	0.466	0.459	0.462
<b><i>Candidate income match (excluding incumbents)</i></b>						
Mean	0.767	0.767	0.784	0.784	0.757	0.757
<i>StrongPartyShare</i>	-0.533	-0.569	-0.515	-0.590	-0.569	-0.611*
	(0.417)	(0.405)	(0.906)	(0.901)	(0.357)	(0.341)
Gender, ethnicity, income	✓	✓	✓	✓	✓	✓
Last election result	✓	✓	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓	✓	✓
District-cycle-party FE's	✓	✓				
District-cycle FE's			✓	✓	✓	✓
Additional controls		✓		✓		✓
N	1737	1737	654	654	1083	1083
R-squared	0.414	0.421	0.391	0.402	0.439	0.443

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. Singleton groups are excluded from the estimation.



Table A20: Decreased competition and experience: Heterogeneity in pre-redistricting *StrongPartyShare*

Quantile: Pre-redistricting <i>StrongPartyShare</i>	Strong party		Weak party	
	1-3	4	1-3	4
<b><i>Experienced candidate</i></b>				
Mean	0.462	0.693	0.352	0.094
<i>StrongPartyShare</i>	1.221** (0.493)	-0.160 (0.811)	-1.817*** (0.432)	-0.421 (0.315)
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	2197	605	2216	435
R-squared	0.434	0.216	0.438	0.422
<b><i>Experienced candidate (excluding incumbents)</i></b>				
Mean	0.226	0.369	0.208	0.066
<i>StrongPartyShare</i>	2.069** (0.880)	0.543 (1.381)	-1.511*** (0.532)	-0.474 (0.282)
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	1494	271	1799	422
R-squared	0.327	0.316	0.299	0.265

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. The sample in each column is restricted to include only districts which lie in the mentioned quantile for the pre-redistricting *StrongPartyShare*. Singleton groups are excluded from the estimation.

Table A21: Decreased competition and ethnicity match: Heterogeneity in pre-redistricting *StrongPartyShare*

Quantile: Pre-redistricting <i>StrongPartyShare</i>	Strong party		Weak party	
	1-3	4	1-3	4
<b><i>Candidate ethnicity match</i></b>				
Mean	0.722	0.433	0.725	0.441
<i>StrongPartyShare</i>	0.082 (0.249)	-0.736** (0.304)	-0.637*** (0.188)	-0.515*** (0.179)
Gender, ethnicity, income	✓	✓	✓	✓
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	1819	510	1887	382
R-squared	0.829	0.902	0.903	0.853
<b><i>Candidate ethnicity match (excluding incumbents)</i></b>				
Mean	0.713	0.423	0.717	0.434
<i>StrongPartyShare</i>	-0.169 (0.302)	-0.449 (0.476)	-0.544** (0.225)	-0.498*** (0.172)
Gender, ethnicity, income	✓	✓	✓	✓
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	1303	229	1566	373
R-squared	0.829	0.889	0.894	0.845

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. The sample in each column is restricted to include only districts which lie in the mentioned quantile for the pre-redistricting *StrongPartyShare*. Singleton groups are excluded from the estimation.

Table A22: Decreased competition and gender match: Heterogeneity in pre-redistricting *StrongPartyShare*

Quantile: Pre-redistricting <i>StrongPartyShare</i>	Strong party		Weak party	
	1-3	4	1-3	4
<b><i>Candidate gender match</i></b>				
Mean	0.344	0.323	0.343	0.337
<i>StrongPartyShare</i>	-0.012 (0.027)	-0.063 (0.041)	-0.005 (0.026)	0.075 (0.069)
Gender, ethnicity, income	✓	✓	✓	✓
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	1819	510	1887	382
R-squared	0.961	0.901	0.965	0.870
<b><i>Candidate gender match (excluding incumbents)</i></b>				
Mean	0.343	0.324	0.343	0.337
<i>StrongPartyShare</i>	-0.046 (0.032)	0.023 (0.121)	0.008 (0.020)	0.081 (0.067)
Gender, ethnicity, income	✓	✓	✓	✓
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	1303	229	1566	373
R-squared	0.964	0.922	0.964	0.870

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. The sample in each column is restricted to include only districts which lie in the mentioned quantile for the pre-redistricting *StrongPartyShare*. Singleton groups are excluded from the estimation.

Table A23: Decreased competition and income match: Heterogeneity in pre-redistricting *StrongPartyShare*

Quantile: Pre-redistricting <i>StrongPartyShare</i>	Strong party		Weak party	
	1-3	4	1-3	4
<b><i>Candidate income match</i></b>				
Mean	0.771	0.762	0.771	0.742
<i>StrongPartyShare</i>	-0.129 (0.324)	0.491 (0.502)	-0.754*** (0.221)	0.491 (0.337)
Gender, ethnicity, income	✓	✓	✓	✓
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	1819	510	1887	382
R-squared	0.438	0.481	0.379	0.490
<b><i>Candidate income match (excluding incumbents)</i></b>				
Mean	0.771	0.763	0.768	0.738
<i>StrongPartyShare</i>	-0.565 (0.437)	0.382 (0.850)	-0.748*** (0.256)	0.476 (0.319)
Gender, ethnicity, income	✓	✓	✓	✓
Last election result	✓	✓	✓	✓
Year FE's	✓	✓	✓	✓
District-cycle FE's	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓
N	1303	229	1566	373
R-squared	0.400	0.457	0.356	0.480

Notes: \*\*\* :  $p < 0.01$ , \*\* :  $p < 0.05$ , \* :  $p < 0.1$ . Robust standard errors in parentheses are clustered at the state-cycle level. Control variables include indicator variables for seats held by the strong party and open seats. Furthermore a dummy variable for elections with two incumbents and five dummy variables for primary type are included. The sample in each column is restricted to include only districts which lie in the mentioned quantile for the pre-redistricting *StrongPartyShare*. Singleton groups are excluded from the estimation.