Estimating electoral bias against candidate traits

Georgina Evans[*], Jorge M. Fernandes[†], Sascha Riaz[‡]

October 12, 2021

Abstract

Are voters biased against certain types of candidate, such as women? This question is a subject of immense scholarly and political interest, since parties are disinclined to nominate candidates whom they expect to incur an electoral penalty. Yet, despite widespread interest in the subject, we identify a set of shared methodological shortcomings in prominent studies that limit the reliability of existing inferences. Most significantly, few studies provide a precise definition of the notion of voter bias being estimated. In response to this, we provide a unified framework for defining types of bias against candidates and propose an estimation strategy to identify the electoral penalty parties have incurred in the past for nominating certain candidates, under clearly specified assumptions. We apply our estimator in two settings: First, we study the effect of gender in the United Kingdom and Germany. In contrast to prior experimental work, we find precisely estimated null effects for voter bias against women. Second, we examine the effect of education in Germany, where we find an electoral premium of about 0.5 percentage points for candidates with doctoral degrees.

[*]PhD Candidate, Department of Government, Harvard University, georginaevans@g.harvard.edu
[†]Assistant Research Professor, University of Lisbon jorge.fernandes@ics.ulisboa.pt
[‡]PhD Candidate, Department of Government, Harvard University riaz@g.harvard.edu

1 Introduction

Do parties face an electoral penalty when they nominate candidates with certain demographic characteristics? A growing body of scholarly work investigates how voters react to candidates' personal features, with a particular emphasis on gender (Seltzer, Newman and Leighton 1997; Dolan 2018; Fujiwara, Hilbig and Raffler 2020; Schwarz, Hunt and Coppock 2020), race, and immigration status (Fisher et al. 2015; Portmann and Stojanović 2019, 2021), sexual orientation (Magni and Reynolds 2020), and age (Eshima and Smith 2021). The resulting evidence is mixed. While some studies find that voters do not penalize candidates for personal features, others show that candidates belonging to underrepresented groups suffer electoral loses.

The electoral success of candidates from historically underrepresented minorities has important political ramifications. First, the presence of minorities in politics creates an opportunity to counter stereotypes, send a strong signal of competence, and to act as role models for other members of their group (Wolbrecht and Campbell 2007; Alexander 2012; Weiss 2020). Second, the inclusion of historically marginalized groups adds their voice to the policy-making process, which increases the likelihood to reap policy benefits for that group (Reynolds 2013; Broockman 2013; Butler and Broockman 2011; Juenke and Preuhs 2012).

Over the past decades, a number of observational and experimental studies have advanced our understanding of voters bias against candidates. However, while observational studies offer strong external validity, most of them are plagued with internal validity issues. We identified a number of common identification issues in prior research based on observational data.[1]

First, previous studies on voter bias commonly lack a precise definition of the key quantity of interest the study seeks to learn about. When the estimand is not clearly defined, it is impossible to evaluate whether the research design and method employed in a paper

[1]A list of the studies we considered is available online.

are appropriate to identify it (Rubin 2005; Blair et al. 2019). Terms such as 'inference', 'unbiased', or 'identification strategy' only have meaning in relation to at least one precisely defined quantity. In this context, it is important to note that regression coefficients do not necessarily correspond to meaningful quantities that researchers or policy-makers care about (Aronow and Samii 2016). Terms such as the 'effect' of a given variable on the outcome are needlessly imprecise.

Second, most previous studies do not explicitly state the assumptions required for their method of choice to return an informative estimate of electoral bias. This issue in part relates to the previous point: when the estimand is not precisely defined, it is impossible to state the identification assumptions required to obtain an unbiased estimate of it.

Third, the matching estimator we propose constitutes a substantial improvement over previous studies in terms of model dependence. Prior studies of voter bias generally adjust for potential confounders as 'control variables' in a multivariate regression framework. However, regression methods for covariate adjustment can be highly sensitive when there are substantial differences in the covariate distributions between the treated and control groups (Imbens 2015; Ho et al. 2007; Athey and Imbens 2015). We view this as a common issue in observational studies of electoral bias against candidates: constituencies in which minority candidates are/are not nominated to run as political candidates are likely strongly dissimilar in terms of their covariate profiles (e.g. with respect to urbanity, demographics, ideology, etc.).

Fourth, we emphasize that party vote-shares measured in the zero-one interval are by construction interdependent within constituencies. To the extent that the causal effect of nominating a candidate with a certain profile on a given party's vote-share is nonzero, this implies that the vote-share of at least one other party must shift by definition. This is a type of spillover effect that can be viewed as a violation of the Stable Unit Treatment Value Assumption (SUTVA) in causal inference (Rosenbaum 2007). The potential outcome (vote-share) of a given party is by definition affected by the treatment status of other parties within the same constituency, unless the causal effect of nominating a candidate with a certain trait on his or her party's vote-share is precisely zero. We address this issue by explicitly comparing candidates of the same party *across* constituencies, holding constant the treatment status of other parties. Generally, our proposed matching algorithm clearly delineates the kinds of treated-control comparisons we leverage to estimate the causal effect of interest.

One way that researchers have dealt with common issues of causal inference using observational data is to turn to survey experiments instead. This type of experimental research has greatly contributed to our understanding of voters' biases against candidates. However, we argue that researchers should preferably seek to draw causal inferences from observational 'real' election data whenever possible. Prior research suggests that respondents tend to evaluate hypothetical candidates and real candidates differently in experiments (McDonald 2020). In addition, survey respondents have experimental demand effects – that is, they adapt their answers to conform to socially accepted norms to confirm the researcher's hypothesis (Zizzo 2010; Berinsky, Huber and Lenz 2012). Finally, experimental studies often rely on convenience samples, which may not be representative of the population of eligible voters as a whole (Findley, Kikuta and Denly 2021). Taken together, internal validity problems in existing observational studies and external validity issues in survey experiments motivate us to find a common framework bringing together the best of both worlds.

In writing this paper, our primary motivation is to square the circle between external and internal validity. We argue that scholars do not need to face a trade-off between them. Our contribution shows that we can causally identify electoral bias using observational data. Our approach starts with a clear definition of the theoretical quantity of interest: electoral bias against candidates that possess a certain trait. In a second step, we proceed to construct an unbiased estimator for this quantity on the basis of panel election data. Our estimator has desirable properties under clearly defined assumptions. We propose a design-based strategy (Rubin et al. 2008) to draw inferences about aggregate electoral bias from observational data.

Our method closely mirrors the design-based strategy for causal inference in difference-in-differences designs.

Our method is useful in institutional contexts with plurality electoral systems with a personalization component, in which voters have the capacity to vote or reward candidates based on their personal traits (Cain, Ferejohn and Fiorina 1987). Furthermore, our approach only applies to electoral systems with a single-member districts component. Our scope conditions make it ideal to study, among others, the United States, the United Kingdom, France, Germany, Canada, New Zealand, and India.

We implement our estimator by looking at the effects of gender and education in electoral bias in the United Kingdom and Germany. We look at the former because the wide scholarly body examining it offers a benchmark against which we can compare our results. In contrast, education has received only scant attention in the literature. We estimate the electoral penalty parties face when nominating female candidates on the basis of more than 41,000 political candidates during the time period 1979 – 2019. We find precisely estimated null effects for electoral bias against female candidates in both countries. All other things being equal, female candidates do not obtain an electoral premium (or penalty). This finding contradicts prior experimental work, which has found a positive bias of about 2 percentage points in favor of female candidates (Schwarz, Hunt and Coppock 2020).

Second, we examine the extent to which voters reward candidates with higher education. We study the case of Germany, where candidates typically list doctoral degrees next to their name on election ballots. We find an electoral premium of about 0.5 percentage points for candidates with PhDs. Our findings suggest pronounced differences between studying bias in experimental vs. observational settings.

2 Measuring Bias Against Candidates

Electoral bias against candidate traits is not a single concept. Yet, prior work has often been equivocal about which quantity the employed empirical strategy is intended to estimate.[2] As we will show, this is not mere technicality: different estimands correspond to different political processes and it is possible that bias exists on some metric and not another. In this section we thus introduce unifying notation that allows us to specify a range of quantities researchers may be interested in.[3]

2.1 Notation

We propose to estimate electoral bias on the basis of panel election data. We observe the election results for P political parties in C constituencies over T elections. In addition to the number of votes each party gained in a given constituency, we observe its treatment status, i.e. whether the party nominated a candidate with a certain trait. Our approach generalizes to any binary candidate characteristic researchers that can be operationalized as a binary variable (e.g. gender, sexual orientation, ethnicity, education, etc.). We note that we draw on the general framework for causal inference with time-series cross-sectional data proposed by Imai, Kim and Wang (2020).

Formally, we let $\mathbf{W}_{c,t} = (W_{1,c,t}, ..., W_{P,c,t})$ denote the treatment status for the P parties in constituency c in election period t. In our running example, we will estimate bias against female candidates, that is $W_{p,c,t} = 1$ if party p nominates a female candidate in constituency c in election t, and $W_{p,c,t} = 0$ otherwise. We use the notation $\mathbf{w}_{-p,c,t}$ to denote the treatment status vector of all parties in constituency c, except for party p. Then we denote by $\mathbf{Y_{c,t}}(\mathbf{w}_{c,t}) = (Y_{1,c,t}(\mathbf{w}_{c,t})), ..., Y_{P,c,t}(\mathbf{w}_{c,t})))$ the potential outcomes of the P parties in constituency c under treatment allocation $\mathbf{w}_{c,t}$. The observed outcome is thus $\mathbf{Y_{c,t}} = \mathbf{Y_{c,t}}(\mathbf{W}_{c,t})$.

[2]We note Schwarz, Hunt and Coppock (2020) as an important exception in this regard.
[3]We closely follow the notation from the interference literature (Tchetgen and VanderWeele 2012).

The outcome will typically be measured as the vote share won by a given party p in a given constituency c in election period t, such that $Y_{p,c,t}(\mathbf{w}_{c,t}))) \in [0, 1]$.

We use this notation to formalize and account for the inherent dependence between party vote shares within a constituency, which we view as a specific form of 'interference' (citation). Whether a voter casts a ballot for a specific candidate (party) in a constituency by definition depends on the qualities and characteristics of other candidates competing in the same constituency.[4] To the extent that the treatment effect is nonzero, the potential outcomes for any party within a constituency necessarily vary with the treatments assigned to other parties. The standard regression assumption of (conditional) independence across observations does not hold by construction when the outcome variable are aggregate voting results. To account for spillover effects of this nature, we define the party-level potential outcome as a function of a constituency-level treatment assignment vector $\mathbf{w}_{c,t}$ rather than the party-level treatment indicator $W_{p,c,t}$.

2.2 Estimands

In what follows, we introduce three important causal estimands. While interconnected, these estimands are empirically distinct and relate to diferent strategic political considerations.

2.2.1 Individual Direct Effect

At the level of the individual party, we can define the individual direct effect of treatment (Hudgens and Halloran 2008):

$$\tau_{p,c,t}(\mathbf{w}_{-p,c,t}) = Y_{p,c,t}(\mathbf{w}_{-p,c,t}, w_{p,c,t} = 1) - Y_{p,c,t}(\mathbf{w}_{-p,c,t}, w_{p,c,t} = 0) \quad (1)$$

This considers the difference in vote-share party p would experience if it were to nominate a female candidate compared to a male candidate, holding the candidate gender of the

[4] Put differently, the events of voting for two candidates A and B in a constituency are mutually exclusive (and therefore dependent) events. The Stable Unit Treatment Values Assumption (SUTVA) is violated by construction.

candidates of all remaining parties represented by $\mathbf{w}_{-p,c,t}$ fixed. We explicitly define this estimand as a function of $\mathbf{w}_{-p,c,t}$ because the treatment effect may depend on the treatment status of the remaining parties in the same constituency. For example, if all other parties except p in a given constituency already nominate female candidates, the electoral penalty for party p for nominating a woman might be much smaller compared to a scenario in which all competing parties nominate male candidates (see also Section 2.1).

When choosing between different estimands and identification strategies, it is important to think about strategic interdependence between the candidate nominations of different parties within the same constituency in the same election period. We consider candidate nominations to be independent when parties do not condition their candidate nominations on the candidate nominations of other parties. Formally:

$$P(W_p) = P(W_p|\mathbf{W}_{-p})$$

In our running example, strategic interdependence would imply that parties are more (less) likely to nominate a female candidate if a competing party in the same constituency chose to nominate a female candidate. To the extent that there is no strategic interdependence between the candidate nominations of parties within a given constituency, the individual direct effect of treatment is arguably the most relevant quantity from a political point of view. When parties decide whether or not to nominate a female candidate, they might ask: assuming that other parties will not condition their candidate nominations on which candidate we chose to nominate, should we expect an electoral penalty when nominating a female vs. male candidate?

2.2.2 Individual Total Effect

To the extent that candidate nominations are interdependent between parties within constituencies, parties might care about the individual total effect of the treatment, which we define as:

$$\delta_{p,c,t} = Y_{p,c,t}(w_{p,c,t} = 1) - Y_{p,c,t}(w_{p,c,t} = 0) \tag{2}$$

The main difference between this estimand and the Individual Direct Effect defined previously is that we do not hold fixed the treatment status of all other parties in the same constituency. Rather, we're interested in the total effect of nominating a woman, which might also include the downstream effects that nominating a female candidate might have on the candidate nominations of other parties. If there is no strategic interdependence in candidate nomination, then this quantity is identical to the individual direct effect evaluated at the realized values of candidate traits of the remaining parties.

2.2.3 Indirect Causal Effect

Finally, parties might be interested in the effect that the candidate nominations of other parties have on their own vote shares. To the extent that the direct treatment effect is nonzero, the vote share of a given party p will depend on the characteristics of the candidates nominated by all other parties. This notion is captured by the indirect causal effect of the treatment, which we define as:

$$\gamma_{p,c,t}(\mathbf{w}_{-p,c,t}, \mathbf{w'}_{-p,c,t}) = Y_{p,c,t}(\mathbf{w}_{-p,c,t}, w_{p,c,t} = g) - Y_{p,c,t}(\mathbf{w'}_{-p,c,t}, w_{p,c,t} = g) \tag{3}$$

This estimand considers the difference in the vote share of party p under two different treatment assignments for the remaining parties, holding fixed the gender of their own candidate at $g \in \{0, 1\}$.

2.3 Identification

Our contention is that all three estimands are interesting from both a political and theoretical point of view. Going forward, we will focus our attention on estimating the individual direct effect of the treatment. We chose this estimand as our primary target because considerations of strategic interdependence are likely limited when parties decide whether or not to nominate a female candidate. Prior research suggests that candidate selection is primarily driven by internal party dynamics (Hazan and Rahat 2010). We leave estimation of other quantities of interest to future research.

The fundamental problem of causal inference in our setting stems from the fact that we only observe the potential outcome under one specific treatment allocation for one party in a given constituency in a given election period. That is, we only observe the realized vote shares given the candidates that parties actually nominated. Estimating bias against certain candidates (e.g. women) requires making inferences about what vote shares would have been observed if parties had chosen to nominate different candidates, with different characteristics, in the *same* constituency and election period.

We hence cannot identify causal effects at the individual candidate level. Instead, we focus on estimating average causal effects. For identification, we take advantage of panel data using a difference-in-difference matching estimator. This will allow us to identify the finite sample average direct effect for a subset of party-constituency-years. In combination with an assumption about the representativeness of the units in our matched sample, we can generalize to the sample average direct effect.

We start by writing $\tau_{p,c,t}(\mathbf{W}_{-p,c,t}) =$

$$[Y_{p,c,t}(w_{p,c,t} = 1, \mathbf{W}_{-p,c,t}) - Y_{p,c,t-1}(w_{p,c,t-1} = 0, \mathbf{W}_{-p,c,t})] -$$
$$[Y_{p,c,t}(w_{p,c,t} = 0, \mathbf{W}_{-p,c,t}) - Y_{p,c,t-1}(w_{p,c,t-1} = 0, \mathbf{W}_{-p,c,t})]$$

$$= \Delta Y_{p,c,t}(1, \mathbf{W}_{-p,c,t}) - \Delta Y_{p,c,t}(0, \mathbf{W}_{-p,c,t})$$

Where $\Delta Y_{p,c,t}(1, \mathbf{W}_{-p,c,t})$ represents the change in vote share party p would experience if it nominated a man at time $t-1$ and a woman at time t, and $\Delta Y_{p,c,t}(0, \mathbf{W}_{-p,c,t})$ represents the change the same party would experience if it had nominated a man in both periods. $\mathbf{W}_{-p,c,t}$ is the observed treatment vector of the remaining parties. We then note that $\Delta Y_{p,c,t}(1, \mathbf{W}_{-p,c,t})$ is observed for party p in constituency c at time t if the party nominated a man at time $t-1$ and a woman at time t, and all other parties nominated a candidate of the same gender in the two period. To estimate direct effects, we will therefore focus on the subset of party-constituency-election units for which we observe this change. For treated units, we observe a change in their treatment status from one period to the next such that $w_{p,c,t-1} = 0$ and $w_{p,c,t} = 1$, while the treatment status of all other parties remains unchanged, such that $\mathbf{w}_{-p,c,t} = \mathbf{w}_{-p,c,t-1}$. We will denote this set of *treated* units as $\mathcal{T}$.

Our estimand can therefore be written as:

$$\frac{1}{\mathcal{T}} \sum_{p,c,t \in \mathcal{T}} \tau_{p,c,t}(\mathbf{W}_{-p,c,t}) = \frac{1}{\mathcal{T}} \sum_{p,c,t \in \mathcal{T}} \Delta Y^{obs}_{p,c,t} - \Delta Y_{p,c,t}(0, \mathbf{W}_{-p,c,t})$$

Below, we illustrate an example of a treated unit on the left, with the corresponding profile for the constituency under control on the right. Each row corresponds to one of three parties, A, B, and C. Each column corresponds to one election period. The entries $W_{p,t}$ indicate the gender of the candidate a given party nominated in an election.

$$\text{Treated:} \quad \begin{array}{c} \\ A \\ B \\ C \end{array} \!\!\begin{array}{c} \begin{array}{cc} t-1 & t \end{array} \\ \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \end{array} \qquad \text{Control:} \quad \begin{array}{c} \\ A \\ B \\ C \end{array} \!\!\begin{array}{c} \begin{array}{cc} t-1 & t \end{array} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \end{array}$$

All that is left then is to construct a valid estimate of $\Delta Y_{p,c,t}(0, \mathbf{W}_{-p,c,t})$.

To make progress, we assume parallel trends in potential outcomes conditional on a vector of past covariates.

Assumption 1

$$\Delta Y_{p,c,t}(1, \mathbf{W}_{-p,c,t}), \Delta Y_{p,c,t}(0, \mathbf{W}_{-p,c,t}) \perp W_{p,c,t} | X_{p,c,t-1}$$

We also impose the assumption of no interference *across* constituencies. This implies that voters in a given constituency c do not condition their vote choice on the characteristics of candidates in other constituencies c'.

Assumption 2

$$\mathbf{Y_{c,t}}(\mathbf{W}_{c,t}) \perp\!\!\!\perp (\mathbf{w}_{c',t})$$

Relying on these assumptions, we can use matching to estimate $\Delta Y_{p,c,t}(0, \mathbf{W}_{-p,c,t})$. For each treated unit, we identify a set of potential control units $M_{p,c,t}$. This matched set contains all units that share the same pre-treatment candidate gender profile as the treated unit *and* did not experience a change in the candidte-gender profile between period $t-1$ and t. Formally, this means that for constituency c treated in period t, we identify all units j for which $\mathbf{W_{c,t-1}} = \mathbf{w_{j,t-1}} = \mathbf{w_{j,t}}$.

We then further refine this matched set. For each treated unit, we find the M control units that are most similar in terms of observable characteristics. Researchers are free to choose among a variety of distance metrics to measure the similarity of treated and control units in this step. Options include the Mahalanobis distance in terms of observed covariates, estimated propensity scores, or CEM matching. Researchers can also choose to match exactly on some discrete covariates, e.g. geographic region.

In our running example, we implement one-to-one matching ($M = 1$) and match exactly on geographic region. We require that treated and control units are in the same NUTS 1 region, which corresponds to states (*Bundesland*) in Germany and statistical regions in the UK (e.g. the South East or West Midlands). We also require that the same party is the incumbent in treated and control units. To further address the potentially confounding effect of incumbency, we also drop those (few) cases in which a woman is nominated for the incumbent party in a given constituency.

From the set of available control units with the same incumbent party and within the same region, we choose the one that is closest in terms of pre-treatment electoral outcomes. We operationalize this on the basis of the Pedersen index:

$$D(c, j, t) = \frac{1}{2} \sum_{p} |Y_{p,c,t} - Y_{p,j,t}|$$

which gives the electoral distance between constituencies c and j in a given period t. We stress that our estimation never compares across parties. We compare changes in the voteshare of *the same party* across constituencies that are i) within the same geographic region and ii) similar in terms of electoral outcomes in the pre-treatement period $t-1$.

We use the set of matched control units to estimate $Y_{i,c,t}(0) - Y_{i,c,t-1}(0)$ and thereby $\tau_{i,c,t}$.

$$\widehat{\tau}_{p,c,t} = \Big(Y_{p,c,t} - Y_{p,c,t-1}\Big) - \Big(\frac{1}{|M_{p,c,t}|} \sum_{j \in M_{p,c,t}} Y_{j,c,t} - Y_{j,c,t-1}\Big)$$

Finally, we obtain our estimate of τ_{ATT} as

$$\widehat{\tau} = \frac{1}{\mathcal{T}} \sum_{p,c,t \in \mathcal{T}} \widehat{\tau_{p,c,t}}$$

3 Empirical Application

Next, we turn to an empirical application of our method to illustrate its potentialities and limitations. We examine voter bias in the United Kingdom and in the single-member district tier in Germany. We focus on candidate-centered elections because personal features of candidates tend to be more important for electoral outcomes in such settings than in party-centered contexts (Cain, Ferejohn and Fiorina 1987; Shugart, Valdini and Suominen 2005).

Our empirical illustration looks at the impact of gender and education in eliciting voter bias at the ballot box. More specifically, we examine the extent to which political parties suffer an erosion in their electoral results in constituencies where they select a woman as candidate. Furthermore, test whether the selection of candidates who hold a doctoral degree yields parties a more favorable result in the German context.

In selecting gender and education as personal features to analyze, we have different motivations. Gender and its effects on voter bias have received extensive scholarly attention in both observational and experimental studies. On aggregate, prior experimental work suggests that voters positively discriminate in favor of female candidates. In an extensive meta analysis of 67 conjoint experiments, Schwarz, Hunt and Coppock (2020) estimate that the average effect of being a woman (relative to a man) is a gain of approximately two percentage points. Observational studies have yielded more mixed results, with evidence for voter bias in some contexts but not in others (Schwindt-Bayer, Malecki and Crisp 2010). Existing work, however, struggles to square the circle between external and internal validity, as discussed previously. Our motivation in looking at gender is to use observational data over three decades to causally identify the effect of having a woman on the ballot on parties' electoral results.

Unlike gender, the effects of education in triggering voter bias at the ballot box have received scant attention. While some conjoint experiments have examined the effects of education (Hainmueller, Hopkins and Yamamoto 2014; Arnesen, Duell and Johannesson 2019; Horiuchi, Smith and Yamamoto 2020), our study is – to the best of our knowledge – the first to estimate the electoral benefits of academic titles in an observational setting. Thus, our primary motivation in looking at the effect of education is to add to the literature on voter bias by unpacking whether parties benefit from selecting candidates with higher levels of education. Prior research suggests that voters might favor highly educated candidates to the extent that they perceive academic titles as a signal of candidate quality or ability. Studies show that highly educated leaders are associated with higher rates of economic growth

(Besley, Montalvo and Reynal-Querol 2011), lower levels of corruption (Efobi 2015), and better educational policies (Diaz-Serrano and Pérez 2013), without hurting social representation and diversity (Dal Bó et al. 2017).[5]

3.1 Data

We obtained official data on all candidates running in German federal elections between 1980 and 2017 from the *Bundeswahlleiter*. For each candidate, we observe their gender and full name as listed on the ballot. We also observe the vote share each candidate obtained in their constituency.[6] We subset our data to candidates running for one of the six major parties which generally make up the German parliament (CDU/CSU, SPD, FDP, Greens, Linke, and AfD). In total, our data includes individual level information on more than 16,000 candidates. The share of female candidates running in German federal elections has consistently increased over time, from about 10% in 1980 to about one third in the 2017 federal election (see Figure 1). Doctoral degrees are generally listed next to candidate names on German election ballots (see Figure A.2 for an example).[7] Between 1980 and 2017, between 15% and 20% of candidates held doctoral degrees.[8]

For the United Kingdom, we obtained candidate level data for all general elections between 1979 and 2019. Our data stems from two sources. For elections up to 2010, we draw on data collected by Eggers and Spirling (2014). For the 2015, 2017, and 2019 election, we obtained official data from the *House of Commons Library* (Baker et al. 2017). We note that for the British election data prior to 2010, we do not observe candidate gender. We impute each candidate's gender based on their name and prefix. The 2015, 2017, and 2019 election data comes with precise information on candidate sex. In total, our data includes individual

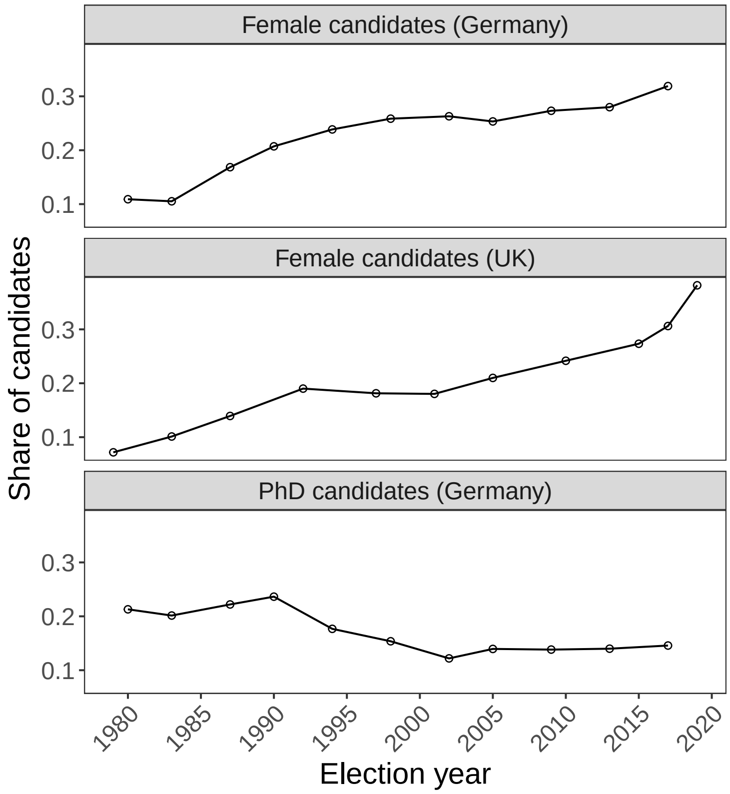
[5]However, see Carnes and Lupu (2016) for a rebuttal of these results. The authors find a weak link between leaders with college degrees and public policy improvement.

[6]We focus on the first vote *('Erststimme')* that is casted for candidates. We discard the second vote *('Zweitstimme')*, which determines the share of seats each party obtains in the German parliament.

[7]We note that in addition to doctoral titles, the place of residence and occupation of each candidate is generally also listed on the ballot in smaller font size (see Figure A.2).

[8]In Figure A.3 in the appendix, we show descriptive data on the number of PhD titles awarded in Germany between 1993 and 2020.

Figure 1: Female/PhD candidates in sample by election year.



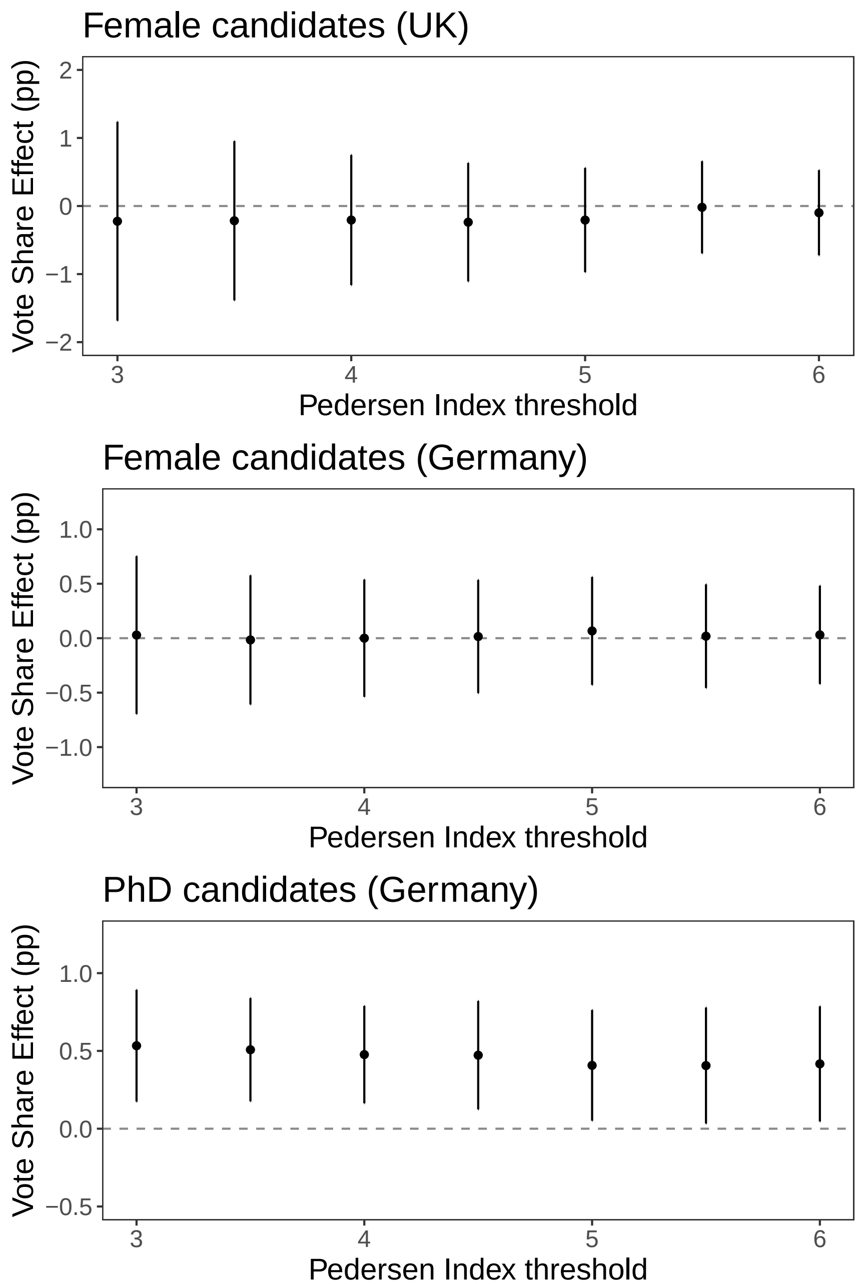
Note: The figure shows the share of female and PhD candidates in our sample over time.

level information on more than 27,000 candidates. We subset our data to candidates running for one of the six major parties (Conservative, Labour, Lib Dems, UKIP, SNP , Green Party). Similar to the German case, the share of female candidates for office has steadily increased over the last decades to about one third in the most recent British general election in 2019.

3.2 Results

We present the main results from our analysis in Figure 2. The y-axis shows our estimates for the individual direct effect of each candidate trait (female/PhD title) on the aggregate vote share a candidate receives in his or her constituency. The x-axis shows electoral Pedersen index threshold we use when matching treated and control constituencies. The lower the

Figure 2: Main results



Note: The vertical axis shows the estimated causal effects of candidate traits. The horizontal axis varies the threshold for including a matched pair of constituencies in our estimation, with lower levels of the Pedersen Index corresponding to more similar vote share distributions within the matched set. The upper panel compares the vote shares of female candidates compared to matched male counterparts in the UK. The middle panel is the analogous result for Germany. The bottom panel compares matched pairs vote shares of candidates in Germany with a PhD to those without.

Pedersen index threshold, the more similar we require treated and control constituencies to be in the election period immediately prior to treatment.

We find precisely estimated null results for bias against female candidates in both the United Kingdom and Germany. Our point estimates are close to zero. This result is robust for varying levels of electoral similarity between treated and control constituencies, as measured by the Pedersen index. Female candidates for political office, on average, do as well as their male counterparts when they are nominated in similar constituencies.

In contrast, we find a sizable, statistically significant electoral benefit of about 0.5 percentage points for German candidates who hold a PhD. Candidates who are listed with a doctoral degree next to their names on the ballot achieve significantly better electoral results. Again, our effect size estimates are not sensitive to varying the level of electoral similarity between treated and control constituencies. We stress that we are isolating the raw effect of the academic title, while accounting for other determinants of electoral results. Most importantly, our method explicitly conditions on the strength of the candidate's party in a given constituency and election year. To put the effect size magnitude into perspective, we note that about 3% of the districts in the 2013 German federal election were won with a vote margin of less than half a percentage point.

We see that as we increase the Pedersen threshold, the standard errors of our estimates decrease. However, we note that we obtain precise estimates of electoral bias even at very low Pedersen thresholds. For example, a Pedersen index value of 3 corresponds to a situation in which the overall difference in the distribution of votes across parties between two constituencies is a mere 3 percentage points.

Our sample size increases as we allow for greater electoral dissimilarity between treated and control units (see Figure A.1). The number of treated units we can match to control units depends on how much dissimilarity between treated and control units researchers are willing to tolerate. Our approach makes this trade off between bias and variance – inherent to all statistical analyses of voter bias – transparent and an inherent feature of the analysis.

For example: while about 5,800 female candidates ran for office in the UK between 1979 and 2019, we can only credibly estimate the treatment effect for a subset of about 800, even at relatively high Pedersen thresholds. We do not view this as a downside of our approach. Rather, our approach highlights the fact that credible units of comparison are not available for many cases in which candidates from historically underrepresented groups run for office. Researchers are of course free to choose less restrictive criteria for matching to increase the sample size. We could, for example, allow treated and control units to be matched across different regions. Standard regression approaches do not circumvent this issue, but instead make it hard for researchers to make transparent decisions about what kinds of treated-control comparisons they are willing to make.

Finally, we conduct a test for our key identification assumption: parallel trends between treated and matched control units in the absence of treatment. Specifically, we test whether the vote share of the party which nominated a female/PhD candidate evolved in parallel in the treated and matched control units *prior* to the nomination of the respective candidate. We use our matched data set to calculate placebo treatment effect estimates for the time periods $t-2$ to $t-1$.[9] If our identification assumption is met, we would expect the distribution of placebo effect estimates to be centered around zero. We present the results in Figure A.4. We find that for all the analyses we conduct, electoral results in treated and matched control units on average evolved in parallel prior to the nomination of a female/PhD candidate.

4 Discussion

Are voters biased against certain types of candidate? Our paper makes a twofold contribution to the literature on voter bias. First, we introduce a methodological framework that allows researchers to use observational data to isolate the causal effects of candidate attributes on electoral results. While we choose to examine gender and education in our empirical

[9] We note that for some units that are treated in early time periods (i.e. in the early 1980s, we do not observe the time period $t-2$.

application, our framework allows researchers to study and candidate trait they might be interested in. For example, researchers might be interested in examining how voters react to candidates depending on their race/ethnicity, age or even beauty. Our framework intends to strike a balance between analyzing observational data from real-world elections with all their vagaries and idiosyncrasies, and the need to establish causal relationships.

Second, in addition to its methodological contribution, our paper adds to existing work on voter bias based on gender and education. Focusing on Germany and the United Kingdom, our data set covers an extended period. Our results on gender confirm prior research that voters do not discriminate against women at the ballot box. Of course, the absence of bias at the ballot box does not mean that women are not subject of discrimination. Crucially, studies document the uphill battle that women have to face *before* gaining access to the ticket (Thomsen and King 2020; Fox and Lawless 2014). Thus, our result that voters do not display bias against women should be interpreted with caution. It it possible that female candidates are more qualified and better prepared than their male counterparts, as they might be subject to a harder intraparty selection process prior to candidate nomination (Bauhr and Charron 2021). Our approach exclusively focuses on aggregate voter bias in casting ballots for candidates who have already been selected to run. Discrimination against candidates from historically underrepresented groups might occur at multiple other stages of the election process.[10]

To the best of our knowledge, we are the first to study voter bias with respect to education, specifically academic titles, in an observational setting. We establish a robust and sizable electoral premium for candidates with doctoral degrees in Germany. Although prior research suggests a link between educated leaders and good public policy, bias against less educated candidates might have negative consequences in terms of descriptive representation of lower social classes. However, we caution against generalizing from these findings to other contexts. Given the high social status associated with academic titles in Germany, more research is

[10]See for example Folke and Rickne (2016) on discrimination against women in promotion decisions within political organizations.

needed to determine whether our results on voter bias in favor of highly educated candidates travel to other countries.

Before we conclude, we discuss a number of potential limitations of our proposed methodological framework. First, while our methods allows researchers to adjust for a variety of constituency-level features in a flexible way, adjusting for covariates at the candidate level is usually a much harder task. Data on candidate-level characteristics is usually sparse. Moreover, candidate-level characteristics (e.g. sex, education) generally do not vary over time. This puts practical limits on the degree to which researchers can adjust for unobserved candidate traits that might be correlated with the candidate characteristic they seek to study the effects of. For example, prior research suggests that candidates have differential access to campaign resources and funding, depending on their race and gender (Crowder-Meyer and Cooperman 2018; Grumbach and Sahn 2020).

Second, our method can only be employed in settings where voters cast a ballot at the candidate-level. While we expect that the flexibility of our framework appeals to researchers focusing on a constellation of candidate features, our method is of limited use to study voter bias in party-centered electoral systems. In settings where votes can only be attributed to parties but not candidates, we cannot use observational data to estimate bias against candidates (e.g. in Spain, Norway, Argentina, Uruguay, and Portugal). In these contexts, experimental work remains the best option to study voter bias.

While our study takes the literature one step forward by offering a framework to causally identify voter bias using observational data, there remain several important avenues to explore. Future research should study whether candidate nominations are strategically interdependent between parties in a given constituency. When a party selects a woman or a highly educated candidate and garners electoral benefits, does this create incentives for other parties to also select a candidate with similar features in the next election cycle? Whether or not parties condition their candidate nominations on the decisions of other parties has crucial implications for the study of voter bias, as we have lined out in this paper. Understanding the sources and magnitude of voter bias has crucial implications for descriptive representation in increasingly diverse democracies.

References

Alexander, Amy C. 2012. "Change in women's descriptive representation and the belief in women's ability to govern: A virtuous cycle." *Politics & Gender* 8(4):437–464.

Arnesen, Sveinung, Dominik Duell and Mikael Poul Johannesson. 2019. "Do citizens make inferences from political candidate characteristics when aiming for substantive representation?" *Electoral Studies* 57:46–60.

Aronow, Peter M and Cyrus Samii. 2016. "Does regression produce representative estimates of causal effects?" *American Journal of Political Science* 60(1):250–267.

Athey, Susan and Guido Imbens. 2015. "A measure of robustness to misspecification." *American Economic Review* 105(5):476–80.

Baker, Carl, Olivier Hawkins, Lukas Audickas, Alex Bate, Richard Cracknell, Vyara Apostolova, Noel Dempsey, Roderick McInnes, Tom Rutherford and Elise Uberoi. 2017. "General Election 2017: Full Results and Analysis." *Parliament UK. July* 13.

Bauhr, Monika and Nicholas Charron. 2021. "Will Women Executives Reduce Corruption? Marginalization and Network Inclusion." *Comparative Political Studies* 54(7):1292–1322.
URL: *https://doi.org/10.1177/0010414020970218*

Berinsky, Adam J, Gregory A Huber and Gabriel S Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk." *Political Analysis* 20(3):351–368.

Besley, Timothy, Jose G Montalvo and Marta Reynal-Querol. 2011. "Do educated leaders matter?" *The Economic Journal* 121(554):F205–227.

Blair, Graeme, Jasper Cooper, Alexander Coppock and Macartan Humphreys. 2019. "Declaring and diagnosing research designs." *American Political Science Review* 113(3):838–859.

Broockman, David E. 2013. "Black politicians are more intrinsically motivated to advance blacks' interests: A field experiment manipulating political incentives." *American Journal of Political Science* 57(3):521–536.

Butler, Daniel M and David E Broockman. 2011. "Do politicians racially discriminate against constituents? A field experiment on state legislators." *American Journal of Political Science* 55(3):463–477.

Cain, Bruce, John Ferejohn and Morris Fiorina. 1987. *The personal vote: Constituency service and electoral independence.* Harvard University Press.

Carnes, Nicholas and Noam Lupu. 2016. "What good is a college degree? Education and leader quality reconsidered." *The Journal of Politics* 78(1):35–49.

Crowder-Meyer, Melody and Rosalyn Cooperman. 2018. "Can't buy them love: How party culture among donors contributes to the party gap in women's representation." *The Journal of Politics* 80(4):1211–1224.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson and Johanna Rickne. 2017. "Who becomes a politician?" *The Quarterly Journal of Economics* 132(4):1877–1914.

Diaz-Serrano, Luis and Jessica Pérez. 2013. "Do More Educated Leaders Raise Citizens' Education?".

Dolan, Kathy. 2018. *Voting for women: How the public evaluates women candidates.* London: Routledge.

Efobi, Uchenna. 2015. "Politicians' attributes and institutional quality in Africa: A focus on corruption." *Journal of Economic Issues* 49(3):787–813.

Eggers, Andrew C and Arthur Spirling. 2014. "Electoral security as a determinant of legislator activity, 1832–1918: New data and methods for analyzing british political development." *Legislative Studies Quarterly* 39(4):593–620.

Eshima, Shusei and Daniel M. Smith. 2021. "Just a Number? Voter Evaluations of Age in

Candidate Choice Experiments." (ID 3704473).

Findley, Michael G, Kyosuke Kikuta and Michael Denly. 2021. "External Validity." *Annual Review of Political Science* .

Fisher, Stephen D, Anthony F Heath, David Sanders and Maria Sobolewska. 2015. "Candidate ethnicity and vote choice in Britain." *British Journal of Political Science* 45(4):883–905.

Folke, Olle and Johanna Rickne. 2016. "The glass ceiling in politics: Formalization and empirical tests." *Comparative Political Studies* 49(5):567–599.

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

Fujiwara, Thomas, Hanno Hilbig and Pia Raffler. 2020. "Party Nominations and Female Electoral Performance: Evidence from Germany.".

Grumbach, Jacob M and Alexander Sahn. 2020. "Race and representation in campaign finance." *American Political Science Review* 114(1):206–221.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments." *Political analysis* 22(1):1–30.

Hazan, Reuven and Gideon Rahat. 2010. *Democracy within parties: Candidate selection methods and their political consequences*. Oxford University Press.

Ho, Daniel E, Kosuke Imai, Gary King and Elizabeth A Stuart. 2007. "Matching as non-parametric preprocessing for reducing model dependence in parametric causal inference." *Political analysis* 15(3):199–236.

Horiuchi, Yusaku, Daniel M Smith and Teppei Yamamoto. 2020. "Identifying voter preferences for politicians' personal attributes: A conjoint experiment in Japan." *Political Science Research and Methods* 8(1):75–91.

Hudgens, Michael G and M Elizabeth Halloran. 2008. "Toward causal inference with interference." *Journal of the American Statistical Association* 103(482):832–842.

Imai, Kosuke, In Song Kim and Erik Wang. 2020. "Matching methods for causal inference with time-series cross-section data." 1.

URL: *https://imai.fas.harvard.edu/research/tscs.html*

Imbens, Guido W. 2015. "Matching methods in practice: Three examples." *Journal of Human Resources* 50(2):373–419.

Juenke, Eric Gonzalez and Robert R Preuhs. 2012. "Irreplaceable legislators? Rethinking minority representatives in the new century." *American Journal of Political Science* 56(3):705–715.

Magni, Gabriele and Andrew Reynolds. 2020. "Voter Preferences and the Political Underrepresentation of Minority Groups: Lesbian, Gay, and Transgender Candidates in Advanced Democracies." *Journal of Politics* .

McDonald, Jared. 2020. "Avoiding the Hypothetical: Why "Mirror Experiments" are an Essential Part of Survey Research." *International Journal of Public Opinion Research* 32(2):266–283.

Portmann, Lea and Nenad Stojanović. 2019. "Electoral discrimination against immigrant-origin candidates." *Political Behavior* 41(1):105–134.

Portmann, Lea and Nenad Stojanović. 2021. "Are Immigrant-Origin Candidates Penalized Due to Ingroup Favoritism or Outgroup Hostility?" *Comparative Political Studies* p. 00104140211024293.

Reynolds, Andrew. 2013. "Representation and rights: The impact of LGBT legislators in comparative perspective." *American Political Science Review* 107(2):259–274.

Rosenbaum, Paul R. 2007. "Interference between units in randomized experiments." *Journal of the American Statistical Association* 102(477):191–200.

Rubin, Donald B. 2005. "Causal inference using potential outcomes: Design, modeling, decisions." *Journal of the American Statistical Association* 100(469):322–331.

Rubin, Donald B et al. 2008. "For objective causal inference, design trumps analysis." *Annals of Applied Statistics* 2(3):808–840.

Schwarz, Susanne, William Hunt and Alexander Coppock. 2020. "What have we learned about gender from candidate choice experiments? A meta-analysis of 30 factorial survey experiments." *Journal of Politics* .

Schwindt-Bayer, Leslie A, Michael Malecki and Brian F Crisp. 2010. "Candidate gender and electoral success in single transferable vote systems." *British Journal of Political Science* 40(3):693–709.

Seltzer, Richard A, Jody Newman and Melissa Voorhees Leighton. 1997. *Sex as a political variable: Women as candidates and voters in US elections*. Boulder, CO: Lynne Rienner Publishers.

Shugart, Matthew Søberg, Melody Ellis Valdini and Kati Suominen. 2005. "Looking for locals: Voter information demands and personal vote-earning attributes of legislators under proportional representation." *American Journal of Political Science* 49(2):437–449.

Statista. 2021. "Promotionen an Hochschulen in Deutschland bis 2020 — Statista." https://de.statista.com/statistik/daten/studie/1244447/umfrage/bestandene-promotionen-an-hochschulen/. (Accessed on 10/06/2021).

Tchetgen, Eric J Tchetgen and Tyler J VanderWeele. 2012. "On causal inference in the presence of interference." *Statistical methods in medical research* 21(1):55–75.

Thomsen, Danielle M and Aaron S King. 2020. "Women's Representation and the Gendered Pipeline to Power." *American Political Science Review* 114(4):989–1000.

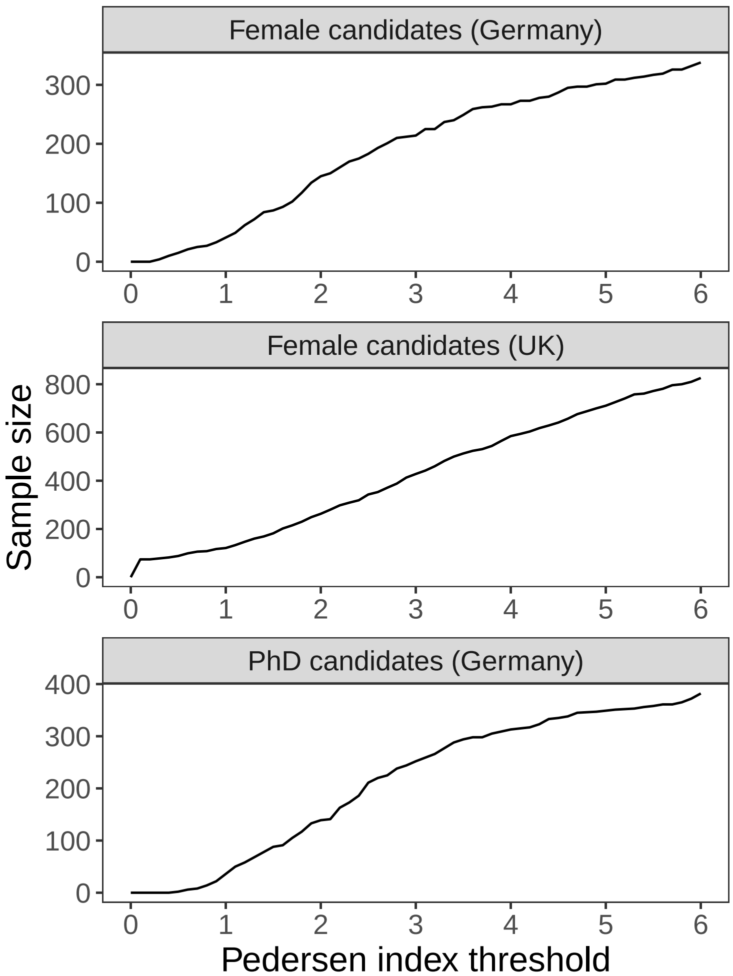
Weiss, Chagai M. 2020. "Reducing Prejudice through descriptive representation: Evidence From A Natural Experiment in Israeli Medical Clinics.".

Wolbrecht, Christina and David E Campbell. 2007. "Leading by example: Female members of parliament as political role models." *American Journal of Political Science* 51(4):921–939.

Zizzo, Daniel John. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics* 13(1):75–98.

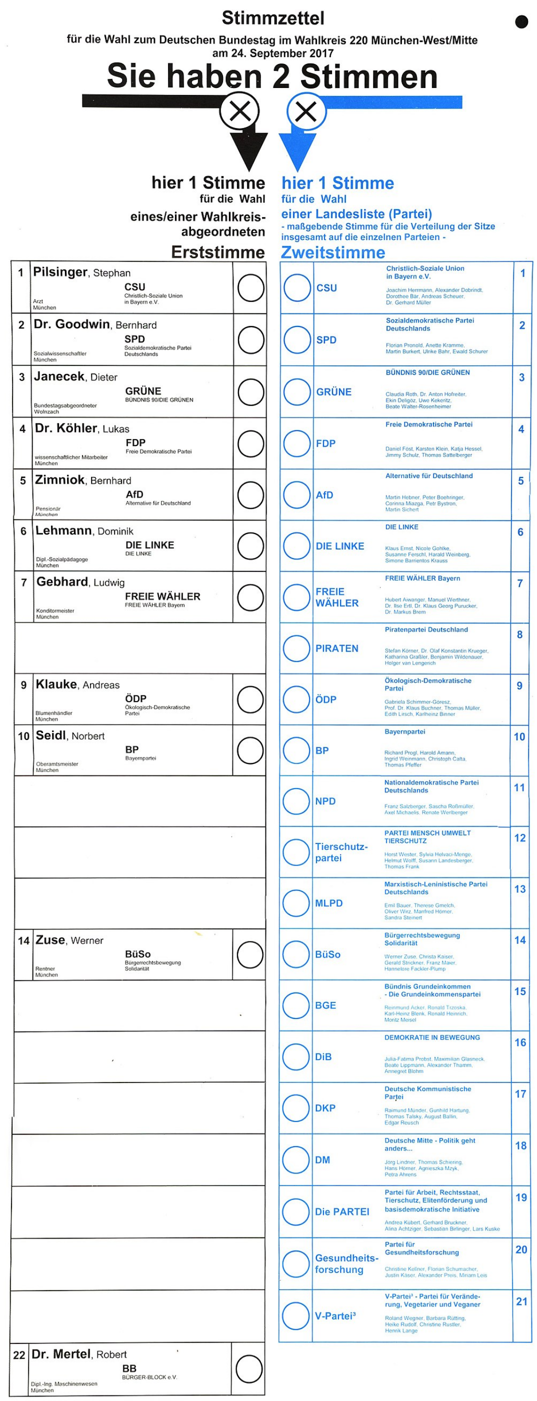
A Supporting Information (Online Only)

Figure A.1: Sample size by Pedersen index threshold



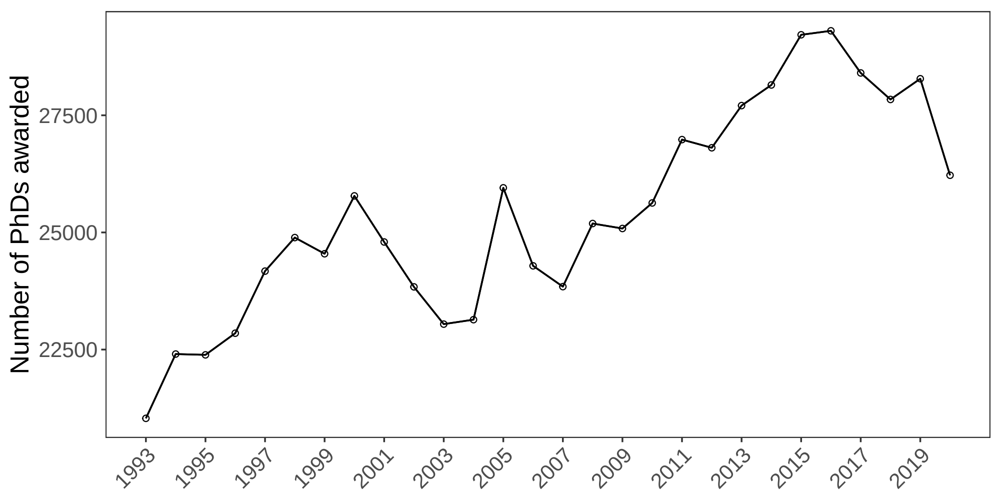
Note: The Figure illustrates the total sample size of our matched dataset (treated and control units) as a function of the Pedersen index threshold. The sample size increases as we allow for greater electoral dissimilarity between treated and control units.

Figure A.2: Example of German election ballot



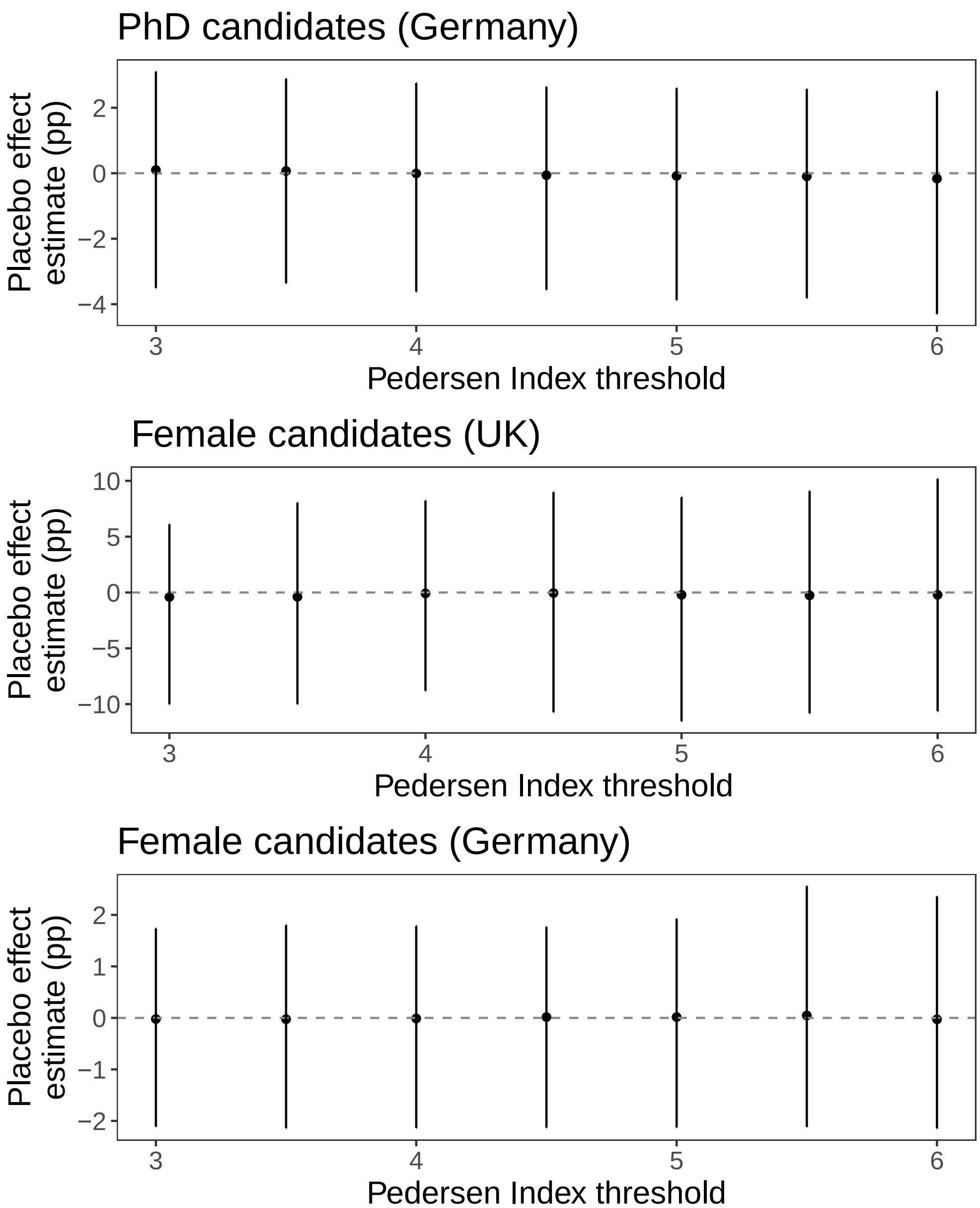
Note: Example ballot from the 2017 German federal election. Doctoral degrees ('Dr.') appear next to the names of each candidate.

Figure A.3: Number of doctoral degrees awarded in Germany over time



Note: The Figure shows the number of doctoral degrees awarded by German universities in each year between 1993 and 2020. The original data source is the German Statistical Office. We obtained the data from Statista (2021).

Figure A.4: Test for parallel trends



Note: The Figure shows the mean, the 2.5th, and the 97.5th percentile of unit-level treatment effect estimates from placebo difference in differences analyses for varying levels of the Pedersen index threshold. We calculate the difference in the change in the vote share of the treated party for treated and control units in the period prior to treatment ($t-2$ to $t-1$.