Democratizing the Party: The Effects of Primary Election Reforms in Ghana

Nahomi Ichino   Noah L. Nathan[*]

January 13, 2021

Abstract

The recent expansion of the primary electorate by one of Ghana's major parties offers a rare opportunity to assess the effects of franchise extensions in contemporary new democracies. Using an original dataset on candidate entry and nominations, this article shows that expanding the primary electorate opened paths to office for politicians from social groups that were previously excluded, such as women and ethnic groups outside the party's core national coalition. The authors propose that democratizing candidate selection has two consequences in patronage-oriented political systems: vote buying will become a less effective strategy and the electorate will become more diverse. These changes, in turn, affect the types of politicians who seek and win legislative nominations. This suggests that a simple shift in who votes in intraparty primaries can be a key institutional mechanism for improving the descriptive representation of women and other under-represented groups.

Forthcoming, *British Journal of Political Science*

The social identities of elected politicians – such as their gender and ethnicity – affect who in society is favored with state resources and whose preferences are represented in government (Pande, 2003; Chattopadhyay and Duflo, 2004). But politicians with particular social identities can only be elected if they appear as candidates on the ballot. The choice sets presented to voters are determined by the candidate selection institutions adopted by political parties. In new democracies, parties increasingly use primary elections to select their nominees for legislative office (Field and Siavelis, 2008), which may have important consequences for the representation of different identity groups. Primaries increase the number and types of voters who decide which politicians will reach the general election. Characteristics of the primary electorate may affect politicians' decisions about whether to even compete for a party's nomination.

We investigate how expanding the size and demographic diversity of the primary electorate affects the types of candidates who choose to compete in primaries and the characteristics of the nominees who appear on the general election ballot. We focus on the short-run effects of a reform to the legislative primary electorate in the National Democratic Congress (NDC), one of Ghana's two major political parties. Previously, the NDC primary electorate in each parliamentary constituency comprised roughly 200 to 800 local party branch executives. Leading up to the 2016 elections, however, the NDC opened primary voting to all rank-and-file members, expanding the primary electorate by an order of magnitude to roughly 2,000–10,000 voters per constituency. This increased the proportion of primary voters who are women or identify with groups outside the party's core ethnic base.

This reform allows us to assess the effects of expanding the primary electorate on candidate entry and selection because the New Patriotic Party (NPP), Ghana's other major party, provides an unusually good set of primaries for comparison. The two parties have similarly-sized, stable bases and have alternated in power. Crucially, they also have very similar organizational structures (Riedl, 2014), including similar gender gaps and ethnic disparities in local leadership and the same primary rules before 2016. Moreover, before the 2012 elections the NPP considered the same electorate expansion as the NDC for the same reasons and may still adopt it. While the stability and organizational parallels between these parties are unusual, they allow us to make a tight

comparison of the primary processes of two very similar parties in similar constituencies in the same electoral environment. Such a comparison would not be possible in other new democracies with more inchoate party systems or a dominant party.

More broadly, this case offers a rare window into the effects of franchise extensions on descriptive representation in the developing world, in terms of both gender and ethnicity. Because democratization is typically a bundle of reforms and universal suffrage was frequently established at independence, past attempts to study the effects of franchise extensions on the social identities of candidates in these settings have struggled to disentangle their effects from other contemporaneous changes (Josefsson, 2014). The effects of franchise extensions within primary elections can be more clearly isolated from those of other factors in our case.

We estimate the average effects of expanding the NDC's electorate on its pool of aspirants and the characteristics of its nominees (average treatment effect on the treated). The NDC and NPP are not exactly the same, and our setting does not provide a "natural" experiment. Instead, our analysis attempts to generate reasonable counterfactuals to NDC parliamentary primaries by creating matched sets with weighted NPP primaries. We apply optimal full matching (Rosenbaum, 1991, 2010) to an original dataset of biographical information on all aspirants seeking nominations in these two major parties in 2016, as well as a similar dataset for Ghana's 2012 elections.

We find that expanding the electorate changed the gender, ethnicity and political experience profiles of the pool of aspiring candidates and eventual nominees. It increased the average number of female aspirants and the probability that the nominee is a woman. In addition, more aspirants from ethnic groups outside the NDC's core coalition sought and won nominations. Neither change came at the expense of the party's incumbent Members of Parliament (MPs), however, as they became more likely to be renominated. Instead, the electorate expansion reduced the probability of nominating wealthy political newcomers – outsiders who were previously able to buy their way to nominations, sometimes even over long-serving incumbents.

We propose that these effects resulted from two shifts. First, expanding the electorate reduced the feasibility of direct vote buying (Lizzeri and Persico, 2004), the dominant mode of campaigning in small-electorate primaries. Second, having more women and ethnic groups outside the party's

3

core base in the electorate increased the expected success of aspirants from these groups.

After the reform, prospective female aspirants with ambitions to seek office, but without the financial resources and connections to the male-dominated patronage networks needed to buy votes, could expect to be more successful. They may also have expected to perform better in an electorate with more female voters. Aspirants from ethnic groups outside the party's core coalition could expect similar changes to their viability in primaries with more co-ethnic voters. These shifts also reduced the viability of wealthy political outsiders who lacked clear appeal among the broader party membership by limiting their ability to buy nominations. However, incumbent MPs often had the resources and reputations of constituency service needed to win primaries with any size of electorate.

We do not analyze the reforms' effects on who is eventually elected to Parliament. It is difficult to construct a reasonable counterfactual of who would have won the general election in 2016 had the NDC kept its previous primary system and potentially selected different nominees, but still competed against the same 2016 NPP nominees.[1] Descriptively, however, the 2016 general election results appear to be consistent with the NDC's reform opening a path to office for a new group of politicians. The NDC had the same number of female MPs (fourteen) in 2016 as in 2012, even though it won 42 fewer seats overall (106 vs. 148) – a 40% increase in the female share of the party's caucus in Parliament. Moreover, eight of the fourteen women elected in 2016 were new MPs who had not been nominated in 2012 under the smaller electorate.

This paper makes several contributions. First, we extend the literature on how electoral and party institutions, including parties' internal rules, affect women's descriptive representation (Norris and Lovenduski, 1995; Wangnerud, 2009; Bjarnegård and Zetterberg, 2019). Our finding that opening up primary voting aids female candidacy contrasts with qualitative evidence that the democratization of candidate selection reduces women's access to nominations (Hinojosa, 2012).[2] It

---

[1]Comparisons with NPP candidates selected through small-electorate primaries are not appropriate, as the two parties are competitors and the observed performance of NPP nominees is not independent of the NDC's candidates. Any within-constituency comparison of NDC candidates' performance between 2012 and 2016 is significantly confounded by the overall increase in the NPP's popularity between the elections, which occurred for reasons unrelated to the NDC's choice of candidates (Bob-Milliar and Paller, 2018).

[2]Hinojosa (2012) draws on case studies of Mexico and Chile to argue that decentralized procedures with larger electorates result in the nomination of fewer women compared with selection by narrow sets of party elites. She contends that elites are more concerned than local party actors about women's representation, and that potential

4

also overcomes some methodological challenges in other studies of the effects of franchise extensions on the election of women that are less able to isolate changes to the electorate from other contemporaneous political reforms (Josefsson, 2014).

We show that expanding the primary electorate can improve women's incorporation even in the short run, without changes to deeper constraints such as women's latent ambition (Fox and Lawless, 2014) or party elites' gendered recruitment efforts (Sanbonmatsu, 2002). Our findings also suggest that democratizing candidate selection may be an especially useful tool to improve women's representation in single-member district electoral systems, where the zero-sum nature of the contest is a particularly serious obstacle to women's candidacy. In these contexts, better-studied remedies like quotas and reserved seats may produce larger overall effects on women's representation than expanding the primary electorate (Duflo, 2005; Tripp and Kang, 2008; Krook, 2009), but they are often also more difficult to implement without backlash from men (Clayton, 2015).

Second, and more broadly, we extend the study of franchise extensions to contemporary new democracies and show that changing the social identities of the pool of politicians seeking elective office is a key channel through which these reforms can affect governance. Previous research has found that historical suffrage extensions in advanced democracies improved policy and substantive representation (Lott and Kenny, 1999; Aidt and Dallal, 2008; Berlinski and Dewan, 2011; Morgan-Collins and Teele, 2017), but it is unclear whether these effects can be attributed to changes in the identities of the politicians elected (Berlinski et al., 2014).

Expanding the Electorate: An Open Empirical Question

Whether women and other politicians from marginalized social identity groups seek and win office may be affected by electoral institutions, such as voting rules (Tripp and Kang, 2008) and quotas (Krook, 2009), and internal party institutions, such as candidate eligibility criteria (Bjarnegård and Zetterberg, 2019). However, prior studies offer little theoretical guidance on how expanding the franchise may affect who runs and wins in primaries or general elections in new democracies.[3]

women aspirants are generally reluctant to participate in competitive primaries.

[3]The recent literature on primaries in new democracies mostly explores their implications for general election performance and candidate quality (Carey and Polga-Hecimovich, 2006; Kemahlioglu et al., 2009; Ichino and Nathan,

Analytical results are not available for a general formal model of candidate entry that involves simultaneous entry by an undefinable pool of potential aspirants from multiple groups. Multiple equilibria are possible even with a simple voting model and a known number of candidates and identity groups (Dickson and Scheve, 2010). Different combinations of entrants can also generate different outcomes, depending on how aspirants split votes within segments of the electorate. Consistent with this ambiguity, research on historical suffrage extensions in today's advanced democracies finds mixed effects on the identities of elected politicians (Berlinski and Dewan, 2011; Berlinski et al., 2014; Morgan-Collins and Teele, 2017).

More generally, studies of primaries in advanced democracies offer little guidance because ideology, which is central to most existing models (e.g., Gerber and Morton 1998, Serra 2011), is not a key feature of electoral competition in many new democracies. Primary aspirants in these settings are differentiated less by policy than by their ability to deliver patronage to small groups of primary voters (Ichino and Nathan, 2012).

In new democracies, changes to the size and demographic composition of the electorate should be more likely to affect primaries through changes to these vote-buying exchanges. Increasing the size of an electorate can induce a shift away from clientelism because narrow particularistic appeals are less feasible with larger numbers of voters (Lizzeri and Persico, 2004). As vote buying declines, potential aspirants who have ambition but lack the resources or access to the patronage networks needed to be effective in clientelist primaries may expect to become more viable. Moreover, where there is significant ethnic voting, changing the composition of the electorate may also improve the expected prospects of aspirants from previously under-represented ethnic groups by bringing more potential supporters into the electorate.

Reforms to Parliamentary Primaries in Ghana

Since 1992, Ghana has held quadrennial elections for president and parliament, which has 275 members from single-member constituencies. The NDC and NPP dominate highly competitive elections and have alternated in government, with the NDC in power leading into 2016. These

2012, 2013b; Choi, 2018).

well-institutionalized parties have faced strong "isomorphic competitive pressures" to mirror each other's behavior and organizations; reforms adopted by one party are often quickly matched by the other (Riedl, 2014). Consequently, the NDC and NPP are unusually similar in many respects, including appeals to clientelism, the size of their ethnic bases, and their grassroots organizations and candidate selection procedures up until 2016.

The parties have very similar platforms, and policy is not a major axis of competition (Bob-Milliar, 2012a; Riedl, 2014; Nathan, 2019). Although the parties have rank-and-file members from every group and voting is not exclusively along ethnic lines, ethnicity is a strong determinant of partisanship. The NPP's base is the Akan; the NDC's core is a collection of smaller groups, most notably the Ewe, the Ga and many northern groups. These bases are stable and similarly-sized, leaving the two parties roughly evenly matched, with less subnational electoral volatility compared to more inchoate party systems (Ferree, 2010).

Although the parties differ in some respects such as their elite factional politics (Bob-Milliar, 2012b; Osei, 2016), they have nearly identical organizations. The NDC has a committee of nine and the NPP a committee of five polling-station executives at nearly all of Ghana's approximately 29,000 polling stations. These branch executives are the core "foot soldiers" in each party's campaign efforts and serve as clientelist brokers linking ordinary supporters to politicians and government officials (Bob-Milliar, 2012a; Brierley and Nathan, 2020).

They also comprised the primary electorates for each party for the 2004 through 2012 elections. These electorates were disproportionately male and composed of each party's core ethnic groups. For the NPP, the primary electorate numbered 200–800 in each constituency, with 5 branch leaders from each polling station. Women made up just 26 percent of this electorate. In addition, the data in Brierley and Nathan (2020) show that the proportion of Akans among NPP polling station-level executives is over 14 percentage points higher on average than in the estimated population of their polling stations' catchment areas. The NDC primary electorate in 2012 was similar, with four branch leaders voting per polling station; the women's organizer was typically the only female voter even though women made up approximately 37 percent of ordinary party members.[4] Ewes,

---

[4]See Q89C, Round 5, Afrobarometer.

Gas and Northerners are similarly over-represented in NDC branch leadership (Nathan, 2019).

With electorates of just a few hundred people, the main mode of competition in parliamentary primaries has been vote buying (Ichino and Nathan, 2012). Aspirants with significant personal wealth dominate. In some cases, aspirants return to their hometowns to "buy" nominations despite having little history of working for their party or living in the constituency. Branch-level party leaders often view primaries as their "cocoa season" – a key opportunity to extract rents. Aspirants pursue primary voters with gifts of TVs, motorbikes, school fees for their children and the like. Many aspirants openly admit in interviews to paying primary voters and report spending upwards of USD 75,000 to secure nominations.[5] They go to great lengths to monitor and enforce vote-buying transactions with the small number of party branch leaders voting in these elections.[6]

As an NPP aspirant in the 2011 primaries described his experience: "Everybody [the voters] was about 'What's in it for me? What have you brought for me? Politicians, after this election, they're not going to care about us any more...this is our chance.'"[7] Nominations often go to the highest bidder. One NPP MP described this bluntly, lamenting the "'money-ocracy' that has eaten into the fiber of our politics." He argued the primaries "elect money into office rather than electing people based on competence."[8]

In 2013, NDC leaders expanded their primary electorate for the 2016 election to include all ordinary party members in each constituency. This was an attempt to reduce the influence of vote buying and select more popular nominees. NDC leaders explicitly stated that they hoped vote buying would be less feasible with a larger electorate.[9] This decision was made before the outcome of the 2016 election could be anticipated with any confidence, and without knowledge of what nomination system the NPP would use.

In June 2015, the NPP held primaries for the 2016 elections under the same rules as for 2012.

---

[5]Interviews with aspirants conducted over multiple electoral cycles show that small-electorate primaries in each party are dominated by vote buying (Appendix, A2).

[6]For example, aspirants are frequently alleged to "camp" voters whose support they have paid for in a hotel the night before the primary to ensure that opponents cannot outbid them at the last minute.

[7]Interview with NPP primary aspirant, Central Region, 2 August 2011.

[8]Interview with NPP primary aspirant and incumbent MP, Brong Ahafo Region, 10 November 2015.

[9]For example, see "NDC's Expanded Electoral College Will Cure Vote-Buying - Ade Coker," *Citi FM Online*, 17 August 2015. The size of the electorate was the only change: NDC primaries kept the same electoral system and candidate eligibility criteria as NPP primaries and past NDC primaries.

The NDC primaries followed months later with a significantly expanded electorate, with most primaries held between late November 2015 and early 2016. The 2015/2016 NDC primary electorate was an order of magnitude larger than the NPP's, with 2,000–10,000 ordinary members voting on each nomination. Rather than taking place at a single venue in each constituency in each party's small-electorate primaries, the new NDC primaries involved voting at every polling station. The NDC assembled a new digital register of all its rank-and-file members to use as its voter roll.

This electorate expansion could have occurred in either party. In 2009, the NPP also carefully considered expanding its electorate, deciding between two proposals: marginally increasing the number of executives from each polling station that could vote or allowing all rank-and-file members to vote, similar to the reform later adopted by the NDC. The NPP's reform advocates made nearly identical arguments to those that NDC leaders would make in 2013, claiming that increasing the electorate would "lead to election of people who actually... work for the party", would "ensure that the selected... candidates... represent the popular will", and "reduce expenditure on internal party elections".[10] A main goal was similarly to reduce vote buying.[11] But due to the concern of one of the party's aspiring presidential candidates that a larger electorate would disadvantage his campaign, the party ultimately settled on the first proposal.

Possible Impacts of the Reform

The NDC's reform likely had two effects, which in turn may have changed the gender, ethnicity, and political experience profiles of the NDC's aspirants and eventual nominees for the 2016 election. First, because the party's rank-and-file membership is more gender balanced and ethnically diverse than its branch leadership, the new electorate had a greater proportion of women and non-core group members than before.

Secondly, as NDC leaders hoped, the electorate expansion may have changed the dominant campaign strategy in primaries. When the electorate is very small, primary aspirants can target

---

[10]Talking points circulated to national NPP leaders in 2009, obtained from a NPP National Executive Committee member, August 2011.

[11]For example, see Mahama Haruna, "Should all NPP Members Elect the Party's Presidential Candidate?," *Ghana Web*, 28 March 2009.

specific voters with personalized benefits and enforce exchanges through direct monitoring. But this becomes more costly as the number of voters shifts into the many thousands; previously common forms of individual-level monitoring become logistically infeasible.[12] Aspirants might switch to employing intermediaries to buy votes on their behalf, but this is less efficient than the direct exchanges that are possible with very small electorates (Stokes et al. 2013).

In interviews, many aspirants anticipated that expanding the electorate would have exactly these impacts on their vote buying. Among the aspirants in our survey (Appendix A2), which was conducted prior to the NDC's primaries, expanding the electorate was by far the most common suggestion for reducing vote buying in primaries.[13] For example, one interviewee predicted, "if we allow all registered members to vote, it will make it difficult for vote buying... how are you going to pay money to everybody?"[14] Another explained:

> "That problem can be reduced if we... expand our register. It will be very difficult to buy everybody. But if you leave it to five people [per polling station], hah! You'll see at the voting grounds that they are registering people for Hajj. 'I'll take you to Mecca.' And that could change somebody right there who had been supporting you. Or 'take this key to a motorbike'... All these things happen... But if it's an expanded album [register], I don't think that such things can be done. Nobody has that much resources to do those things."[15]

We expect these two changes to lead women to anticipate that they may perform better in primaries than before for two reasons. First, prospective female aspirants are likely generally disadvantaged relative to male aspirants in their access to economic opportunities needed to amass the private fortunes needed for extensive vote buying (Arriola and Johnson, 2014). They may also be less embedded in male-dominated party networks, with fewer pre-existing social connections to branch executives that would facilitate the monitoring and enforcement of vote buying in small-electorate primaries. Expanding the electorate likely reduced both of these disadvantages, leading women to expect to perform better.

---

[12]Even Ghana's well-institutionalized parties cannot monitor individual vote choices on a large enough scale to enforce a personalized *quid pro quo* with a substantial proportion of general election voters (Brierley and Nathan, 2020).

[13]Over one quarter (32 of 125) of these aspirants proposed expanding the electorate in response to the open-ended question "What do you think could be done to reduce vote buying?"

[14]Interview with NPP primary aspirant and incumbent MP, Ashanti Region, 5 November 2015.

[15]Interview with NPP primary aspirant and incumbent MP, Northern Region, 4 November 2015.

Secondly, aspirants may expect voters to use gender as an important heuristic for shared preferences,[16] leading prospective women aspirants to expect to perform better in an electorate with a significantly greater proportion of women voters. Female potential aspirants may also expect that female voters will be less intrinsically biased against them than male voters, all else equal. We do not have direct evidence about these expectations in Ghana. In qualitative interviews, female politicians make clear that they expect women candidates to be held back by damaging sexist attacks from male voters, similar to what Clayton et al. (2020) observe in Malawi.[17] Regardless of whether female voters truly prefer female candidates, having more female voters in the electorate should encourage more female aspirants to compete as long as these aspirants *expect* to receive more support from female voters. We therefore expect the electorate expansion to increase (a) the number of female aspirants and (b) consequently, the probability the nominee is female.

Based on a similar logic, we expect the electorate expansion to increase (c) the number of aspirants from outside the NDC's core ethnic coalition and (d) consequently, the probability that the nominee is from outside its core ethnic coalition. To the extent that aspirants have an advantage in winning support from voters with whom they share ethnic identities, aspirants from non-core ethnic groups that are under-represented in local party leadership may expect to perform better in primaries with more co-ethnic voters. Anticipating this, more potential aspirants from these groups compete for the nomination. With more aspirants from the party's non-core groups competing, the probability that one is nominated may also increase.[18]

We are agnostic on the reform's effects on (e) the number of aspirants from the NDC's core groups and (f) the total number of aspirants, as these would depend on possible coordination by potential aspirants.

---

[16]Gottlieb et al. (2018) show, for example, that female voters in Africa have different policy preferences in some domains from men.

[17]Focus group with NPP "women's wing" leaders, Greater Accra, 9 October 2015. For example, they describe female candidates and party activists as being denounced as "prostitutes" by male voters for becoming involved in politics.

A meta-analysis of survey experiments also shows that female voters are significantly more supportive of female candidates than male voters, on average (Schwarz and Coppock, 2020, 17-18). But most of these studies are from advanced democracies. The few studies conducted in other African countries have found contrasting results (Clayton et al., 2020; Phillips, 2020).

[18]As we explore further below, this depends on the total number of aspirants and the degree to which they may split the votes of non-core groups.

11

However, we expect that reduced feasibility of vote buying in primaries will reduce (g) the probability that the eventual nominee will be a newcomer to politics who has the wealth for vote buying, but not the public reputation and political experience for more general campaigning. Indeed, this is what NDC leaders hoped to accomplish by expanding the electorate. But we also expect the expanded electorate to (h) not reduce the probability that the nominee will be the incumbent. Incumbents have extended opportunities in office to build a reputation for service among local party members through their control of constituency development funds and by providing personal assistance to party supporters (Lindberg, 2010).[19] Unlike wealthy newcomers, incumbents can fall back on campaigning for re-nomination based on these reputations even if vote buying is more difficult.

Methods

We define the treatment as allowing rank-and-file members in a constituency to vote in primaries, with the party-constituency as the unit of analysis. Our outcomes are the numbers and characteristics of aspirants and the characteristics of nominees, and our estimand is the average treatment effect on the treated. We employ optimal full matching (Rosenbaum, 1991, 2010) to create sets of treatment (NDC) and control (NPP) primaries that are comparable on confounding and prognostic variables.

Optimal full matching creates sets of variable numbers of treated and control units that minimize the overall sum of distances between treated units and their matched control units on the propensity score. Treated units are each matched with at least one control, with each control used only once (Hansen and Klopfer, 2006). This is implemented with `optmatch` v0.9-13 in `R` v3.6.2.[20] Treatment effects are estimated by regressing each outcome on the treatment indicator with effect of the treatment on the treated (ETT) weighting, in which each set is weighted in proportion to the

---

[19]Assuming that incumbents' decisions to seek re-election depend on their performance in office, we expect that incumbents will *on average* have better reputations among party supporters than aspirants who have not had any similar opportunities to distribute state resources in the community.

[20]We use optimal full matching because we have a small number of control units relative to treated units compared with most matching analyses (Appendix, pg. A9). The Appendix (pg. A15) also discusses an alternative differences-in-differences strategy and reports ordinary least squares regressions without matching.

number of treated units in the set (Hansen, 2004). The units within each set are further weighted so that the weighted average of the outcomes in each set is the difference in means between the treatment and control units in the set.

We produce separate matches for each outcome variable, matching exactly on the 2012 value of each outcome (the lagged dependent variable) in each party-constituency and on a propensity score calculated from the remaining matching variables.[21] This means, for example, that for the analysis of the effect on the number of female aspirants, all units in a matched set of 2016 NDC and NPP primaries will have had the same number of female aspirants in 2012, when the parties had similarly-sized primary electorates. Past outcomes can be particularly informative for future outcomes since the Ghanaian party system is very stable, with negligible party switching by aspirants unlike in some other African countries. We regularly observe the same aspirants competing against each other in successive primaries.

Each outcome uses a different set of matching variables for its propensity score, but always includes four common variables from Ichino and Nathan (2012). The first is the party's vote share in the previous presidential election, which measures the attractiveness of a nomination to potential aspirants. The second is the party's vote share in the previous parliamentary election. When combined with presidential vote share, this reflects the performance of the party's past parliamentary nominee relative to the party's underlying support, measuring the previous nominee's quality, which would affect the entry decisions of potential aspirants in the next election. The third is population density, a proxy for economic development and a larger pool of potential wealthy aspirants who can buy votes. The fourth is fractionalization among the ethnic groups in the core coalition of the party. Additional matching variables specific to each outcome are described with the results.

The key non-interference assumption would be violated if aspirants sort across parties or constituencies in response to primary rules. But this is unlikely. To be eligible for a primary, aspirants in both parties had to be card-carrying party members in a constituency for four years, joining long before the electorate was changed. Since NPP primaries preceded NDC primaries, the assumption

---

[21]Missing data for pre-2012 primaries limit us from matching on trends in lagged outcomes (Appendix, pg. A8).

13

could also be violated if NDC primary voters chose a different nominee based on the outcome of the NPP's primary. But this would only make sense in the small set of very competitive constituencies where differences in the quality of candidates would affect the general election outcome. Primary voters generally lack the necessary information about both aspirants and other voters' preferences to successfully coordinate on a particular aspirant; in our extensive interviews, we did not encounter discussion of voters conditioning their choices in this way.

Data

Official results for primaries are not collated or publicly released in Ghana. By combining information from multiple sources, however, we are able to construct a dataset of NDC and NPP aspirants for all primaries that were held by early 2016 (Ichino and Nathan, 2021). Our final dataset includes 1,532 parliamentary aspirants across 272 constituencies for the NDC and 271 constituencies for the NPP, each out of 275. We define an aspirant as any candidate who publicly stated that he or she would seek a nomination. This includes some who announced that they would do so, but subsequently dropped out of the primary or were disqualified.[22] We drop some constituencies where we do not feel confident that we have identified all the aspirants, leaving us with 1494 aspirants across 219 constituencies for the NDC and 252 constituencies for the NPP.[23] Summary statistics are in the Appendix (pg. A2).

We began with official lists of aspirants on the ballot on primary day in each constituency, but lists were missing for many constituencies for both parties due to inconsistencies in record keeping and primaries delayed by legal disputes. We supplemented the official lists with media reports. From March 2015, when politicians in the NPP began announcing their candidacies, through January 2016, we saved every article mentioning primaries in either party from ten prominent news organizations listed in the Appendix (pg. A3). This yielded 1,950 articles, from which we coded names and biographical information, including professions and past government and party positions. This allowed us to identify many more aspirants not on the official lists. We used a similar

---

[22]In the Appendix (pg. A9), we show the robustness of our results to removing aspirants who dropped out before the primary was held.

[23]We examine differences between missing and non-missing party-constituencies in the Appendix (pg. A4).

procedure to compile data on the 1,285 aspirants from primaries held before the 2012 elections. To add additional biographical details, aspirant names were merged with data from the Parliament of Ghana website and official lists of government ministers dating to 2000. We also supplemented the biographical coding with an in-depth survey of 125 aspirants in the NPP primaries in November 2015. With these sources, the number of aspirants in the NDC ranges from 1 to 9 per constituency, with a median of 3 and mean of 3.5, and the number of aspirants in the NPP ranges from 1 to 8, with a median of 3 and a mean of 2.9.

We also coded each aspirant's ethnicity based on their name, which is easily connected to the main ethnic categories in Ghana. We employed a dictionary of 3,503 names of Ghanaian politicians, using the procedure outlined in the Appendix (pg. A4). We measured constituency demographic characteristics, including the ethnic composition of each parliamentary constituency, using geo-coded enumeration area-level 2010 census data.

Results

Gender

Figure 1 plots the average number of women aspirants and share of nominees who are female, by party, before (2012) and after (2016) the NDC's electorate expansion. It shows an overall increase in the number of NDC aspirants and a substantial rise in the share of NDC nominees who are female, but little change in the number of female NDC aspirants.

We turn to matching to adjust for potentially important underlying differences across party-constituencies. For the number of female aspirants in 2016, we match on a propensity score using the four common variables described earlier and the population share of Muslims in the constituency.[24] We exact match on the number of female aspirants in 2012. For whether the nominee is female, we account for a constituency's underlying receptivity to female leaders using the same matching variables as for the number of female aspirants, and again exact match on the lagged dependent

---

[24]Gender norms are generally more conservative in Muslim areas and may affect female aspirants' willingness to compete. We find that our effects on female nominees are concentrated in areas with low Muslim population shares (Appendix, pg. A10)
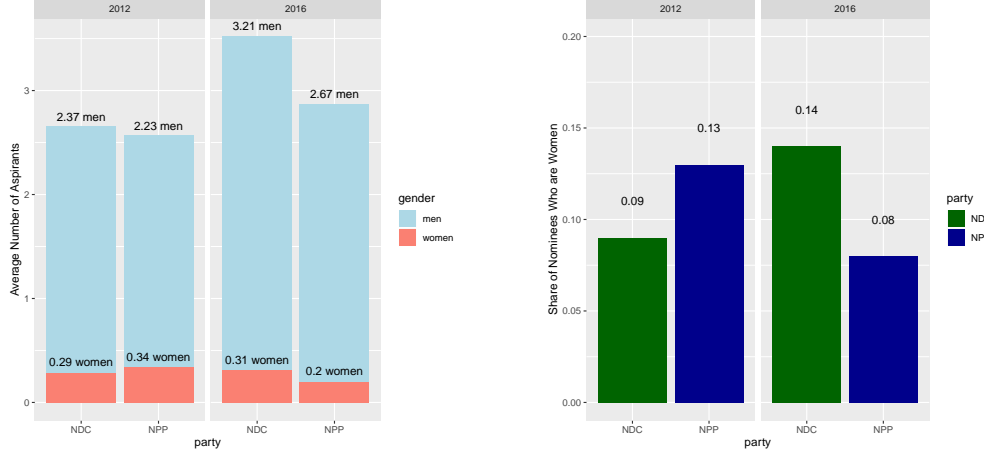
Figure 1: Women aspirants and nominees, by year and party

variable (female nominee in 2012). For the total number of aspirants, we use only the four common variables and exact match on the total number of aspirants in 2012.

Table 1 presents the number of sets and treated units remaining after matching for each outcome. We assess balance following Hansen and Bowers (2008). We compare treatment and control units within each set with the difference in means for each of our covariates, and assess their combination with a χ^2 test. Our matching procedure significantly improves balance (Table 1, rows 1-3).[25]

	Before Matching				After Matching				
	χ^2	df	p-value	n_T	χ^2	df	p-value	n_T	Sets
Number of female aspirants	92.13	6	< 0.01	170	1.87	5	0.87	168	87
Total number of aspirants	94.31	5	< 0.01	170	5.28	4	0.26	170	95
Nominee is female	92.51	7	< 0.01	170	0.29	6	1.00	170	97
Num. aspirants from non-core ethnic groups	79.51	8	< 0.01	150	4.65	7	0.70	150	83
Num. aspirants from party's core ethnic groups	92.10	8	< 0.01	150	4.96	7	0.66	148	73
Nominee belongs to party's core ethnic group	167.28	8	< 0.01	169	3.00	7	0.88	169	55
Nominee has only private sector background	116.41	5	< 0.01	211	1.34	4	0.85	211	114
Nominee is the incumbent	57.78	6	< 0.01	102	2.67	5	0.75	102	53

Table 1: *Overall Balance Before and After Optimal Full Matching.* The fourth and eighth columns show the number of treated units in our sample before and after matching. The last column shows the number of sets created. The number of treated units before and after matching differs when there are no control units available that have the same value of the 2012 outcome.

Effects are estimated by weighted least squares regression, with standard errors calculated as

---

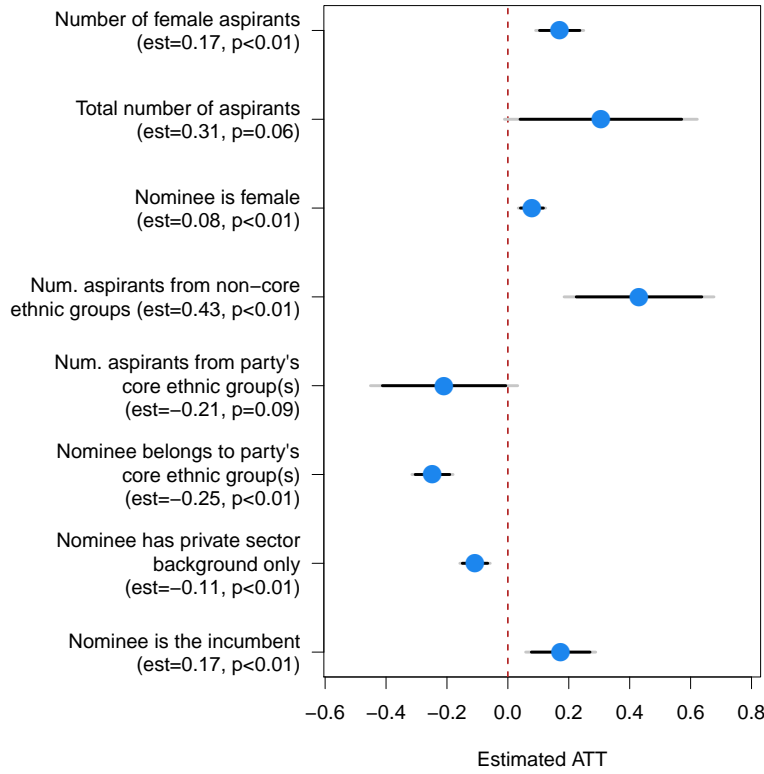[25]Balance for individual covariates is in the Appendix (pg. A12).

Figure 2: *Estimated average effects on the number of aspirants and characteristics of nominees.* The light and dark bars indicate 95% and 90% confidence intervals, respectively.

in Hansen (2004). Our results are at the top of Figure 2. We find that the NDC's reform increased the average number of female aspirants by 0.17 (p <0.01) and the average number of total aspirants by 0.31 (p=0.06). These are substantively large differences. In 2012, NDC primaries averaged 0.29 female aspirants out of 2.65 total, or only 11 per cent of all aspirants. An estimated effect of 0.17 additional female aspirants in Figure 2 is 55 per cent of the estimated average effect on the overall number of aspirants (0.31), equivalent to a 42 per cent (4.6 percentage point) increase in the overall proportion of female contestants compared to 2012. There are few historical estimates to benchmark our estimated effect, but it is comparable to those of more intensive interventions designed to increase female candidacy. For example, Bhavnani (2009) finds that reserving city council seats for women in India increased the proportion of candidates who were female in subsequent elections for those seats after reservations were withdrawn by 7.4 percentage points from a baseline of 4.4 per

cent.

Furthermore, expanding the electorate led to an 8-percentage-point increase in the probability that the NDC nominee was female (p=0.001), a large effect compared with a baseline of 9 per cent. In the Appendix (pg. A14), we show that women's advancements are not concentrated in constituencies where the NDC is unlikely to win.

The gender literature suggests two alternative explanations. First, learning about the NDC's reform may have increased women's ambition to seek office. This clearly affected women's ambition for a specific office. But latent ambition is deeply ingrained from early socialization (Fox and Lawless, 2014) and unlikely to change quickly. More importantly, in order to become eligible for the NDC primaries, potential aspirants needed to become active party members in their constituencies at least a year before the electorate expansion was announced. This leaves the pool of potential aspirants largely fixed in the short-run. It is more plausible that the expansion encouraged the entry of already-ambitious women who had been developing political careers in their constituencies, but were still undecided about whether to seek this particular office this cycle.

Secondly, personal recruitment (Sanbonmatsu, 2002) by party leaders or concurrent changes to formal candidate eligibility criteria (Bjarnegård and Zetterberg, 2019) may lead to changes in women's ambition in the short run (Fox and Lawless, 2011; Karpowitz et al., 2017). However, this is also unlikely to account for the results. Starting with the 2012 elections, the two parties adopted very similar efforts to recruit more women to run for parliament, including reducing registration fees by 50% for female aspirants. Moreover, the NPP has been more aggressive in encouraging female aspirants, moving first to reduce fees for women and publicly discussing reserving nominations for women.

Ethnicity

Next, we turn to ethnicity. Figure 3 shows the average number of aspirants and share of nominees from the parties' core and non-core ethnic groups. It indicates a substantial increase in the number of NDC aspirants from outside its core base, but no overall change in the ethnic composition of its nominees.
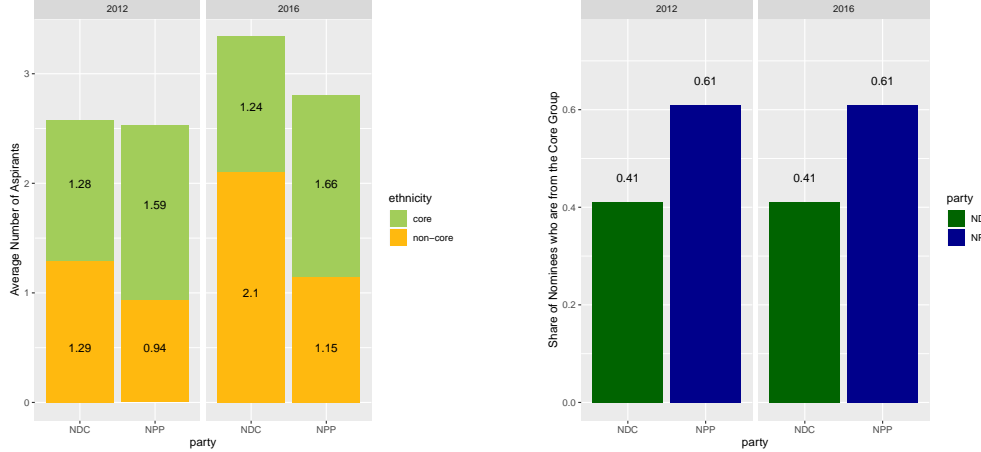
Figure 3: *Aspirants' and nominees' membership in their parties' core ethnic coalition, by party and year.*

As before, we employ optimal full matching to estimate the effects of the electorate expansion, with balance statistics in Table 1. For the number of aspirants from core and non-core ethnic groups, we exact match on the lagged dependent variable. We add the population share of the largest ethnic group in the constituency, segregation between ethnic groups associated with the NDC and the NPP, and segregation among the sub-groups within each party's national ethnic coalition to the four common matching variables for the propensity score.[26] Figure 2 shows that the electorate expansion increased the average number of aspirants from ethnic groups outside the party's core coalition by 0.43 ($p <$0.01), while reducing the average number of core group aspirants by 0.21 (p=0.09).

Whether the entry of new aspirants from non-core ethnic groups leads to more of these aspirants winning primaries depends on several factors, including the total number of aspirants, the ethnic demography of party members and the resources available to each aspirant. Although the latter two are unmeasured, we proxy for these factors using the constituency-level population share of the ethnic groups in the party's coalition, spatial segregation between core and non-core groups in the

---

[26]We include the size of the largest group because the NDC's reform is less likely to have an effect on the ethnic composition of entrants in constituencies dominated by a single group. We include the segregation measures because segregation may affect primary voters' beliefs about which candidates are most likely to target them with local public goods (Ichino and Nathan, 2013a), and hence whether additional aspirants from under-represented groups seek the nomination. There are diverse constituencies in both urban and rural areas (Appendix, pg. A6).

constituency, and the total number of aspirants competing in 2012.[27] Including these as matching variables, we find that expanding the electorate increased the probability that the NDC nominee is from an ethnic group outside of the party's core coalition by 22 percentage points ($p < 0.001$). To put this into context, this is larger than the impact of reserving seats for ethnic minority candidates in India (Bhavnani, 2017).

As with gender, party recruitment is an unlikely alternative explanation for these results. Although recruitment efforts are unobserved, national leaders from both parties have made public efforts to expand their ethnic appeal; the NPP has sought in recent years to shed its label as an "Akan party." If national party leaders favored aspirants from non-core ethnic groups, this is just as likely to have been done in the NPP as in the NDC, which is at odds with our results.[28]

Another alternative explanation is that aspirants from non-core ethnic groups may have won more often after the expansion due to greater coordination failures among voters from the party's core groups, not because non-core aspirants became more viable. However, this is inconsistent with our finding that the electorate expansion appears to have *decreased* the entry of aspirants from the NDC's core ethnic groups (see Figure 2; $p = 0.09$). The core-group vote likely became more concentrated - not less - and non-core aspirants could only have won more often if they gained support among primary voters.

Political Experience

Finally, we examine the political experience of nominees. Figure 4 indicates a large drop in the share of NDC nominees who are wealthy newcomers and a smaller increase in the share who are incumbents. We proxy for wealth earned outside government or party positions by coding whether the nominee has a high-level private sector business background (Pinkston, 2016). We code an aspirant as having private sector experience if there is biographical information that he or she has served in a managerial or executive position in any private business or worked in a well-compensated

---

[27]These variables are discussed in greater detail in the Appendix (pg. A8).

[28]We have anecdotal evidence, for example, of NPP leaders intervening in the 2008 and 2012 primaries to help aspirants from non-core groups, including to nominate Ga aspirants over Akans in multiple constituencies in Greater Accra. Interview with NPP primary aspirant, Greater Accra Region, 26 July 2011; interview with NPP primary aspirant, Greater Accra Region, 27 July 2011.
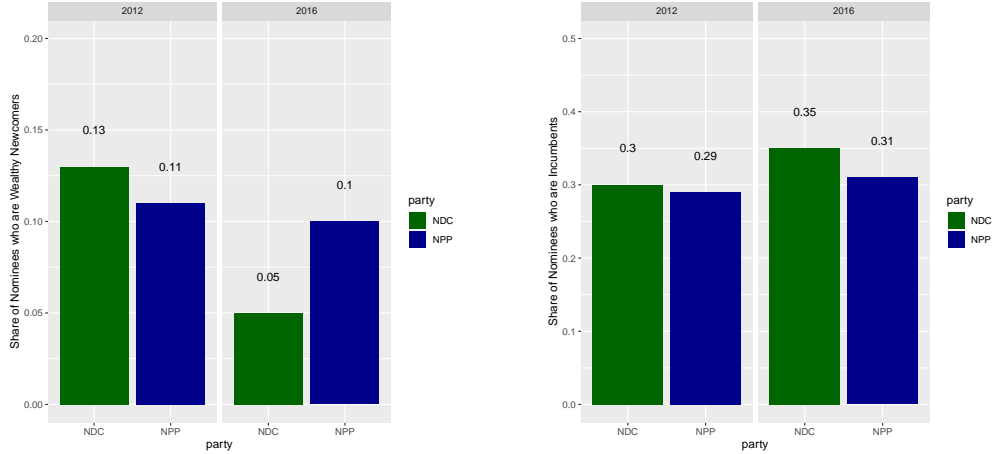
Figure 4: *Political experience of nominees, by party and year.*

profession (e.g., lawyer). Because experience in the private sector and a background in government or party leadership are not mutually exclusive, we code private sector background as 1 *only* if a nominee had not served as an MP or in other government or party positions. This identifies the wealthy political newcomers.[29]

Figure 2 shows that the NDC's electorate expansion made its nominees 11 percentage points less likely ($p < 0.001$) to be someone with a private sector background, but no prior political experience. For this outcome, we use the four common matching variables and exact match on the lagged outcome variable. Table 1 again presents balance statistics.

While the NDC's expansion of the electorate reduced the success of wealthy political newcomers in primaries, as NDC leaders intended, it did not appear to harm the party's existing MPs. For our analysis of whether incumbents were renominated, we include the population share of the incumbent's ethnic group as a matching variable, since the incumbent may be stronger, regardless of electorate size, in constituencies where he or she is from a more numerous group. Restricting the analysis to party-constituencies with incumbent MPs, we find that the expansion of the electorate led to incumbent MPs being 17 percentage points *more* likely to be renominated ($p < 0.01$).

---

[29]Missing data precludes us from analyzing the effects of the NDC's reform on the total number of aspirants with these backgrounds.

Robustness to Alternative Specifications

Our results are generally robust to alternative specifications. We address concerns about potentially poor covariate overlap by rematching using a caliper that restricts matches to party-constituencies that fall within progressively smaller ranges on the propensity score (Appendix, pg. A18). We also restrict the ratio of treated to control units to 10:1 and 1:10 to limit the influence of a single control (treated) constituency that may be matched with many treated (control) units (pg. A16). Neither affects our substantive results, although our point estimates vary somewhat as data is discarded by these procedures. In addition, we show that our results do not depend on matching exactly on the lagged dependent variable (pg. A20).

We also compare our preferred optimal full matching approach to three alternative matching methods: nearest-neighbor propensity score matching, Mahalanobis distance matching and genetic matching. We show that improvements in covariate balance are greater with optimal full matching, indicating that our preferred approach should produce less biased estimates conditional on the choice of matching variables (pg. A21).

Alternative Explanations

We proposed above that these results may result from two changes to the nature of primaries: the shift to a more diverse primary electorate – in terms of both gender and ethnicity – and a reduction in the prevalence of vote buying. While our data does not allow us to directly observe either mechanism, we are able to rule out several alternative explanations that do not operate through changes to the composition of the primary electorate or the extent of vote buying.

First, the NDC may have had more aspirants simply because nominations in the incumbent president's party are perceived to be more valuable. This is unlikely since our analysis only compares nominations with similar demand by exact matching on the number of aspirants in 2012, when the NDC was also the incumbent party and both parties used the same rules. Both parties ran the same presidential candidates in 2016 as in 2012, and the platforms and core bases of each party were unchanged. The main difference leading up to 2016 was the growing popularity of the NPP among swing voters, amid a backlash against the NDC's performance (Bob-Milliar and Paller,

2018). Any increased expectations of becoming an MP as a member of the president's party should have increased the number of aspirants in the NPP relative to the NDC. But we find the opposite.

Secondly, the NDC's electorate expansion may have created a less certain electoral environment, leading aspirants from under-represented groups who did not believe they would have more support from new voters to enter to take advantage of a more open race. However, the result that more female and non-core group aspirants won primaries suggests that these new entrants did not misjudge their increased viability. Moreover, we find similar results when we exclude aspirants with lower viability, proxied by those who declared their candidacy but dropped out before the primary (Appendix, pg. A9).

Thirdly, the fact that the NPP considered, but rejected, the full expansion of its own electorate in the run-up to the 2012 election might indicate that the it is a less democratic party than the NDC and explain why a larger and more diverse group of aspirants competed in and won the NDC's primaries. However, it was easier for the NDC to expand its electorate because the party is *less* democratic than the NPP and doing so did not require building consensus among multiple factions. The NPP national leadership had been riven by factional disputes, particularly between the two main aspirants for the party's presidential nomination in 2008, 2012, and 2016. One of those factions blocked an identical proposal to expand the NPP's primary electorate to all party members before 2012. The NDC's leadership is more centralized and unified around a powerful General Secretary (Osei, 2016), who has greater discretion to change the party rules. This greater factionalism within the NPP should bias against our finding that more aspirants competed in the NDC: interviews with NPP aspirants suggest that the party's factions sometimes put forward rival candidates in the primaries, which should increase the number of aspirants in our control group observations.

Beyond Ghana

Expanding the primary electorate increased the number of aspirants who are female or members of under-represented ethnic groups in one of Ghana's major parties. This reform also increased the probability that the party's nominees were female or members of these ethnic groups and led

to more incumbent MPs being renominated. But it decreased the probability that nominees were wealthy individuals with little political experience. Although we cannot directly test mechanisms, these results are consistent with our interpretation that expanding the electorate made vote buying a less viable strategy and changed the ethnic and gender composition of the electorate, which may have encouraged potential aspirants from groups that became better represented in the electorate to compete for nominations.

Our findings show that changes to primary rules are another institutional means – beyond quotas and reserved seats – to make candidate selection more inclusive. Although expanding the primary electorate has smaller effects than directly reserving seats, its effects are similar in magnitude to the longer-run impacts of seat reservations even after they are withdrawn (Bhavnani, 2009, 2017). It may be easier to expand the primary electorate because it does not restrict competition by men or members of the ethnic majority who may be excluded from reserved seats. By making it more difficult for political outsiders with little reputation among general election voters to buy their way to nominations, this reform also appears to improve candidate quality. This echoes other findings that institutional shifts to improve descriptive representation among candidates can have carry-over effects on candidate competence (Besley et al., 2017).

Should we expect similar short-run effects in other settings? Primaries are increasingly common across the developing world (Field and Siavelis, 2008), including in Africa (Choi, 2018; Warren, 2018), where parties in at least fifteen countries now elect their legislative nominees through primaries (Ichino and Nathan, 2018). Political parties such as the Botswana Democratic Party have also recently expanded their primary electorates in a similar manner to the NDC (Warren, 2018).

While it would not be possible to employ our empirical strategy in countries with single-party dominant or more inchoate party systems, our general theoretical answer is yes, conditional on three scope conditions that affect our proposed mechanism for how expanding the electorate affects the calculations of potential aspirants. The first condition is that primaries are real elections: the outcomes are not decided in advance through the manipulation of party elites. The second is that there are gender gaps and ethnic disparities in elite primary electorates compared to the broader party membership. The third is that both inter- and intra-party competition are patronage based

and non-ideological. While the first condition may not hold where national party leaders strongly pressure primary voters to back favored aspirants (Choi, 2018), the latter two conditions should apply widely, especially in Africa (Logan and Bratton, 2006; Riedl, 2014).

Where these conditions hold, key differences in party systems may only lead to greater effects than for Ghana. For example, in more inchoate party systems, expanding the electorate may generate higher levels of uncertainty that encourage more non-viable aspirants to compete in the primaries, adding to the effects on aspirant entry. Expanding the electorate in such settings may also increase the likelihood of coordination failures within a party's core ethnic groups, adding to the viability of the path to nominations for aspirants from non-core groups. Similarly, the effects on the number and characteristics of aspirants may be even larger where there is greater potential for party switching than in Ghana. If aspirants can easily switch to a political party that expands its electorate to include more female and ethnically non-aligned voters, we expect more potential aspirants from under-represented groups to do so. Future research examining how the strategies of male and non-minority aspirants change when faced with these new competitors and types of voters can also help refine our argument for new settings.

References

Toke S. Aidt and Bianca Dallal. Female voting power: The contribution of women's suffrage to the growth of social spending in western europe (1869-1960). *Public Choice*, 134(3/4):391–417, 2008.

Leonard R. Arriola and Martha C. Johnson. Ethnic politics and women's empowerment in africa: Ministerial appointments to executive cabinets. *American Journal of Political Science*, 58(2): 495–510, 2014.

Samuel Berlinski and Torun Dewan. The political consequences of franchise extension: Evidence from the second reform act. *Quarterly Journal of Political Science*, 6:329–376, 2011.

Samuel Berlinski, Torun Dewan, and Brenda van Coppenolle. Franchise extension and the british aristocracy. *Legislative Studies Quarterly*, 39(4):531–558, 2014.

Timothy Besley, Olle Folke, Torsten Persson, and Johanna Rickne. Gender quotas and the crisis of the mediocre man: Theory and evidence from sweden. *American Economic Review*, 107(8): 2204–2242, 2017.

Rikhil R. Bhavnani. Do electoral quotas work after they are withdrawn? evidence from a natural experiment in india. *American Political Science Review*, 103(1):23–35, 2009.

Rikhil R. Bhavnani. Do the effects of temporary ethnic group quotas persist? evidence from india. *American Economic Journal: Applied Economics*, 9(3):105–123, 2017.

Elin Bjarnegård and Pär Zetterberg. Political parties, formal selection criteria, and gendered parliamentary representation. *Party Politics*, 25(3):325–335, 2019.

George M. Bob-Milliar. Political party activism in ghana: Factors influencing the decision of the politically active to join a political party. *Democratization*, 19(4):668–689, 2012a.

George M. Bob-Milliar. Party factions and power blocs in ghana: A case study of power politics in the national democratic congress. *Journal of Modern African Studies*, 50(4):573–601, 2012b.

George M. Bob-Milliar and Jeffrey W. Paller. Democratic ruptures and electoral outcomes in africa: Ghana's 2016 election. *Africa Spectrum*, 53(1):5–35, 2018.

Sarah Brierley and Noah L. Nathan. The connections of party brokers. Forthcoming, *Journal of Politics*, 2020.

John M. Carey and John Polga-Hecimovich. Primary elections and candidate strength in latin america. *Journal of Politics*, 68(3):530–43, 2006.

Raghabendra Chattopadhyay and Esther Duflo. Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443, 2004.

Donghyun Danny Choi. Party control and primary elections in new democracies. Working Paper, 2018.

Amanda Clayton. Women's political engagement under quota-mandated female representation: Evidence from a randomized policy experiment. *Comparative Political Studies*, 48:333–69, 2015.

Amanda Clayton, Amanda Lea Robinson, Martha C. Johnson, and Ragnhild Muriaas. (how) do voters discriminate against women candidates? experimental and qualitative evidence from malawi. *Comparative Political Studies*, 53(3-4):601–630, 2020.

Eric S. Dickson and Kenneth Scheve. Social identity, electoral institutions, and the number of candidates. *British Journal of Political Science*, 40(2):349–375, 2010.

Esther Duflo. Why Political Reservations? *Journal of the European Economic Association*, 3(2-3): 668–78, 2005.

Karen Ferree. The social origins of electoral volatility in africa. *British Journal of Political Science*, 40:759–79, 2010.

Bonnie N. Field and Peter M. Siavelis. Candidate selection procedures in transitional polities. *Party Politics*, 14(5):620–639, 2008.

Richard L. Fox and Jennifer L. Lawless. Gaining and losing interest in running for public office: The concept of dynamic political ambition. *Journal of Politics*, 73(2):443–462, 2011.

Richard L. Fox and Jennifer L. Lawless. Uncovering the origins of the gender gap in political ambition. *American Political Science Review*, 108(3):499–519, 2014.

Elisabeth R. Gerber and Rebecca B. Morton. Primary Election Systems and Representation. *Journal of Law, Economics and Organization*, 14(2):304–324, 1998.

Jessica Gottlieb, Guy Grossman, and Amanda L. Robinson. Do men and women have different policy preferences in africa? determinants and implications of gender gaps in policy prioritization. *British Journal of Political Science*, 48(3):611–636, 2018.

Ben B. Hansen. Full matching in an observational study of coaching for the sat. *Journal of the American Statistical Association, Applications and Case Studies*, 99(467):609–618, 2004.

Ben B. Hansen and Jake Bowers. Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science*, 23(2):219–236, 2008.

Ben B. Hansen and Stephanie Olsen Klopfer. Optimal full matching and related designs via network flows. *Journal of Computational and Graphical Statistics*, 15(3):609–627, 2006.

Magda Hinojosa. *Selecting Women, Electing Women: Political Representation and Candidate Selection in Latin America*. Philadelphia, PA: Temple University Press, 2012.

Nahomi Ichino and Noah L. Nathan. Primaries on demand? intra-party politics and nominations in ghana. *British Journal of Political Science*, 42(4):769–791, 2012.

Nahomi Ichino and Noah L. Nathan. Crossing the line: Local ethnic geography and voting in ghana. *American Political Science Review*, 107(2):344–361, 2013a.

Nahomi Ichino and Noah L. Nathan. Do primaries improve electoral performance? clientelism and intra-party conflict in ghana. *American Journal of Political Science*, 57(2):428–441, 2013b.

Nahomi Ichino and Noah L. Nathan. Primary elections in new democracies. In Robert G. Boatright, editor, *Routledge Handbook of Primary Elections*, pages 369–383. New York: Routledge Press, 2018.

Nahomi Ichino and Noah L. Nathan. Replication data for: Democratizing the party: The effects of primary election reforms in ghana. 10.7910/DVN/FYOEWG, Harvard Dataverse, V1. UNF:6:NB3It0+tFw5bPNtx9wn91Q==, 2021.

Cecilia Josefsson. Who benefits from gender quotas? assessing the impact of election procedure reform on members of parliament's attributes in uganda. *International Political Science Review*, 35(1):93–105, 2014.

Christopher F. Karpowitz, J. Quin Monson, and Jessica Robinson Preece. How to elect more women: Gender and candidate success in a field experiment. *American Journal of Political Science*, 61(4):927–943, 2017.

Ozge Kemahlioglu, Rebecca Weitz-Shapiro, and Shigeo Hirano. Why primaries in latin american elections? *Journal of Politics*, 71(1):339–352, 2009.

Mona Lena Krook. *Quotas for Women in Politics: Gender and Candidate Selection Reform Worldwide.* Oxford University Press, New York, 2009.

Staffan I. Lindberg. What accountability pressures do mps in africa face and how do they respond? evidence from ghana. *Journal of Modern African Studies*, 48(1):117–142, 2010.

Alessandro Lizzeri and Nicola Persico. Why did the elites extend the suffrage? democracy and the scope of government, with an application to britain's 'age of reform'. *The Quartely Journal of Economics*, 119(2):707–765, 2004.

Carolyn Logan and Michael Bratton. The political gender gap in africa: Similar attitudes, different behaviors. Afrobarometer Working Paper No. 58, 2006.

Jr. Lott, John R. and Lawrence W. Kenny. Did women's suffrage change the size and scope of government? *Journal of Political Economy*, 107(6):1163–1198, 1999.

Mona Morgan-Collins and Dawn L. Teele. Revisiting the gender voting gap in the era of women's suffrage. Working Paper, 2017.

Noah L. Nathan. *Electoral Politics and Africa's Urban Transition: Class and Ethnicity in Ghana.* New York: Cambridge University Press, 2019.

Pippa Norris and Joni Lovenduski. *Political Recruitment: Gender, Race, and Class in the British Parliament.* New York: Cambridge University Press, 1995.

Anja Osei. Formal party organisation and informal relations in african parties: evidence from ghana. *Journal of Modern African Studies*, 54(1):37–66, 2016.

Rohini Pande. Can mandated political representation increase policy influence for disadvantaged minorities? theory and evidence from india. *American Economic Review*, 93(4):1132–1151, 2003.

Melanie L. Phillips. Family matters: Gendered candidate selection by party gatekeepers. Working Paper, 2020.

Amanda Pinkston. *Insider Democracy: Poverty and the Closed Political Class in Democratic Africa*. PhD thesis, Harvard University, 2016.

Rachel Beatty Riedl. *Authoritarian Origins of Democratic Party Systems in Africa*. New York: Cambridge University Press, 2014.

Paul R. Rosenbaum. A characterization of optimal designs for observational studies. *Journal of the Royal Statistical Society*, 53:597–610, 1991.

Paul R. Rosenbaum. *Design of Observational Studies*. Springer, 2010.

Kira Sanbonmatsu. Political parties and the recruitment of women to state legislatures. *Journal of Politics*, 64(4):791–809, 2002.

Susanne Schwarz and Alexander Coppock. What have we learned about gender from candidate choice experiments? a meta-analysis of 67 factorial survey experiments. Forthcoming, *Journal of Politics*, 2020.

Gilles Serra. Why primaries? the party's tradeoff between policy and valence. *Journal of Theoretical Politics*, 23(1):21–51, 2011.

Susan C. Stokes, Thad Dunning, Marcelo Nazareno, and Valeria Brusco. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. New York: Cambridge University Press, 2013.

Aili Mari Tripp and Alice Kang. The global impact of quotas: On the fast track to increased female legislative representation. *Comparative Political Studies*, 41(3):338–61, 2008.

Lena Wangnerud. Women in parliaments: Descriptive and substantive representation. *Annual Review of Political Science*, 12:51–69, 2009.

Shana Warren. Democratizing candidate selection: Evidence from botswana's bulela ditswe primaries. Working Paper, 2018.

Appendix:
Democratizing the Party: The Effects of Primary Election Reforms in Ghana
Nahomi Ichino and Noah L. Nathan

The replication data and code are available at https://doi.org/10.7910/DVN/FYOEWG.

Contents

A   Data

A.1   Qualitative Interview Data

Qualitative information on the conduct of Ghanaian primaries is drawn from 155 interviews conducted by, or on behalf of, the authors with primary contestants in both major parties between 2010 and 2016. The most formal qualitative data collection involved an in-depth survey of 125 NPP primary aspirants competing in the party's 2015 primaries in advance of the 2016 elections. Because the NPP refused to provide us the contact information for its full slate of parliamentary aspirants, the survey sample is non-random and includes all aspirants in the NPP's 2011 primaries who competed again in 2015. The NDC refused to make similar contact information available, making it impossible to draw a similar sample of aspirants in the NDC's primaries before the 2016 elections. Ultimately, 125 of the 213 contacted NPP aspirants agreed to interviews. Almost all of the remainder were aspirants whose phone numbers had gone out of service since 2011 and could not be reached. The 2015 interviews were conducted by Ghanaian research assistants either over the phone or in person. These interviews were audio recorded and research assistants also entered responses into a questionnaire. Data gathered in our interviews is used to supplement the coding of biographical details from media sources for each of these aspirants.

A.2   Summary Statistics

Table A1 presents summary statistics for all variables used in the main specification and the alternative models.

Table A1: *Summary Statistics*

	n	Mean	SD	Min	Max
NDC:					
Total number of 2016 aspirants	219	3.52	1.65	1	9
Number of 2016 female aspirants	219	0.31	0.55	0	3
Number of 2016 non-core female aspirants	219	0.21	0.45	0	2
Num. 2016 aspirants from party's core ethnic groups	195	1.24	1.68	0	7
Num. 2016 aspirants from non-core ethnic groups	195	2.10	1.44	0	7
2016 nominee has only private sector background	219	0.23	0.42	0	1
2016 nominee is the incumbent	219	0.35	0.48	0	1
2016 nominee is female	219	0.14	0.35	0	1
2016 nominee belongs to party's core ethnic group	217	0.41	0.49	0	1
2016 nominee is non-core and female	216	0.09	0.29	0	1
Total number of 2012 aspirants	223	2.65	1.66	1	9
Number of 2012 female aspirants	223	0.29	0.58	0	3
Number of 2012 non-core female aspirants	220	0.18	0.47	0	3
Num. 2012 aspirants from party's core ethnic groups	211	1.28	1.54	0	7
Num. 2012 aspirants from non-core ethnic groups	211	1.29	1.22	0	6
2012 nominee has only private sector background	275	0.26	0.44	0	1
2012 nominee is the incumbent	275	0.30	0.46	0	1
2012 nominee is female	275	0.09	0.29	0	1
2012 nominee belongs to party's core ethnic group	217	0.41	0.49	0	1

Continued on next page

	n	Mean	SD	Min	Max
2012 nominee is non-core and female	216	0.07	0.26	0	1
2016 incumbent's ethnic group share	104	0.55	0.34	0.003	0.97
Vote share in 2012 parliamentary election	269	0.48	0.16	0.13	0.92
Vote share in 2012 presidential election	268	0.53	0.19	0.14	0.96
Fractionalization of party's core ethnic groups	264	0.62	0.26	0.01	0.93
Segregation of party's core groups from other groups	264	0.18	0.11	0.00	0.56
Segregation among party's core groups	263	0.41	0.15	0.02	0.89
NPP:					
Total number of 2016 aspirants	252	2.87	1.57	1	8
Number of 2016 female aspirants	252	0.20	0.46	0	2
Number of 2016 non-core female aspirants	252	0.10	0.34	0.00	2.00
Num. 2016 aspirants from party's core ethnic groups	240	1.66	1.53	0	7
Num. 2016 aspirants from non-core ethnic groups	240	1.15	1.38	0	7
2016 nominee has only private sector background	252	0.32	0.47	0	1
2016 nominee is the incumbent	252	0.31	0.46	0	1
2016 nominee is female	252	0.08	0.27	0	1
2016 nominee belongs to party's core ethnic group	248	0.61	0.49	0	1
2016 nominee is non-core and female	245	0.02	0.13	0	1
Total number of 2012 aspirants	234	2.56	1.41	1	8
Number of 2012 female aspirants	234	0.34	0.59	0	3
Number of 2012 non-core female aspirants	230	0.18	0.42	0.00	2.00
Num. 2012 aspirants from party's core ethnic groups	222	1.59	1.49	0	8
Num. 2012 aspirants from non-core ethnic groups	222	0.94	1.18	0	6
2012 nominee has only private sector background	275	0.30	0.46	0	1
2012 nominee is the incumbent	275	0.29	0.45	0	1
2012 nominee is female	275	0.13	0.33	0	1
2012 nominee belongs to party's core ethnic group	271	0.61	0.49	0	1
2012 nominee is non-core and female	245	0.05	0.22	0	1
2016 Incumbent's ethnic group share	103	0.62	0.23	0.01	0.97
Vote share in 2012 parliamentary election	269	0.45	0.17	0.04	0.85
Vote share in 2012 presidential election	268	0.46	0.19	0.03	0.96
Fractionalization of party's core ethnic groups	263	0.51	0.27	0.04	0.89
Segregation of party's core groups from other groups	263	0.20	0.12	0.00	0.74
Segregation among party's core groups	262	0.40	0.23	0.01	1.00
Constituency:					
Pop density of constituency (log(1000s/sqkm))	264	1.85	5.52	0.01	49.86
Pop share of largest ethnic group in constituency	264	0.73	0.18	0.31	0.97
Muslim population share in constituency	264	0.17	0.21	0.01	0.98

A.3 News Media Sources

We use information from the following news sources in our coding: *The Ghanaian Chronicle* (independent newspaper), *The Daily Guide* (independent newspaper), *The Daily Graphic* (state-owned newspaper), *The Ghanaian Times* (state-owned newspaper), *Citi FM* (independent radio station), *Peace FM* (independent radio station), and *My Joy* (independent radio station), as well

as `modernghana.com`, `vibeghana.com`, and `ghanaweb.com`, which are independently owned news aggregation websites. These sources span non-partisan outlets and those more aligned with each party.

A.4   Missing Party-Constituencies

We drop from our analyses party-constituencies where we do not feel confident that our media sources provide a full accounting of all primary aspirants for the 2016 elections. Out of 550 possible party-constituencies (275 constituencies * 2 parties), we drop 79, leaving 471 party-constituencies (219 NDC constituencies; 252 NPP constituencies) for our analysis. Table A2 presents p-values from simple t-tests for the difference-in-means between the missing and included party-constituencies on key covariates, including several measures of primary characteristics before the 2012 elections.

There are differences on some covariates. On average, the missing party-constituencies had slightly more aspirants competing in their 2012 primaries, had fewer incumbents re-nominated in 2012, were slightly more likely to have supported their party in the 2012 presidential (but not parliamentary) election, and have smaller Muslim population shares. But the missing and included party-constituencies are similar on all of the other covariates.

Table A2: *Comparing missing to non-missing party-constituencies.*

Variable	Differences in means (missing – non-missing)	p-value
Total number of aspirants, 2012	0.57	0.03
Number of female aspirants, 2012	0.07	0.39
Num. aspirants from non-core groups, 2012	0.06	0.74
Female nominee, 2012	-0.03	0.46
Incumbent was nominee, 2012	-0.12	0.01
Private-sector only nominee, 2012	0.04	0.39
2012 parliamentary vote share	0.03	0.14
2012 presidential vote share	0.05	0.03
Ethnic fractionalization of core groups	0.00	0.92
Muslim % (constituency)	-0.05	0.01
Population density (constituency)	1.10	0.16

A.5   Coding Aspirant Ethnicity

We code the ethnicity of each aspirant based on their names, which are generally easily connected to the main ethnic categories in Ghana. We assign ethnicity based on a dictionary of 3,503 names of Ghanaian politicians, comprising aspirants in the 2011-2012 NDC and NPP primaries as well as all candidates in the 2010 district assembly (city council) elections in Greater Accra Region, which as Ghana's largest urban area has numerous candidates from all major ethnic groups.

Each name in the dictionary was coded in triplicate into ethnic categories by a team of five university-student research assistants in Accra who come from different regions of the country and ethnic groups. Anglophone name fragments that lack ethnic content (e.g., "John") were removed, but Anglophone surnames were left in the dictionary, since these frequently indicate Fanti ethnicity. The dictionary was then matched to the list of aspirants, and each aspirant was assigned the

ethnicity of the majority coding of matches to her name. This allows us to identify the overall ethnic category for 87% of the 2015 primary aspirants and over 90% of the 2011-2012 aspirants. Importantly, this method can easily distinguish Ghanaian names among 7 broad categories – Akan (excluding Fanti), Fanti, Ga-Dangme, Ewe, Guan, and Northern – but cannot reliably distinguish ethnic sub-groups within these categories. For example, a name such as "Kwame Owusu" is clearly identifiable as Akan, but could not be consistently coded among Akan subgroups. For Northern names, there are too many small Northern ethnic groups with their own naming conventions for the research assistants to be able to systematically tell them apart.

A.6 Ethnic Diversity by Population Density

Figure A1 plots two measures of constituency-level ethnic diversity – overall ethnic fractionalization (left panel) and ethnic fractionalization among core ethnic groups of each party (right panel) – against constituency population density (logged 1000s/sq km). Population density is moderately correlated with both diversity measures, with more diversity in denser (i.e., more urban) constituencies: $r = 0.25$ for the correlation of overall ethnic fractionalization and population density, while $r = 0.16$ for the correlation of intra-party fractionalization and population density. But Figure A1 shows that, using either measure, there are still many rural constituencies with low population density but significant ethnic diversity.
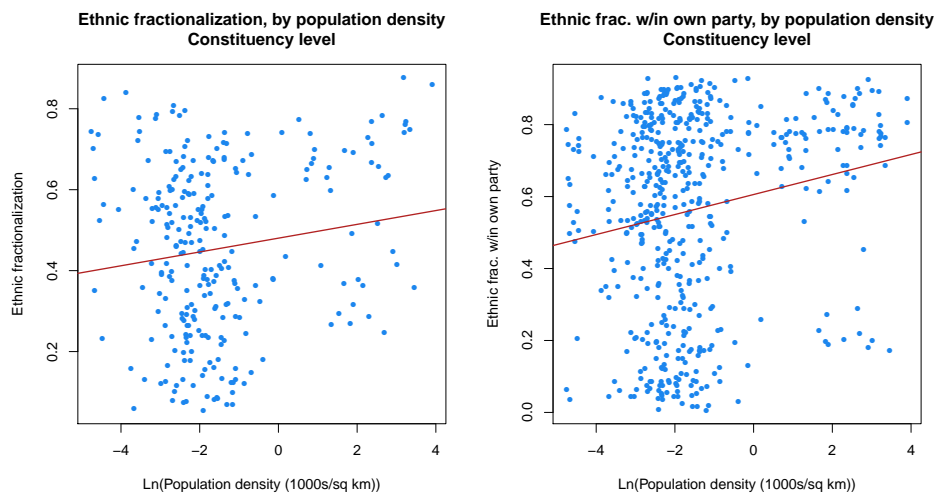

Figure A1: *Ethnic diversity measures by logged constituency population density.*

B Further Details on the Main Specification

B.1 Aspirants and Nominees who are Female and from Non-Core Groups

For the number of female aspirants from non-core groups, we match on a propensity score calculated from the matching variables used for the number of female aspirants and the matching variables used for the number of aspirants from non-core groups, as well as the lagged dependent variable. We estimate that the average effect on the number of female aspirants from non-core groups is 0.12 (p <0.01). For whether the nominee is female and also from a non-core group, we match a propensity score calculated from the matching variables used for the whether the nominee is female and the matching variables used for whether the nominee is from a non-core group, as well as the lagged dependent variable. The estimated effect is an 8 percentage-point increase in the probability the NDC nominee was female and from a non-core group ($p < 0.01$). Figure B1 shows the balance on individual matching variables. Although the sample sizes become small, Table B1 shows that these effects are not concentrated in constituencies where the party has little chance of winning.
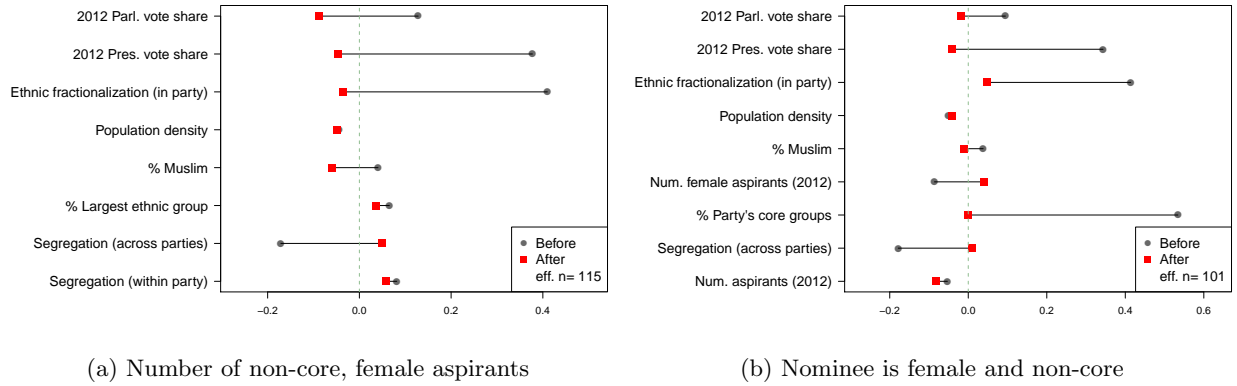

(a) Number of non-core, female aspirants

(b) Nominee is female and non-core

Figure B1: *Balance on individual matching variables for intersectional outcomes:* the x-axis displays standardized differences in means on each variable.

Table B1: *Heterogeneous Effects for Intersectional Outcomes by Constituency Competitiveness.*

Outcome	Subset	Estimate	S.E.	p-value	n_T	Sets
Num. Non-Core, Female	Full data	0.17	0.04	< 0.01	168	84
Aspirants	Non-competitive	0.09	0.04	0.05	57	35
	Competitive	0.08	0.05	0.10	61	31
	Stronghold	0.08	0.07	0.29	49	14
Nominee is Female & Non-Core	Full data	0.08	0.02	< 0.01	169	76
	Non-competitive	0.09	0.03	< 0.01	57	25
	Competitive	0.06	0.03	0.04	63	31
	Stronghold	0.02	0.02	0.32	44	14

A7

B.2　Matching Variables

All models match exactly on the 2012 value of each outcome variable and an outcome-specific propensity score. We are not able to match party-constituencies on pre-2012 trends, for example dating back to the 2000 or 2004 elections. Unfortunately, there is too much missingness in our data for the pre-2012 primaries to implement such a model without restricting analyses to a very small and unrepresentative set of observations. We would also lose additional observations because new constituency boundaries were introduced between the 2008 and 2012 elections and we cannot define 2008 and 2004 primary outcomes for constituencies that did not previously exist.

The following four variables are always included in the calculation of the propensity score: the party's vote share in the 2012 presidential election in the constituency; the party's vote share in the 2012 parliamentary election in the constituency; the constituency population density; and ethnic fractionalization among the aligned ethnic groups of the party in the constituency.

As noted in the main text, for the total number of aspirants from core or non-core ethnic groups the set of matching variables also includes the size of the largest ethnic group in the constituency, segregation between ethnic groups associated with the NDC and the NPP, and segregation among the sub-groups within each party's national ethnic coalition. The first measure of segregation compares the geographic segregation of Northerners, Ewes, and Ga-Dangmes together from Akans (excluding the Fanti). The second measure of segregation captures the spatial distribution of the sub-groups among the Northerners, Ewes, and Ga-Dangmes for NDC primaries and the spatial distribution of the sub-groups within the Akan for NPP primaries.

We include the segregation measures because as primary electorates expand and competition shifts towards promises about the delivery of local public goods and away from vote buying, segregation can affect primary voters' beliefs about which types of candidates are most likely to target them with local public goods. This in turn could affect whether additional aspirants from new ethnic groups enter the primaries looking to better represent their co-ethnics' interests by bringing local public goods to their areas in a constituency. Voters often expect a politician to favor his or her own ethnic communities in the delivery of these goods. This will still benefit voters from other groups living in the same area when ethnic groups are residentially integrated because the benefits of local public goods are non-excludable within local communities (Ejdemyr et al. 2018). But where there is greater ethnic segregation, local public goods targeted to a politician's own group will not benefit other ethnic groups, increasing incentives for these other groups to seek their own co-ethnic MP who will instead target them with benefits. This dynamic accounts for significant variation in vote choice in Ghanaian general elections (Ichino and Nathan 2013). Ethnic segregation within each party's coalition, as well as ethnic segregation between each party's core groups and other ethnic groups in the constituency, may thus affect the extent to which primary competition is polarized along ethnic lines, altering the incentives of politicians from these groups to come forward as aspirants.

We measure segregation using Theil's spatial information theory index H, which is equal to 0 at complete integration and the even spatial distribution of ethnic groups, and to 1 at complete segregation (Reardon and O'Sullivan 2004). This is also known as the multigroup entropy index, and it is the weighted average deviation of each enumeration area's entropy from the constituency entropy. For a constituency with G ethnic groups, each with population share π_g, the entropy of the constituency is $E = \sum_{g=1}^{G}(\pi_g)ln(\frac{1}{\pi_g})$. The entropy of an enumeration area k is $E_k = \sum_{g=1}^{G}(\pi_{gk})ln(\frac{1}{\pi_{gk}})$, where π_{gk} is ethnic group g's population share in enumeration area k. H

for a constituency can then be expressed as $\sum_{k=1}^{K} \frac{n_k(E-E_k)}{nE}$, where n_k is the total population of the enumeration area k and $n = \sum_{k=1}^{K} n_k$ is the total population of the constituency with K enumeration areas. We use seg v0.5-1 in R v3.3.1 to calculate H (Hong 2014).

For whether the nominee is a core ethnic group member we add: the constituency population share of the ethnic groups in the party's core coalition, the segregation of these ethnic groups associated with the party from all other ethnic groups in the constituency, and the number of aspirants from these party-associated groups who competed in the 2012 primary. Including these three additional variables helps account for potential heterogeneity in the effects of the NDC's reforms on the types of nominees who win primaries. Aspirants from ethnic groups outside of the party's core ethnic base may find it more difficult to win a primary in constituencies where the party's core ethnic base makes up a larger proportion of the population. Aspirants from non-associated groups may also have a more difficult time winning primaries in segregated constituencies where voters from the party's core groups may be less likely to believe they will benefit from local public goods promised by non-coethnics. Finally, the baseline number of aspirants from each ethnic group can affect the probability a candidate from that group wins under the new rules, since it affects whether an additional aspirant from a particular ethnic group would split the group's votes.

B.3 Matched Sets from Optimal Full Matching

We prefer optimal matching to nearest-available (greedy) matching algorithms that may be more familiar to political scientists. With the latter, the ordering of the units matters since a control unit that is matched to a treated unit becomes unavailable for matching to another treated unit later on. This can be particularly consequential when the number of available control units is limited as in our study.

Figure B2 shows sets of treatment and control units created by our main specification with optimal full matching for the total number of aspirants; for this example, we show only those sets for units that had 3 aspirants in the 2012 primaries. Treated (NDC) units are in dark green and control (NPP) units are in blue, with the weight of each unit represented by the area of its circle. The units are sorted by their propensity score, which does not include the total number of aspirants in the previous election (on which we exact match). Note that all treated units (green) are weighted the same. Line segments join units of one treatment status to units of the other treatment status within the same matched set.

B.4 Excluding Aspirants Who Drop Out or are Disqualified

Our main analysis counts as aspirants any primary contestants who file nomination forms to compete in the primary. But for each party and each election year, some of these aspirants drop out before the primary actually occurs, while others are formally disqualified by party leaders during the vetting process, as described in Ichino and Nathan (2012). There may be concern that our estimated effects on the entry of new types of aspirants (i.e., women, non-core ethnic groups) are primarily driven by "hopeless" aspirants who initially decide to enter the primary because of the uncertainty created by the NDC's new rules, but then drop out before the primary election date as they realize they are not viable candidates.
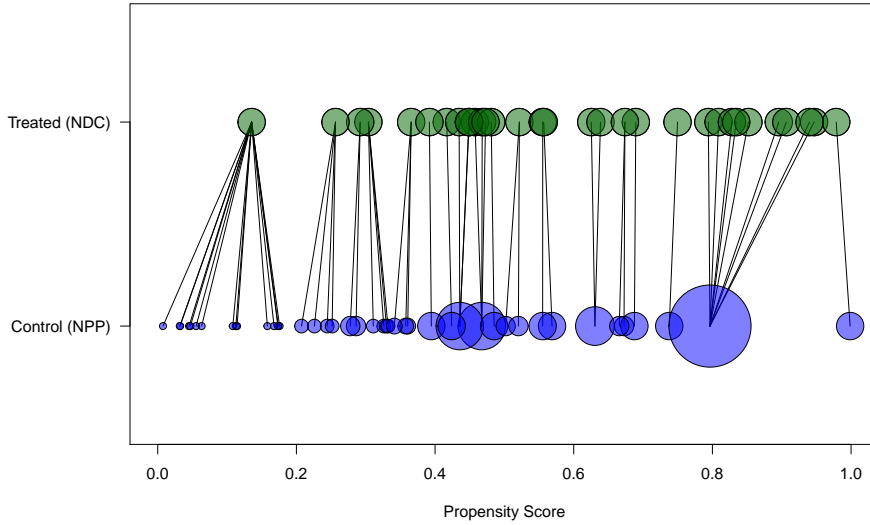
Figure B2: *Matched sets for total number of aspirants among units with 3 aspirants competing in the 2012 primaries.* The area of each circle represents the weight assigned to each unit.

We believe this alternative explanation is unlikely for two reasons. First, it cannot account for our result that significantly more women and aspirants from non-core ethnic groups go on to win nominations due to the NDC's reforms. This strongly suggests that many of these new aspirants were viable contestants. Second, our coding of the media reports allows us to identify which specific aspirants dropped out or were disqualified through the vetting process before the actual primary was held. At least one initial aspirant dropped out before the primary in 131 NDC primaries and 39 NPP primaries leading into the 2016 elections. In Figure B3, we re-do our main analyses for the number of aspirants from non-core groups, the number of female aspirants, and the total number of aspirants after excluding all aspirants who dropped out or were disqualified. We show our original estimates from Figure 1 in the main text for comparison (in blue). Our point estimates remain very similar, and our main findings that the NDC's reforms increased the entry of female aspirants and of aspirants from non-core ethnic groups are still statistically significant at conventional levels. Even after adjusting for all drop outs, Figure B3 confirms that the NDC's reforms led to new types of aspirants competing in its primaries.

B.5   Heterogeneous Effects by Constituency-Level Muslim Population Share

Table B2 reports estimates of the effect of the reforms on whether the nominee is female in constituencies that have below and above median Muslim population shares (9.5%).

Table B2: *Estimated ATT on Whether Nominee is Female.*

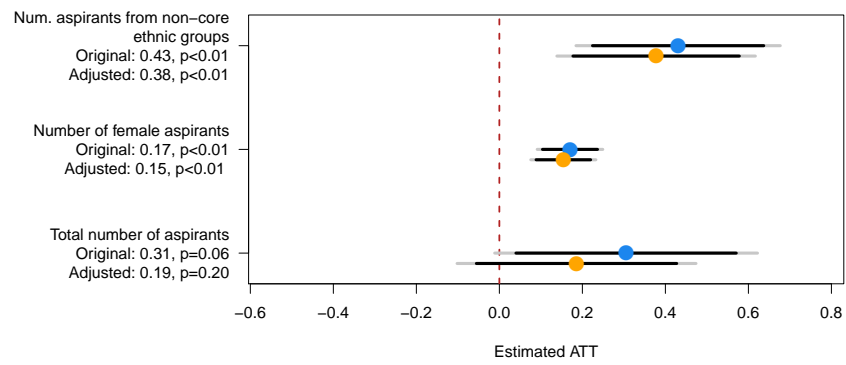|	Estimate	S.E.	p-value	n_T	Sets
Nominee is Female (Low Muslim)	0.14	0.04	< 0.01	78	54
Nominee is Female (High Muslim)	-0.03	0.03	0.32	92	36

A10

Figure B3: *Results for number of aspirants, removing aspirants who dropped out or were disqualified before the primary (adjusted estimate; orange). Original estimates in blue for comparison.*
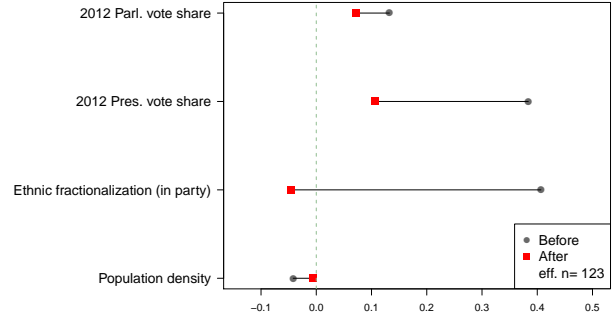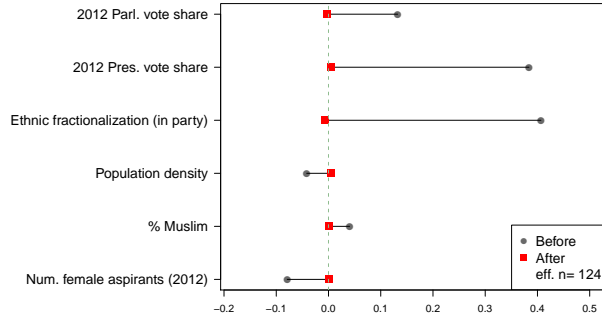
B.6 Balance on Individual Matching Variables

Figures B4 - B6 present standardized mean differences on individual matching variables before and after optimal full matching, using `xBalance` in the package `RItools` version 0.1-15 in `R` version 3.6.2 (Bowers et al. 2016). Before matching, this is the difference in means between treatment and control divided by the pooled standard deviation for each covariate. After matching, the within-set difference in means is weighted in proportion to the harmonic mean of the number of treated units and control units in the set (Hansen and Bowers 2008).
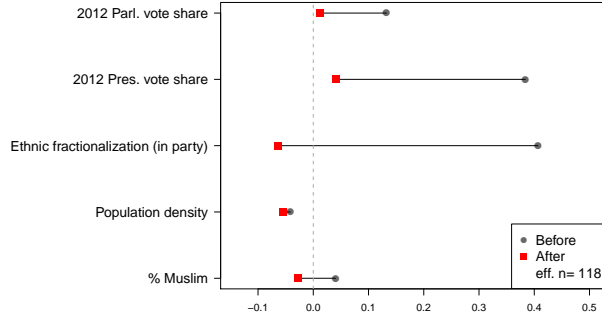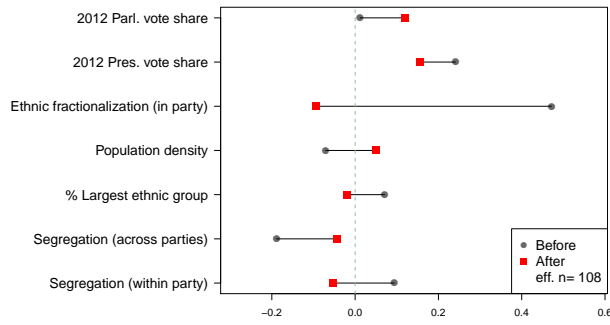

(a) Number of female aspirants
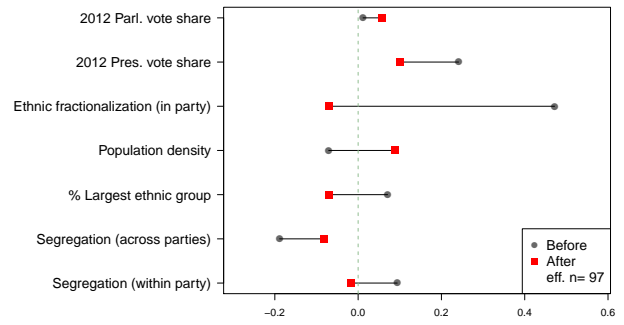
(b) Total number of aspirants

(c) Nominee is female

Figure B4: *Balance on individual matching variables for gender-related outcomes and total number of aspirants:* the *x*-axis displays standardized differences in means on each variable.

(a) Number of aspirants from non-core groups

(b) Number of aspirants from core groups
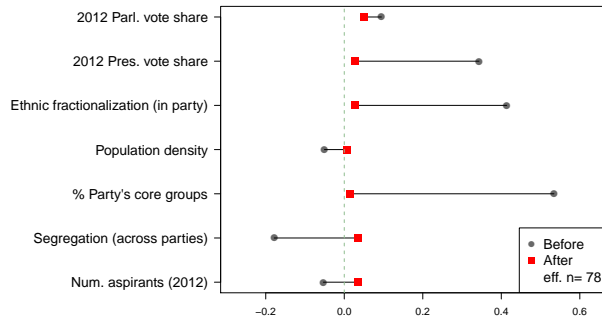
(c) Nominee is from a core group

Figure B5: *Balance on individual matching variables for ethnicity-related outcomes:* the *x*-axis displays standardized differences in means on each variable.
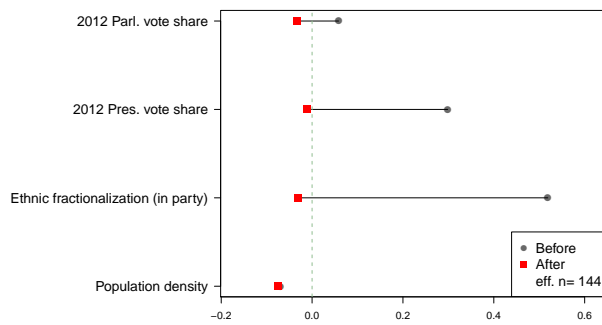
(a) Nominee has only a private sector background
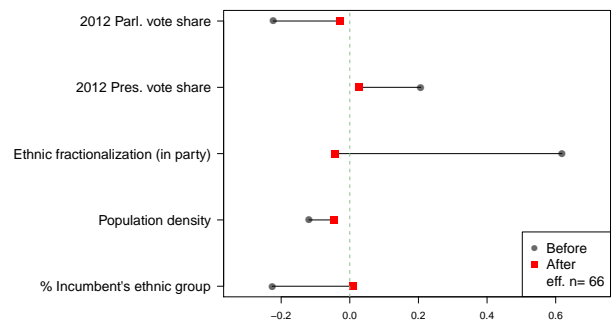
(b) Nominee is the incumbent

Figure B6: *Balance on individual matching variables for outcomes related to political experience:* the *x*-axis displays standardized differences in means on each variable.

A13

B.7 Heterogeneous Effects by Constituency Competitiveness

In Table B3 we report heterogeneous effects by three levels of electoral competitiveness: non-competitive constituencies in which the party received less than 45% in the 2012 parliamentary election; competitive constituencies where the party received between 45% and 55%; and stronghold constituencies where the party received more than 55%. We rematch for each outcome within each strata of competitiveness. We report estimates from Figure 1 in the main text using the full data for comparison. The last two columns give the number of treated units remaining and the number of sets created by optimal full matching. Table B3 shows that our main results are not concentrated exclusively in "hopeless" (i.e., non-competitive) seats for each party.

Table B3: *Heterogeneous Effects by Constituency Competitiveness.*

Outcome	Subset	Estimate	S.E.	p-value	n_T	Sets
Num. Female Aspirants	Full data	0.17	0.04	< 0.01	168	89
	Non-competitive	0.20	0.06	< 0.01	57	36
	Competitive	0.12	0.07	0.09	62	30
	Stronghold	0.18	0.09	0.05	48	18
Total Number of Aspirants	Full data	0.31	0.16	0.06	170	95
	Non-competitive	0.71	0.24	< 0.01	57	28
	Competitive	0.14	0.24	0.56	63	33
	Stronghold	0.61	0.32	0.06	45	20
Nominee is Female	Full data	0.08	0.02	< 0.01	170	97
	Non-competitive	0.09	0.04	0.01	57	35
	Competitive	0.02	0.03	0.49	63	32
	Stronghold	0.11	0.05	0.04	50	20
Number of Aspirants from Non-Core Ethnic Groups	Full data	0.43	0.12	< 0.01	150	83
	Non-competitive	0.17	0.21	0.42	52	32
	Competitive	0.86	0.18	< 0.01	56	28
	Stronghold	1.37	0.26	< 0.01	28	10
Number of Aspirants from from Core Ethnic Groups	Full data	-0.21	0.12	0.09	148	77
	Non-competitive	0.15	0.17	0.39	52	28
	Competitive	0.18	0.21	0.41	56	24
	Stronghold	0.18	0.28	0.52	40	15
Nominee is a Core Ethnic Group Member	Full data	-0.22	0.03	< 0.01	170	54
	Non-competitive	-0.05	0.05	0.35	57	19
	Competitive	-0.12	0.06	0.05	63	24
	Stronghold	0.02	0.05	0.70	50	8
Nominee has Private Sector Background	Full data	-0.11	0.03	< 0.01	211	114
	Non-competitive	-0.21	0.04	< 0.01	81	45
	Competitive	-0.00	0.04	0.93	73	37
	Stronghold	-0.00	0.02	1.00	57	28
Nominee is the Incumbent	Full data	0.17	0.06	< 0.01	102	53
	Competitive	0.12	0.09	0.21	46	21
	Stronghold	0.22	0.08	0.01	51	24

C   Alternative Specifications and Estimation Procedures

C.1   Differences-in-Differences

The standard linear differences-in-differences (DD) model is an alternative approach to estimating the ATT. The DD model and the lagged dependent variable model make different identifying assumptions, and the choice of models depends on whether one believes the most important omitted variables are time invariant or not. The DD is a fixed effects approach that differences out unspecified unmeasured time-invariant variables. Without weighting using covariates, this is simply the difference between the average of the changes in the outcomes for the NDC primaries from 2012 to 2016 and the average of the change in the outcomes for the NPP primaries from 2012 to 2016. This is also equivalent to pairing the NDC to the NPP in *each* constituency and calculating the differences in outcomes, since the difference of averages is the same as the average of differences. It relies on a "parallel trends" assumption, which we are unable to check because we only have data for two time periods (2016 and 2012) due to redistricting ahead of the 2012 elections.

By contrast, our preferred approach emphasizes what we think is a more important time-varying omitted variable – the normalization or expectation of the feasibility of success of aspirants from under-represented groups – through a proxy of the 2012 (i.e., lagged) outcome. Having female aspirants or nominee in the previous primary demonstrates that a path to political office may be open for women in that constituency. The same argument applies to having aspirants or a nominee from a non-core ethnic group. Our approach reflects the belief that this normalization is an important part of the model for the counterfactual. We would implement a lagged dependent variable model if data for additional elections were to become available, as potential aspirants would adjust their expectations as outcomes are realized. However, the models in this analysis are very similar because we only have two time periods and all our covariates are measured only at one point in time (for example, from the 2010 census).

Table C1: *Estimated ATT with Differences-in-Differences.*

	Estimate	S.E.	*p*-value
Number of female aspirants	0.17	0.07	0.02
Total number of aspirants	0.57	0.20	0.01
Nominee is female	0.08	0.04	0.04
Num. aspirants from non-core ethnic groups	0.45	0.17	0.01
Num. aspirants from party's core ethnic groups	0.09	0.18	0.62
Nominee belongs to party's core ethnic group	0.03	0.05	0.61
Nominee has only private sector background	-0.07	0.04	0.07
Nominee is the incumbent	0.04	0.09	0.68

We can estimate the ATT in the DD framework by regressing the outcome on an indicator for NDC, year, and their interaction, and our matching variables. Table C1 reports the coefficients on the interaction term, the estimates of the ATT for each outcome, from regressions that include covariates. Six of the 8 results are signed in the same direction as the results in our preferred analysis; the two that are signed in the opposite direction have *p*-values greater than 0.6. The DD-estimated effects are very similar to our main specification for (a) the number of aspirants from non-core ethnic groups, (b) number of female aspirants, (c) whether the nominee is female,

and (d) whether the nominee has only a private sector background. The DD-estimated effect is much larger than our full optimal matching-estimated effect for the (e) total number of aspirants.

C.2 Ordinary Least Squares

Table C2 reports the coefficients on the NDC variable from OLS regressions. For each outcome, the matching variables and the lagged outcome are included as controls.

Table C2: *Estimated ATT with OLS.*

	Estimate	S.E.	p-value
Number of female aspirants	0.13	0.05	0.02
Total number of aspirants	0.49	0.19	0.01
Nominee is female	0.08	0.03	0.01
Num. aspirants from non-core ethnic groups	0.68	0.15	< 0.01
Num. aspirants from party's core ethnic groups	-0.14	0.16	0.39
Nominee belongs to party's core ethnic group	-0.17	0.04	< 0.01
Nominee has only private sector background	-0.11	0.03	< 0.01
Nominee is the incumbent	0.04	0.07	0.61

C.3 Restricting the Number of Treated or Control Units in Each Set

As a robustness check, we restrict matched sets to have at most 1 treated to 10 control or 10 treated to 1 control units (Tables C3 and C4). Note that the left side of Table C3 ("Before matching") simply replicates the left side of Table 1 from the main text. The balance is not quite as good as in our main specification, but the results are substantively similar.

Table C3: *Balance before and after full optimal matching with restrictions on the number of treated or control units in each set.*

	Before matching			After matching		
	χ^2	df	p-value	χ^2	df	p-value
Number of female aspirants	92.13	6	< 0.01	5.27	5	0.38
Total number of aspirants	94.31	5	< 0.01	8.88	4	0.06
Nominee is female	92.51	7	< 0.01	2.22	6	0.90
Num. aspirants from non-core ethnic groups	79.51	8	< 0.01	6.47	7	0.49
Num. aspirants from party's core ethnic groups	92.10	8	< 0.01	6.91	7	0.44
Nominee belongs to party's core ethnic group	168.19	8	< 0.01	21.54	7	< 0.01
Nominee has only private sector background	116.41	5	< 0.01	4.08	4	0.39
Nominee is the incumbent	57.78	6	< 0.01	4.96	5	0.42

We can further show that there are significant advantages to allowing matched sets to have variable numbers of treated and control units, with some sets having many constituencies from one

Table C4: *Estimated ATT, with restrictions on the number of treated or control units in each matched set.*

	Estimate	S.E.	p-value	n_T	Sets
Number of female aspirants	0.17	0.04	< 0.01	168	87
Total number of aspirants	0.27	0.16	0.09	170	95
Nominee is female	0.08	0.02	< 0.01	170	97
Num. aspirants from non-core ethnic groups	0.46	0.12	< 0.01	150	83
Num. aspirants from party's core ethnic groups	-0.20	0.12	0.11	148	73
Nominee belongs to party's core ethnic group	-0.28	0.04	< 0.01	169	60
Nominee has only private sector background	-0.09	0.02	< 0.01	211	114
Nominee is the incumbent	0.10	0.06	0.08	102	53

party being matched to just one constituency from the other party. If we conduct optimal matching with the restriction that only one control unit be matched to each treated unit (pair matching), significant imbalances remain (Table C5).

Table C5: *Imbalance remains after optimal pair matching.*

	χ^2	df	p-value	n_T	Sets
Number of female aspirants	70.67	5	< 0.01	168	168
Total number of aspirants	75.13	4	< 0.01	168	168
Nominee is female	77.98	6	< 0.01	170	170
Num. aspirants from non-core ethnic groups	51.68	7	< 0.01	137	137
Num. aspirants from party's core ethnic groups	49.85	7	< 0.01	119	119
Nominee belongs to party's core ethnic group	91.95	7	< 0.01	152	152
Nominee has only private sector background	104.64	4	< 0.01	210	210
Nominee is the incumbent	46.71	5	< 0.01	95	95

C.4 Matching with Calipers

We set caliper restrictions and re-estimate our models. The caliper restricts matched sets to party-constituencies falling within c standard deviations of each other on the propensity score. In Figure C1, we vary c from 0.1 to 1.5 standard deviations and present the estimated effects with their 90% and 95% confidence intervals. For comparison, our original effect estimates are represented by the dashed blue horizontal line in each panel. The red diamonds indicate the number of treated units remaining after optimal full matching with calipers. Larger numbers of treated units with no available matches from the control group are discarded when the calipers are set to be very small.

There are small variations in estimated effect sizes depending on the caliper size. But overall, the main results for most outcomes hold consistently across most caliper sizes, suggesting that these results are not due to extrapolation or interpolation from poor overlap.
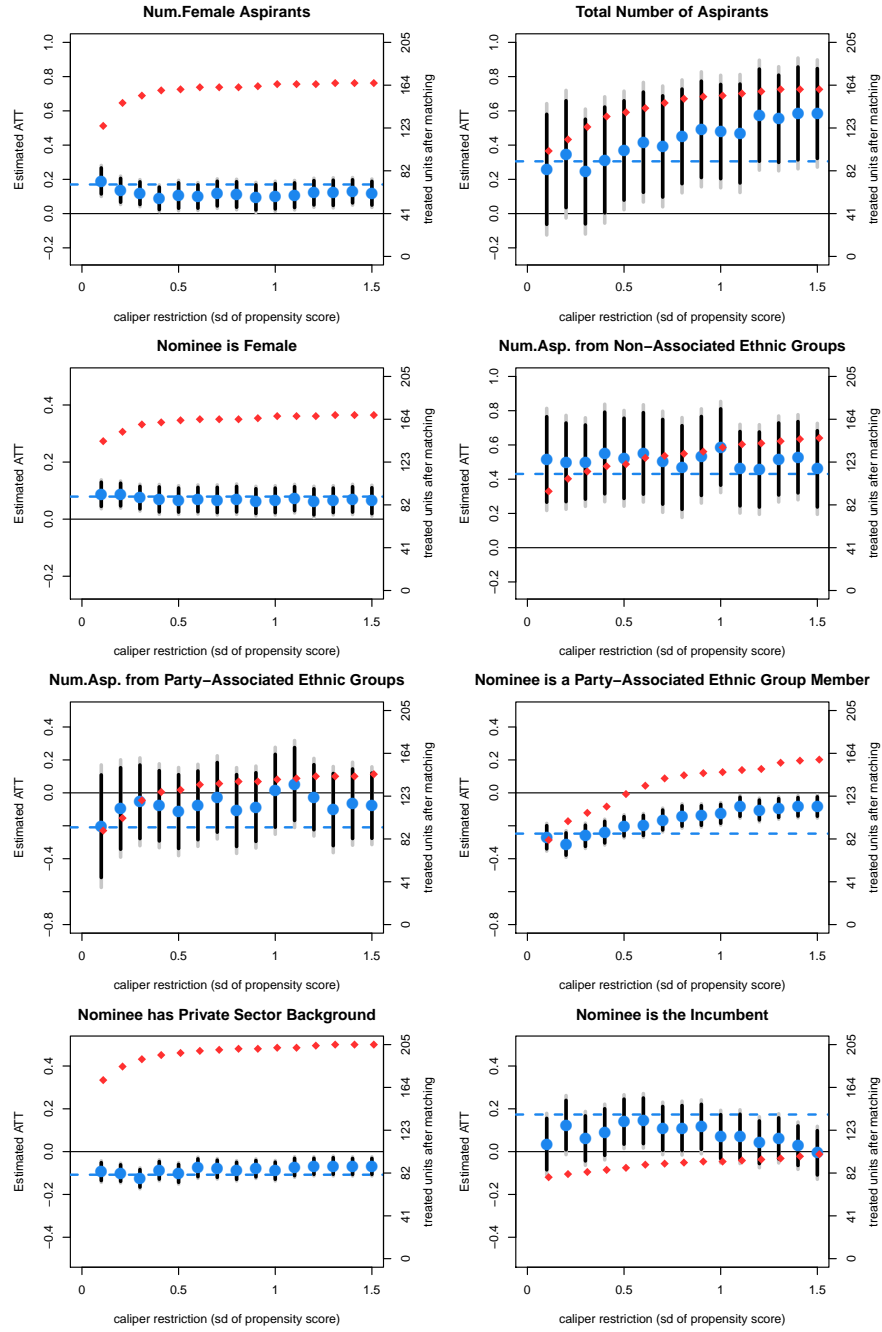
Figure C1: *Estimated average effects after optimal full matching with calipers.*

C.5   Analysis without Exact Matching on 2012 Outcome

Figure C2 shows that all except one of our results hold when we no longer require exact matches on the 2012 value of each outcome variable. The estimate for the average effect on the number of female candidates is still signed in the same direction, but no longer statistically significant at conventional levels. Corresponding balance statistics are in Table C6.

Table C6: *Overall Balance without Exact Matching on 2012 Outcome.*

|	χ^2	df	p-value	n_T	Sets
Num. Female Aspirants	0.48	6	1.00	170	93
Total Number of Aspirants	0.15	5	1.00	170	94
Nominee is Female	0.97	7	1.00	170	96
Num. Asp. from Non-Associated Ethnic Groups	2.66	8	0.95	150	82
Num. Asp. from Party-Associated Ethnic Groups	6.43	8	0.60	150	77
Nominee is a Party-Associated Ethnic Group Member	0.85	8	1.00	169	55
Nominee has Private Sector Background	1.54	5	0.91	211	113
Nominee is the Incumbent	1.83	6	0.93	102	50
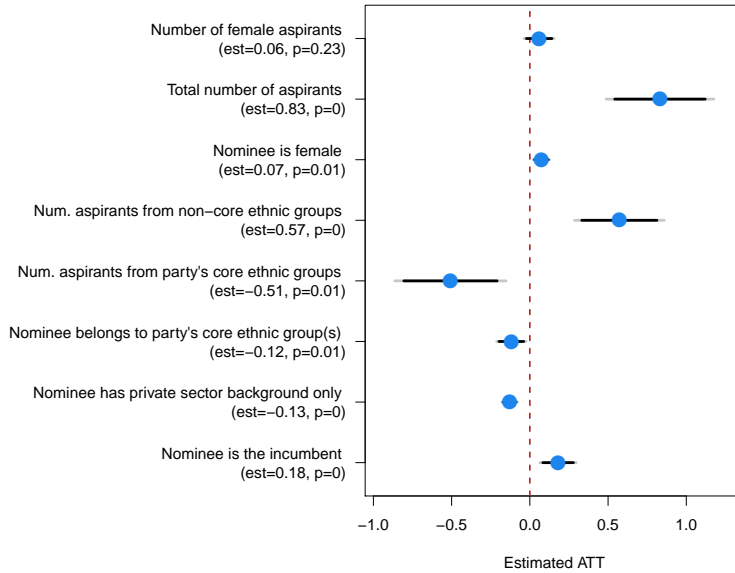
Figure C2: *Estimated average effects after matching, without exact matching on 2012 outcome.*

C.6   Alternative Matching Approaches

A variety of matching methods are available for applied research. Figures C3 and C4 present summaries of balance and the estimated ATT using optimal full matching (our preferred model), along with three matching methods with 1:1 matching – propensity score matching with a logit model to calculate the propensity score, multivariate matching using Mahalanobis distance, and genetic matching. The first two of these alternative methods are popular in social scientific research, while the third method uses a genetic search algorithm to look for the optimal weight to be given to each covariate and can be considered a generalization of the first two methods.

For each outcome, we use the same set of matching variables as in the main specification, including exact matching on the 2012 value of the outcome variable, and we match, with replacement, one control for each treated unit. We use regression to estimate the ATT while adjusting for bias and report Abadie-Imbens standard errors. These are implemented with the package `Matching` version 4.9-6 in `R`, version 3.6.0 (Sekhon 2011).

For each method, we have 8 outcomes, each with their own set of matching variables. We summarize the covariate balance achieved by each of these methods in Figure C3, with each panel corresponding to a different matching method (or no matching). Each panel is a frequency distribution of the absolute value of the standardized mean difference between treatment and control for each of the individual matching variables for all eight outcomes. For the panel on optimal full matching approach, the numerator is a weighted average of within-stratum differences in means on the covariate. Each panel is a histogram of the absolute values of the magnitudes represented by the red squares in the two figures from the subsection on "Balance on Individual Matching Variables". The dashed line in each panel indicates the mean of these absolute values for each method. We can see that all four matching methods improve balance by this measure, shifting the distributions towards zero from the baseline of no matching, but that optimal full matching is notably more successful than the other three approaches. We thus prefer optimal full matching because it produces better balance than the alternatives.

No matching
mean= 0.208

Propensity score
mean= 0.148

Genetic matching
mean= 0.097

Mahalanobis distance
mean= 0.093

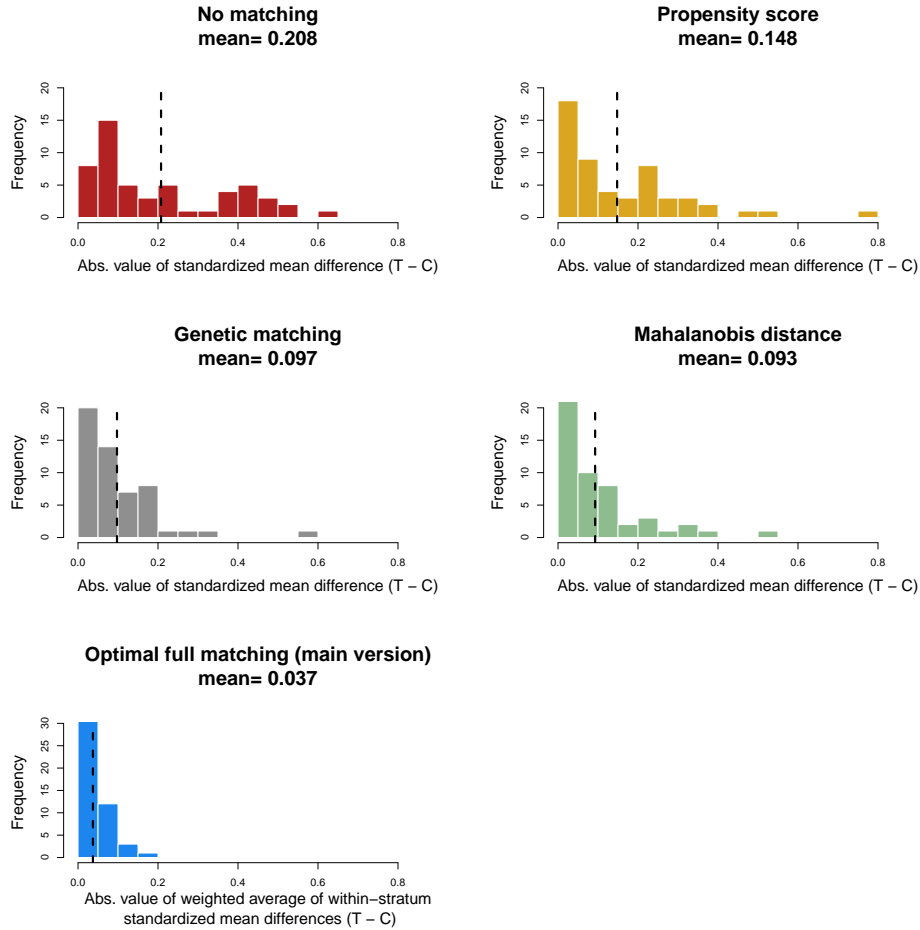**Optimal full matching (main version)**
mean= 0.037

Figure C3: *Summary of Balance on Individual Matching Variables with Alternative Matching Methods.* Each panel summarizes a different matching method. The dotted line indiates the average of the absolute values of the standardized mean differences between treatment and control on all the individual matching variables for all eight outcomes.
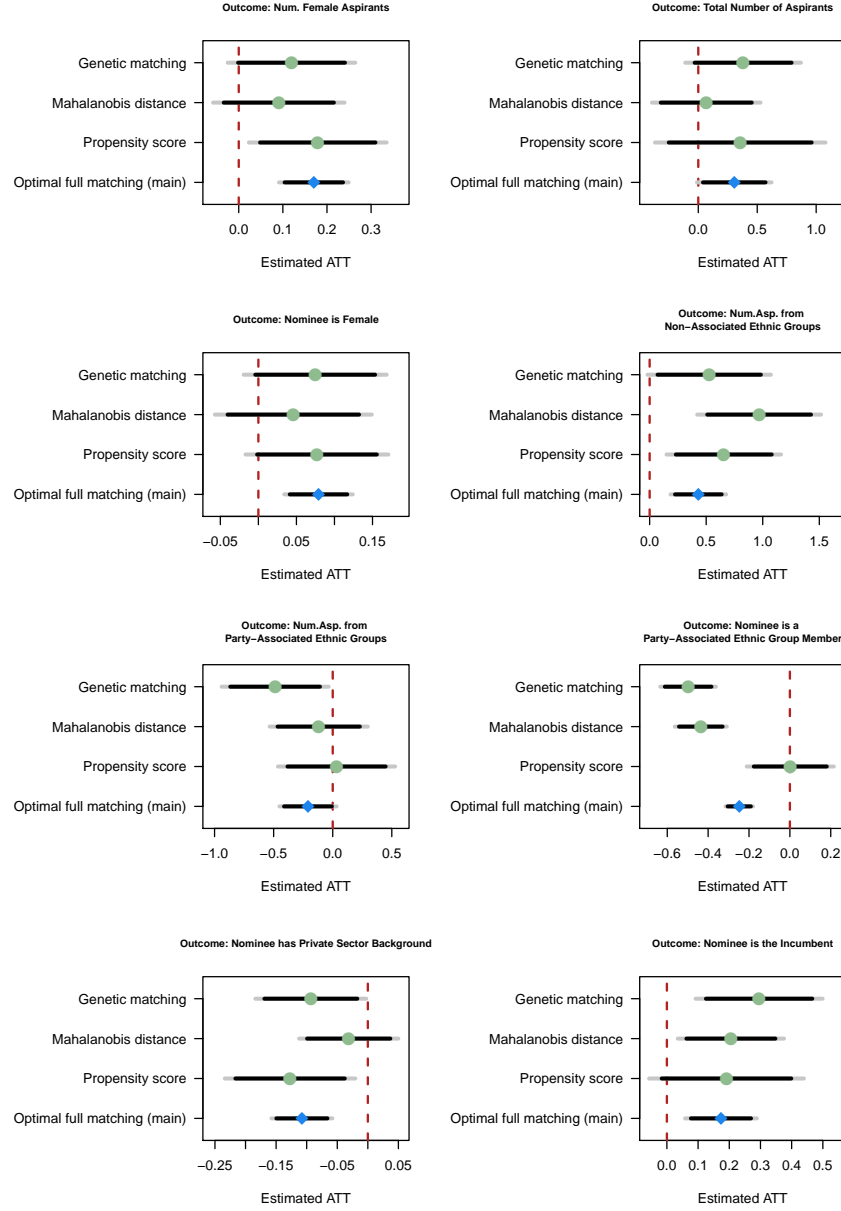
Figure C4: *Estimated ATT with Alternative Matching Methods.* The circles indicate the estimated effect, with the darker and lighter lines indicating 90% and 95% confidence intervals, respectively.