Incumbency Effects in German and British Elections: A Quasi-Experimental Approach [1]

Draft version. Please do not cite without permission of the authors

Jens Hainmueller [2]
Department of Government
Harvard University

Holger Lutz Kern [3]
Department of Government
Cornell University

May 19, 2005

[1] We would like to thank David S. Lee, Gary King, and Jonathan N. Katz for sharing their data. David S. Lee, Jonathan N. Katz, and Walter R. Mebane, Jr. provided greatly appreciated comments. As always, all remaining errors are our own responsibility.

[2] jens_hainmueller@ksg05.harvard.edu

[3] hlk23@cornell.edu

1 Abstract

Following the recent turn towards quasi-experimental approaches in the US literature on the incumbency advantage (Lee, 2001; Lee, forthcoming), we employ a Regression Discontinuity Design (RDD) to identify the causal effects of party incumbency in British and German post-World War II elections. The RDD framework exploits the randomized variation in incumbency status that occurs when a district race is close. Based on the assumption that parties do not exert perfect control over their observed vote shares, incumbents that barely won a race should be similar in their distribution of observed and unobserved confounders to non-incumbents that barely lost. This provides us with a naturally occurring counterfactual exploitable for causal inference under a weaker set of assumptions than conventional regressions designs commonly used in the incumbency literature. In both British and German federal elections, we find that party incumbency has a significant positive impact on vote shares and the probability of winning in marginal districts, the subpopulation of interest for which incumbency advantage is likely to make a difference. This stands in contrast to previous more ambiguous findings.

2 Introduction

The electoral advantage that incumbency bestows upon members of Congress is one of the most studied aspects of congressional politics in the United States (Gelman and King, 1990). Since the early 1970s, we have learnt a great deal about the growth, the causes, and the consequences of incumbency advantage in American elections, both at the federal and state level.[1] But more than twenty years ago, Cain, Ferejohn and Fiorina (1984) already lamented the exclusive focus of this literature on American political institutions and the American political context. Since then, the study of the incumbency advantage in democracies other than the Unites States has made comparatively little progress. In the cases of Great Britain and Germany, for example, we still do not have any reliable estimates of the causal effects of incumbency, and the empirical findings in both cases are ambiguous in terms of signs and significance.

Our paper aims to help fill this lacuna. Following the recent turn towards quasi-experimental approaches in the US literature on the incumbency advantage (Lee, 2001; Lee, forthcoming), we employ a Regression Discontinuity Design (RDD) to identify the causal effects of party incumbency in British and German post-World War II elections. The RDD allows for causal in-

[1] A by no means exhaustive list of works includes Erikson (1972), Alford and Hibbing (1981), Born (1979), Collie (1981), Cover (1977), Cover and Mayhew (1977), Jacobson (1987), Krehbiel and Wright (1983), Mayhew (1974), Ansolabehere, Snyder and Stewart (2000), Ansolabehere and Snyder (2004), Ansolabehere, Snowberg and Snyder (2004), Cox and Katz (2002), Gelman and King (1990), Gelman and Huang (2004), King and Gelman (1991), Krashinsky and Milne (1993), Lee (forthcoming), Levitt and Wolfram (1997).

ference under much weaker assumptions than commonly employed identification strategies such as the Gelman and King regression model (Gelman and King, 1990) or more recent techniques such as instrumental variable estimation (Ansolabehere and Snyder, 2004).

Conceptualized in terms of the Rubin Causal model, the RDD framework exploits the randomized treatment assignment to incumbency that occurs when a district race is close. Since parties do not exert perfect control over their observed vote shares, incumbents that did barely win the previous election are virtually identical to non-incumbents that did barely lose, thus providing us with a naturally occurring counterfactual exploitable for causal inference. Whereas conventional estimations of incumbency advantage must rely on the assumption that they can statistically control for all observed and unobserved confounders, a condition that very rarely holds in practice, the RDD estimate relies on the assumption of local random assignment at the threshold of winning (Lee, 2001; Lee, forthcoming). If it holds, this ensures an unconfounded identification of the incumbency advantage in marginal districts, the subpopulation of interest for which incumbency advantage is likely to make a difference.

Since assignment to treatment is randomized at the threshold (i.e. orthogonal to potential outcomes), inferences at this point can be as credible as those obtained from a classical randomized experiment. Our estimates should be unaffected by the inclusion or exclusion of any pre-treatment controls; their inclusion should only result in lower sampling variability. In contrast, conventional identification strategies are largely based on observations located far from the threshold, where assignment to treatment is confounded by self-selection. Given that self-selection is usually based on unobserved characteristics, causal inference seems like a hopeless endeavor for such observations; the identification of unbiased causal estimates using conventional regression designs is very unlikely (LaLonde, 1986; Dehejia and Wahba, 1999). Our theories are simply too crude and our data are too limited to allow us any real confidence that we can "control" for all confounding factors. (Local) random assignment obviates this concern and enables us to draw reliable causal inferences.

In contrast to the mixed findings of previous studies, we will demonstrate below that in close races incumbency in both British and German federal elections has a robust and positive effect on a party's fortune. The magnitude of the party incumbency advantage is fairly similar in both countries; we estimate it to be around 1-2 percentage points of vote share or an approximate .20 increase in the probability of winning. While not nearly as large as the incumbency advantage observed in recent US House elections, these effects can nonetheless have an impact on races in marginal districts. Our results are also robust to a series of falsification tests.

The paper is structured as follows. The next section provides a brief review of existing studies on the incumbency effect in British and German elections, as well as the potential threats to validity these studies are faced with. The following two sections introduce our model and estimation strategy and discuss the conditions under which RDDs will (not) allow for reliable causal inferences. Section five presents our empirical findings. Sections six and seven present balance and robustness tests. Section eight concludes.

3 The Quest for Identification – Prior Estimates of the Incumbency Advantage

Scholars working on the incumbency advantage in non-American settings have generally followed the identification strategies employed by their American colleagues. It is therefore instructive to quickly review the American literature on the incumbency advantage before turning to the work done on Great Britain and Germany.

For U.S. House elections, the literature can be roughly divided into three stages. Early works relied on "sophomore surges" and "retirement slumps" (among others) to estimate the incumbency advantage (e.g. Erikson, 1972; Cover and Mayhew, 1977; Born, 1979; Payne, 1980).[2] In a path-breaking article, Gelman and King (1990) have shown that the identification strategies employed in these studies were all problematic. They introduced an improved estimator, often referred to as the 'Gelman and King' model.[3] Many scholars

[2]Gelman and King (1990) provide a review of this early literature.

[3]Gelman and King proposed a linear regression of the Democratic vote share in district j at time t on three independent variables: the lagged Democratic vote share in district j at time $t-1$, an incumbency status variable (-1 if there is a Republican incumbent running, 0 for open seats, and 1 for a Democrat incumbent running), and a dummy variable indicating which party won the seat in the last election. Under certain assumptions, the least squares estimate of the coefficient for the incumbency status variable provides an

directly adopted their approach or added twists and tweaks such as non-linear incumbency effects and selection bias correction procedures (Cox and Katz, 2002; Krashinsky and Milne, 1993; Levitt and Wolfram, 1997).

For about a decade, the Gelman and King approach has been regarded as "state-of-the-art" (Cox and Katz, 2002, p. 33). An increasing awareness of the strong assumptions on which it rests, however, has led scholars to look for alternatives such as quasi-experiments.[4] Quasi-experiments represent an important departure from conventional regression models, because they explicitly capitalize on exogenous or random variation in treatment assignment to identify the incumbency advantage instead of trying to "control" for potentially confounding covariates by including them in the regression equation. For example, in a highly original article, Ansolabehere et al. (2000) use the decennial redistricting as a natural experiment to contrast the incumbent's vote in the old parts of his district with his vote in the new parts of his district. This allows them to estimate the electoral benefits of each incumbent's "home-style."[5]

---

unbiased estimate of how much larger the incumbent's party vote share would have been with an incumbent defending the district as opposed to a non-incumbent.

[4]As Gelman and King themselves admit, their model assumes that the decision to run for reelection is orthogonal to the number of votes that the incumbent would get if he decided to run. To the extent that there is strategic exit (Cox and Katz, 2002; Jacobson and Dimock, 1994), this will lead to biased estimates. Another problem is the exclusion of variables such as incumbent quality. If incumbent quality is related both to treatment assignment at time $t-1$ and vote share at time t, which appears plausible, estimates will be biased upwards due to omitted variable bias.

[5]Note that the counterfactual implicit in their work is more narrow than usual, since it only contrasts the incumbent's vote share in new and old districts. Their model excludes other possible sources of incumbency advantage such as the cue contained in the label "incumbent" and also assumes that redistricting is not strategic (but see Cox and Katz,

Most relevant to our paper, economist David Lee (2001; forthcoming) employs a quasi-experimental research design to estimate the causal effects of U.S. House incumbency. Relying on a regression-discontinuity design (RDD), he uses close elections at time $t-1$, for which treatment assignment can be considered random, to estimate the causal effects of incumbency at time t. We will discuss Lee's model more carefully below.

Quasi-experimental approaches have not been applied to German or British federal elections yet. In fact, all of the existing work relies on some variant of Gelman and King's regression model or less sophisticated identification strategies. We now turn to a brief review of these studies.

3.1 Germany

Lancaster's (1998) study was the first to address the incumbency advantage in German post-World War II elections.[6] Lancaster regresses a binary dependent variable for winning or losing a district on candidate-specific independent variables (age, gender, and incumbency status) and finds a positive and statistically significant effect of incumbency on the probability of winning. But omitted factors such as candidate quality or district-specific factors related both to treatment assignment and potential outcomes are likely to

2002).

[6]Under Germany's mixed electoral system, each voter has two votes. The first or "district" vote is cast for candidates in SMDs under plurality voting. The second or "party list" vote determines the overall balance of seats in the Bundestag. Half the members of the Bundestag are elected in SMDs, the other half is elected through party lists. Bawn (1999) provides a concise review of the German electoral system.

lead to biased inferences.

Bawn (1999) examines a similar sample of German federal elections (1969-87) by looking at candidate vote gaps, defined as the district candidate's share of the first vote minus his party's share of the second vote in the same district. Bawn then considers positive candidate vote gaps as evidence for incumbency effects, which she estimates to be between .5 and 1 percentage points. Note, however, that Bawn's analysis cannot distinguish between incumbency advantage and any personal characteristics of incumbents. Indeed, Bawn (1999, p.494) herself notes that incumbency status "may also capture intangible elements . . . such as the degree to which the representative 'fits' the district in terms of personal ideology or style." Moreover, Bawn's approach assumes that party votes are exogenous to first votes. If a popular incumbent is able to not only garner a large share of the first vote, but also to raise the standing of his party in the district more generally, the gap between first and party vote shares will be a biased estimate of the incumbency advantage. Cox and Schoppa (2002) find evidence for such interaction effects across first and second tiers in mixed-member electoral systems, including Germany's.

3.2 Great Britain

Textbook accounts of British politics characterize British elections as purely party-or leader-centered and leave little room for any incumbency advantage. More recent work, well summarized in Gaines (1998), has challenged this traditional picture. Cain et al. (1984) have argued that district services provided by members of Parliament give rise to a noticeable incumbency advantage. But while their findings (based on a survey at the time of the 1979 general election as well as interviews with MPs and party agents) are quite suggestive, they fall short of a systematic assessment.

Gaines (1998) uses a regression model close to Gelman and King's but estimates separate equations for each major party and also includes a set of fixed effects. The results are quite mixed; depending on the exact specification, Gaines finds a small positive or sometimes even negative incumbency advantage. He concludes that "in Great Britain, incumbency advantage is either very small or else very elusive" (Gaines, 1998, p. 188)

Katz and King (1999), the most recent contribution, also applies the Gelman and King model to the British multi-party system, but impute data for partially contested districts and model the multiparty data according to a additive logistic t distribution. The authors find a small positive incumbency advantage in post-World War II British elections, with an effect size of .5 percentage points for the conservatives, 1 percentage point for Labour, and about 3 percentage points for the Alliance.

Given that the existing literature has largely adopted Gelman and King's (1990) identification strategy (or even more problematic ones), our knowledge about the incumbency advantage in British and German elections remains quite limited. The quasi-experimental research design that we rely on avoids many of the strong and implausible assumptions hitherto common in the literature. Of course, RDDs also rest on certain assumptions. But as we will show in the next section, these assumptions turn out to be weaker than the assumptions necessary for more traditional estimations strategies to work.

There is one important caveat, however. The causal effect estimated by our RDD is not directly comparable to earlier estimates of incumbency advantage. The existing literature focuses on the *legislator* incumbency advantage, while the RDD approach identifies an overall *party* incumbency advantage. As defined by Gelman and King (1990), the legislator incumbency advantage measures the difference between the proportion of the vote received by an incumbent legislator in his district, and the proportion of the vote received by the incumbent party in that district, *if the incumbent legislator does not run*. The party incumbency advantage identifies a related, though not identical counterfactual: the "electoral gain to being the incumbent party in a district, relative to *not* being the incumbent party" (Lee, forthcoming, p. 23). So, in a sense, party incumbency is a broader concept than legislative incumbency. It subsumes legislator incumbency but also contains the advantage that a party might enjoy simply from having won the district the last time, no matter whether the incumbent legislator runs again or not. Given the overarching role of political parties in both British and German politics compared to individual legislators, *party* incumbency seems to be the more relevant concept.

4 Regression-Discontinuity Designs

In our empirical analysis we rely on a regression-discontinuity design (RDD) to obtain estimates of the causal effect of party incumbency on a variety of outcomes such as vote share changes or the probability of winning the district in the next election. The RDD is a quasi-experimental framework that allows for identification of treatment effects in settings in which assignment to treatment changes discontinuously as a function of one or more underlying variables.

The earliest published example of RDDs dates back to Thistlethwaite and Campbell (1960). They examine the effect of scholarships on career outcomes by comparing students just above and below a threshold in tests scores that determined whether students were granted the award. The underlying idea is that in the close neighborhood of the threshold, assignment to treatment is as good as random. Accordingly, unlucky students that just missed the threshold are virtually identical to lucky ones who scored just above the cutoff value. The only difference between them is that the latter received the treatment while the former did not, thus providing us with a natural counterfactual for causal inference.

A data generating process characterized by assignment to treatment being solely based on exceeding a threshold on a predetermined covariate arises surprisingly often in empirical settings. Since the early work by Thistlethwaite and Campbell, RDDs have been frequently used in various disciplines such as medicine and public health, education, economics, and sociology.[7] Whereas in political science this identification strategy still seems somewhat under-exploited, the econometrics and program evaluation literature has recently seen a renewed interest in RDDs.[8] Most relevant for the purpose of this paper, David Lee (2001; forthcoming) for the first time applied a RDD framework to the estimation of incumbency advantage in U.S. House elections. Similar RDDs have been employed to study incumbency effects in India (Linden, 2004) and Ghana (Miguel and Zaidi, 2003) as well as split-party delegations in the U.S. Senate (Butler and Butler, 2005). Our empirical strategy largely follows the methodology employed in these papers.

5 The Model

Due to its quasi-experimental character, the RDD[9] is best formulated in the framework of the Rubin Causal Model and related methods that conceptualize causal inference in terms of potential outcomes under treatment and control (Rubin, 1974; Rubin, 1978; Rubin, 1990; Holland, 1986; Angrist and Krueger, 1999; Rosenbaum, 2002).

[7] Shadish, Cook, and Campbell (2002, p. 208) provide a long list of applications. Also see Trochim (1984) for further examples and an introduction to the RDD.

[8] For recent theoretical work on identification and estimation of RDDs see Hahn, Todd, and van der Klaauw (2001); Porter (2002); Battistin and Rettore (2002; 2003); Lee and Card (2004); and Lee (forthcoming). Recent empirical applications in economics include Angrist and Lavy (1999); Black (1999); Berk and de Leeuw (1999); Lee, Moretti, and Butler (2004); DiNardo and Lee (2004); Martorell (2004); and Matsudaira (2004).

[9] There are generally two types of RDDs: the *fuzzy* and the *sharp* design (Trochim, 1984). Here we only focus on the sharp design.

First, we need to introduce some notation. Let $TVS_{i,j,t}$ denote the actual (i.e. "true") vote share of party j in district i in an election at time t. For tractability, we assume that the same parties stage candidates in election t and $t-1$.[10] Assume further that the observed vote share party j receives is represented by $OVS_{i,j,t}$, which is the sum of two components:

$$OVS_{i,j,t} = TVS_{i,j,t} + \eta_{i,j,t} \tag{1}$$

Here $TVS_{i,j,t}$ reflects a systematic, or predictable component that is a function of the party's individual attributes or actions (such as the experience or likability of the party's candidate, campaigning efforts, etc.), and $\eta_{i,j,t}$ is a exogenous, random chance component (such as the weather on election day, traffic conditions, etc.) with mean zero and a continuous density (more on this point below).

Following Lee (forthcoming), we define "incumbency effect" as the overall causal impact of being the current incumbent party on the votes obtained in a district election. Thus, let $D_{i,j,t}$ be a binary indicator of treatment status that takes on the value of 1 if party j is the incumbent in district i at time t; and 0 otherwise. Under a first-past-the-post system, treatment status (i.e. whether a party becomes the incumbent or not) is determined by whether its observed vote share exceeds that of its strongest opponent in district i in the election at $t-1$. To compute this *margin of victory*, we rank parties in

[10] As we will explain below, this assumption is not a problem here since all parties generally stage candidates in all district in our sample.

each district by their observed vote shares in the election at $t-1$. Let W be the party with the largest observed vote share and $W-1$ be the party with the second largest observed vote. Then we can easily derive the *margin of victory* as:

$$MV_{i,j,t-1} = OVS_{i,j,t-1} - OVS_{i,W-1,t-1} \tag{2}$$

if a party is the winning party in a district and for every other party

$$MV_{i,j,t-1} = OVS_{i,j,t-1} - OVS_{i,W,t-1} \tag{3}$$

Note that by construction MV will be positive for winning and negative for loosing parties. The threshold, label it $\bar{MV}$, is zero. Once we define MV this way, we obtain a setting suitable for the RDD framework. Note that assignment to treatment is a deterministic function of whether a party's MV_{t-1} exceeds the threshold of $\bar{MV}_{t-1}$. We can write this as $D_t = 1[MV_{t-1} \geq \bar{MV}_{t-1}]$ where $[\cdot]$ is an indicator function that takes on the value of 1 if the condition within $[\cdot]$ is true and 0 otherwise.[11]

Finally, let Y_{1ij} denote the potential outcome party j would attain in district i if exposed to the treatment and let Y_{0ij} denote its potential outcome without the treatment in the same district. Our outcomes of interest are changes in vote share or the probability of winning in the election at time t.

Potential outcomes refer to possibly counterfactual events. The well known "fundamental problem of causal inference" (Holland, 1986) is that

[11]Note that, although possible in theory, there are no ties in our data.

for each unit i, we never observe both potential outcomes Y_{1ij} and Y_{0ij} simultaneously, but only the realized outcome $Y_{ij} = D_{ij} \cdot Y_{1ij} + (1 - D_{ij}) \cdot Y_{0ij}$. For example, given that a party is the incumbent, we never observe the outcome it would have attained had it not been the incumbent in the same district race. It is thus impossible to estimate individual causal effects $(Y_{1ij} - Y_{0ij})$. However, we can, under certain assumptions, estimate the average treatment effect $ATE = E[Y_1 - Y_0]$ or the average treatment effect for the treated $ATT = E[Y_1 - Y_0|D = 1]$ for a given population.

The strength of the RDD derives from the fact that we know the assignment mechanism. Under fairly weak assumptions, this allows for an identification of the ATE at the discontinuity of the covariate that determines treatment assignment. Such an estimate can be as credible as a causal inference drawn from a randomized experiment. More formally, when the support of the assignment variable, in our case MV_{t-1}, is continuous and $E[Y_0]$ is "smooth" at the threshold that determines assignment to treatment, in our case $\bar{MV}_{t-1}$, then the ATE is identified at the threshold:[12]

$$
\begin{aligned}
ATE &= E[Y_1 - Y_0|MV_{t-1} = \bar{MV}_{t-1}] && (4)\\
&= E[Y_1|MV_{t-1} = \bar{MV}_{t-1}] - \lim_{\varepsilon \to 0} E[Y_0|MV_{t-1} = \bar{MV}_{t-1} - \varepsilon] && (5)\\
&= E[Y|MV_{t-1} = \bar{MV}_{t-1}] - \lim_{\varepsilon \to 0} E[Y|MV_{t-1} = \bar{MV}_{t-1} - \varepsilon] && (6)
\end{aligned}
$$

[12]The proof that the RDD provides an unbiased estimate of the treatment at the discontinuity has been derived by several authors. For details see Goldberger (1972*a*; 1972*b*); Rubin (1977); Cappelleri (1991); Hahn, Todd, and van der Klauuw (2001); Lee (2001; forthcoming).

The identification assumption here is that $E[Y_0|MV_{t-1}]$ is smooth at $\bar{M}V_{t-1}$; there is no discontinuity at this threshold.[13] This assumption allows to attribute any discontinuity at the threshold to the treatment, because it implies that given a small enough ε the incumbents marginally above the threshold provide a valid counterfactual for those non-incumbents marginally below.[14] Note that one limitation here is that this ATE is in fact only a local ATE, because identification relies on marginal elections which may not be representative of all elections (more on this point below).

Why would equation 6 yield a quasi-experimental, unbiased causal estimate of the party incumbency effect, given the non-random selection to treatment? After all, MV_{t-1} is usually correlated with potential outcomes and the threshold ($\bar{M}V_{t-1}$) is known a priori to the parties. The reason why the RDD nonetheless provides an unbiased estimate is the assumption of local random assignment in the close neighborhood of the threshold.[15] Recall

[13] As usual in causal inference, for this result to hold we also need two additional assumptions: First, following Cox (1958), we need to assume "no interference between units," i.e. what Rubin (1978) refers to as the stable unit treatment value assumption (SUVTA). SUVTA implies that the treatment status of one unit does not affect potential outcomes for the other units (see also Rubin (1974) regarding this assumption in the RDD context). In our case this means that whether or not a party is the incumbent in one district has no impact on the outcomes in any other district. This assumption is non-testable, but if it is violated, causal inference becomes very difficult if not impossible (**?**). Second, we need to assume that treatment does not cause the assignment variable (Rubin, 1974). This is not a problem here because whether or not a party becomes the incumbent is determined after the election.

[14] As Shadish, Cook and Campbell (2002, p.237) put it: "even if the latter statistics are not transparent to the reader, a more common sense rationale should be: If participants with assignment scores at 50.05 [...] perform remarkably better on the outcome variable than participants at 49.95, surely, the .05 (sic) difference between them on the assignment variable is unlikely to account for that improvement."

[15] This logic is fully developed by Lee (forthcoming).

from equations $1-3$ that the margin of victory is a function of observed vote shares. Observed vote shares in turn consist of a systematic component ($TVS_{i,j,t}$) that parties can influence, but also a random component ($\eta_{i,j,t}$) over which parties exert no control. It can be proven that as long as the covariate that determines assignment to treatment includes such a random component with a continuous density, treatment status at the threshold is statistically randomized (Lee, forthcoming).

This assures that at the threshold, all observed and unobserved confounders determined prior to assignment will be orthogonal to treatment status (formally, we can write this conditional independence assumption as $Y_{1ij}, Y_{0ij} \perp D_{i,j,t} | MV_{i,j,t-1} = \bar{M}V_{i,j,t-1}$). Accordingly, as in a randomized experiment, our ATE estimate will not be confounded by omitted variables and should be unaffected by the inclusion of pre-treatment covariates, which would only result in lower sampling variability. This assumption of local conditional independence is testable at least for the observed confounders using common covariate balance tests known from randomized experiments or matching estimators. When the local randomization "worked," treatment and control group should yield similar distributions of baseline characteristics in the neighborhood of the threshold. We provide such balance tests in a separate section below.

It is important to briefly consider the condition under which the identification assumption of local random assignment breaks down. Local random assignment critically hinges on the presence of the random component ($\eta_{i,j,t}$).

We do not imply that each election has to be decided by this random component; in many non-close elections the random component will not be decisive. The key idea is that as elections become closer and closer, confounders no longer systematically affect assignment to treatment. In the limit, i.e. at the threshold, we should obtain conditional independence for all confounders. The plausibility of this assumption is a function of the degree to which parties are able to sort around the threshold (Lee, 2001; Lee, forthcoming). For example, if parties had perfect control over their observed vote shares or were at least able to perfectly predict them, they would never run if they knew they would loose. Alternatively, they would just invest enough effort to get exactly one more vote than the strongest district opponent. This would clearly violate our identifying assumption, as it would lead to a discontinuity in $E[Y_0]$ at the threshold. However, given the randomness inherent in election outcomes, such a scenario seems implausible. Just imagine the weather had been different on election day (Knack, 1994).

Before we proceed to estimation, a potential limitation of the RDD needs to be addressed. As noted above, our ATE estimate does not identify the average treatment effect for the entire population, but only for close elections. However, these marginal elections may not be representative of the whole population of elections. In fact, as we move further away from the threshold, local random assignment will cease to hold. We have no reason to believe that marginal and non-marginal districts are sufficiently similar. So unless we retreat to additional homogeneity assumptions, our incumbency effect estimate will not be applicable to the latter districts. Our data per se are only informative about the effect of D at the threshold.

However, when looking at the incumbency advantage, marginal districts are exactly the subpopulation of interest. It is only in these close elections that the incumbency advantage is likely to be decisive for the election outcome. Moreover, marginal districts are not uncommon in both our cases. For example, over 20 (40) percent of all races in an average German federal election are close, with the winner of each district leading by less than 5 (10) percentage points.[16] In the British data, about 14 (25) percent of all races in an average general election are close, with the winner of each district leading by less than 5 (10) percentage points.

6 Estimation Strategy

Our model suggests the following data generating process for the observed outcomes:

$$Y_{i,j,t} = f(Z_{i,j}, \theta) + h(MV_{i,j,t-1}, \delta) + D_{i,j}\beta + \varepsilon_{i,j} \tag{7}$$

where f is some function according to which Z, a vector of district level covariates (e.g. campaigns efforts, etc.), may affect votes in the current election with coefficients θ. h is some function that relates the margin of victory in the previous election to votes in the next election. Parties that

[16] Bawn (1999, 493) also provides evidence that winning gaps in German district races are fairly small.

did well in the last election are more likely to do well in the current election. Finally, D is our incumbent indicator and β is the central parameter of interest that identifies the party incumbency advantage.

At a first glance, equation 7 is just a conventional regression setup (like the Gelman and King Model). The fundamental problem that has plagued these conventional estimates of incumbency advantage is that there may be some unobserved Z that we cannot control for (like unobserved candidate characteristics, etc.), and since Z is likely to be correlated with MV and thus D estimates of β tend to be biased. The key advantage of the RDD is as follows: If our (fairly weak) assumptions hold and conditional on the parties' attributes and actions, there exist a random chance component with a continuous density as part of the assignment variable, then thanks to local random assignment our estimate of β will be un-confounded at the threshold $MV_{i,j,t-1} = \bar{MV}_{i,j,t-1} = 0$. Since treatment variation at the threshold is randomized, we do not need to control for additional covariates. Just like in a randomized experiment, inclusion of covariates should not affect β, apart from reducing sampling variability, because randomization ensures conditional independence.

As is well known, the real problem in drawing inferences from a RDD is getting the functional form of h right. Since D should only pick up the potential "jump" in the conditional expectation of $E[Y|MV_{i,j,t-1}]$ at the threshold, it is clear that a miss-specification of the functional form of h feeds into bias in β. Two main solutions to this problem have been proposed.[17] The first and most common method is to stick with a parametric model, but allow for a highly flexible functional functional of $E[Y|MV_{i,j,t-1}]$ by including higher order polynomials in MV plus all interactions with the treatment indicator into h.[18] Polynomials usually ensure a good fit of the functional form on both sides of the threshold and render it relatively insensitive to outliers. The second and more recent solution relies on non-parametric or semi-parametric estimators (Hahn, Todd and van der Klaauw, 2001; Porter, 2002) to identify the conditional expectation function at the threshold. No functional form assumption is required. Instead, one fits a semi- or non-parametric relationship to both sides of the threshold and then takes the difference at the discontinuity point. Porter (2002) recently derived a local polynomial estimator (based on a kernel regression) that achieves the optimal rate of convergence under a broad set of conditions. One problem is that there is no commonly accepted optimality criterion for the bandwidth choice of the kernel.[19] For ease of presentation, we display the parametric solution below. The non-parametric solution is used in the robustness section.

[17]Note that these two solutions require that the variable that determines assignment to treatment is continuous, as in our case. If this variable is discrete, a combination estimator can be used. See Lee and Card (2004) for details.

[18]Thus we estimate $Y = MV + MV^2 + MV^3 + MV^4 + D + D * MV + D * MV^2 + D * MV^3 + D * MV^4$. The coefficient on D identifies the ATE, because at the threshold MV equals zero.

[19]A problem here is the bias arising from the bad boundary behavior of the kernel; at the boundary the bias of the kernel converges to zero at a slower rate than at interior points. For details see Hahn, Todd, and van der Klauuw (2001); Porter (2002); and also Fan (1992) and Haerdle (1990).

7 Data

We use two data-sets in our analysis. For the German case, our data are taken from Caramani (2000). We originally planned to examine all federal elections to the Bundestag in the 1957-2002 period. But due to several rounds of redistricting, we had to exclude the 1957, 1976, 1980, and 2002 elections. This leaves us with 9 elections. There are about 248 districts per election up to 1990, and 328 districts since Unification. In our analysis, we also exclude the districts in East Germany due to lack of variation. For these districts, information on lagged vote shares or incumbency status only exists for two elections (1994 and 1998); there are also only 38 districts in which party incumbency switched between parties.[20]

In virtually all of the district races, the strongest two parties are the Social Democratic Party (SPD) and the Christian Democratic Party (CDU), although third parties obtain some share of the vote. Accordingly, we only focus on the incumbency advantage for these two parties. Note that partially contested districts are not a problem in Germany because the same parties generally run candidates in all districts.[21]

[20] All our findings are substantively identical if we include East German districts. Results are available upon request.

[21] The two big parties as well as the Free Democrats (FDP) have staged candidates in virtually all districts throughout our sample period. The Greens have only entered the political landscape in the early 80s. Fortunately, their pattern of entry has been rather "clean." In the election of 1983, which was the first elections in which the Greens ran district candidates at all, they already did so in all but four districts. In later elections they ran candidates in all but two districts. Exclusion of these few partially contested districts leaves all our results unaffected.

For the British case, we draw upon the data-set used in Katz and King (1999), which includes English constituency level results for the 1955-1992 period. The data-set is adjusted for redistricting and contains about 480-500 races per election. The parties of interest are the Labour Party and the Conservative Party (hereafter "Tories"). Smaller regional parties are excluded (for details see King and Katz (1999, 17)). Due to lack of variation, we provide no incumbency estimates for the Alliance.[22]

One issue that needs to be addressed for the British case is the problem of partially contested districts. As is well known, inferences from multiparty electoral data are problematic when parties strategically decide to run candidates in only some districts, because "different numbers of parties composing the 'other' category will generally have a large effect on a variable such as the percentage of votes for the governing party." (Katz and King (1999, 16)). One potential solution to this problem is to impute "expected" vote shares for all parties in partially contested districts (Katz and King, 1999; King, Honaker, Joseph and Scheve, 2001; Honaker, Katz and King, 2002). However, our focus is not to solve this problem here. Thus, in order to keep the analysis tractable and to avoid the necessity of making additional assumption for the imputation procedure, we opt for a simpler, yet very common solution. We exclude all those current and lagged district races in which the Alliance did not stage candidates. Note that this results in some informa-

[22]The Alliance only had 68 incumbents over the whole sample period. Given the low power of the RDD, this is not enough variation to obtain reliable estimates.

tion loss, but effectively eliminates any bias that may result from partially contested districts.[23] As Katz and King show, no bias exists for those districts in which all parties contest. Most previous studies of the incumbency advantage have made similar exclusions (Gelman and King, 1990; Levitt and Wolfram, 1997).

Our analysis is focused on the incumbency advantage at the level of the party in a district, regardless of the identity of the specific candidate a party stages. For each of the four parties of interest (SPD, CDU, Tories, and Labour), we compute the margin of victory for each district as described in equations $2-3$. We then estimate the causal effect of party incumbency by regressing our outcomes on a fourth order polynomial in the margin of victory in the previous election and all interactions with the incumbency dummy (so that we achieve a good fit of the conditional expectation of Y on both sides of the threshold). At this point, we include no additional covariates in the specification (we will add covariates in the robustness section). The causal effect of party incumbency is simply the "gap" in the conditional expectation of Y at the threshold, contrasting the situation when a party is and is not the incumbency in a particular district. In order to avoid strong assumptions about the error terms, we use a robust estimator for the variance/covariance matrix.

As our outcome variables, we employ two measures of a party's success in

[23] We only lose about 22 percent of the district races, mostly in early elections. After 1970, the Alliance ran candidates in almost all districts.

the next election: the party's vote share, and the probability that the party will win the race.

8 Findings

8.1 German Case

Table 1 shows our central findings for the German case. Party incumbency has a robust and sizeable causal effect on both a party's vote share and the probability of winning a district race. This incumbency effect is very similar for both the SDP and the CDU in terms of magnitude and statistical significance. On average, incumbency raises a party's vote share by about 1.5 percentage points; it also increases the predicted probability of winning a race by about 0.20. All effects are significant at least at the .10 level.

A graphical representation of the incumbency effect on vote share is provided in figure 1 for the CDU and figure 2 for the SPD. In these graphs, the vertical axis displays the respective party's vote share in the district election at time t. The horizontal axis is the margin of victory at time $t-1$ (the party's vote share minus that of its strongest opponent), with the dashed vertical line at zero marking the threshold. All observations to the right (left) of the dashed line represent incumbents (non-incumbents). The red curve plots the fitted values from our polynomial fit on both sides of the threshold. Finally, each data-point represents a local average of the outcome variable for intervals of the margin of victory variable (each interval is 0.05

wide).

Several features are apparent in the graphs. First, note that for both parties there is a positive relationship between the margin of victory and the election outcome. This of course is what we would expect to find. If a party wins (looses) by a larger margin at time $t-1$, it is more likely to receive a higher (lower) vote share at time t. Second, and most importantly, there is a noticeable discontinuity right at the threshold of zero. This represents the causal effect of party incumbency. For both parties, candidates that barely won the previous election are significantly more likely to obtain a higher vote share (or win) in the current election than those that barely lost. This effect is un-confounded, given that local random assignment at the threshold ensures that candidates just below and above the cutoff are likely to be similar in all respects expect their treatment status (see our balance tests below). If party incumbency had no causal effect, we would expect no such discontinuity at the threshold. Note that nowhere except at the threshold do we see a discontinues jump in the conditional expectation function; there exists a smooth, rather well-behaved relationship between the two variables that is well approximated by our multiplicative polynomial fit.[24]

8.2 British Case

Our central findings for the British case are displayed in table 2. Very similar to our results for Germany, we find a robust positive causal effect of party

[24]Similar graphs for the probability of winning are available upon request.

incumbency for both parties. The magnitude of the incumbency effect is slightly bigger for the Tories than for the Labour party, although not significantly so. In his regard our findings differ from Katz and King (1999), which finds significant differences between the two parties. Party incumbency is estimated to increase the vote share by about 1.8 percentage points for Labour, and by about 2.6 percentage points for the Tories. Similarly, party incumbency increases the probability of winning a race by about 0.18 for Labour and by about 0.23 for the Tories. All effects are significant at least at the .10 level.

Graphical representations of the incumbency effect on vote share are provided in figure 3 for the Tories and in figure 4 for Labour. The patterns are very similar to the figures shown for the German case. Again, there is a clear "jump" at the threshold of winning, while elsewhere the conditional expectation of the outcome variable is smooth and well-behaved.[25]

It needs to be emphasized that, as we indicated above, these estimates of the party incumbency effect are informative only for marginal districts. They cannot be extrapolated to non-marginal district without making additional assumptions about the homogeneity of close and non-close district elections. Yet, given both the number of marginal district races and the magnitudes of the party incumbency effects estimated above, the incumbency advantage in marginal districts is likely to have a substantial impact on the overall election outcome. Moreover, given the inherent self-selection to treatment

[25]Similar graphs for the probability of winning are available upon request.

in the rest of the data, there is little hope of estimating causal effects at a greater distance from the threshold.

9 Balance Tests

As we have explained above, our inferences rest on the assumption of local random assignment at the threshold. In this section, we test this assumption for the pre-treatment covariates that we have data for.[26] If, in the limit, there is randomized variation in treatment status, we expect the covariates that are determined prior to treatment assignment to be balanced in the close neighborhood of the threshold. Take turnout at time $t-1$, for example. At the threshold, there should be no systematic difference between districts that were barely won and districts that were barely lost. In other words, there should be no "jump" in the conditional expectation of pre-determined covariates at the threshold.

Table 3 provides evidence for the German case that is consistent with the assumption that treatment variation at the threshold is randomized — at least for the covariates that we do observe. For both parties, we do not find any significant differences at the threshold for vote share, turnout, and various (linear) combinations of the two. We also do not find any significant differences for state dummies except for one, which is what one may

[26] Unfortunately, we are currently left with a rather limited set of covariates. But hopefully we will be able to find more covariates. We plan to add balance tests for other covariates in future redrafts.

reasonably expect even in a randomized experiment because of sampling variability.[27]

Figures 5 and 6 present graphical representations of the balance tests for the lagged vote share variables. These graphs are identical to the ones shown in the previous section, except that this time we put the party's lagged vote share on the vertical axis. Since the vote share of a party is determined before incumbency status is assigned, incumbency should not have any effect on this variable, i.e. there should be no discontinuity at the threshold. Clearly, this is the case for both parties, lending confidence to our assumption of local random assignment regarding this important confounder.

Table 4 displays similar balance tests for the British case. Unfortunately, given our data constraints, we are currently restricted to tests of lagged vote share only.[28] Yet, at least with regard to this important confounder, the null of balance cannot be rejected at conventional levels for either party. Figures 7 and 8 confirm this graphically; again, we do not find any noticeable discontinuity at the threshold.

Overall, these balance tests lend confidence to the assumption that treatment status is indeed randomized at the threshold, although clearly more

[27]As is well known, in randomized experiments balance across the universe of observed and unobserved confounders obtains only in the limit as $N \longrightarrow \infty$. In any finite sample, one may end up with a bad draw. Also note the possibility of type I errors. At the .10 level, we expect one in ten significance tests to incorrectly reject the null of covariate balance, even if we had true balance in all ten tests.

[28]Balance is particular important regarding lagged vote share, because this confounder by itself impounds many important characteristics of a district race. Yet, more covariate tests would clearly be desirable here.

balance tests using additional predetermined covariates would be desirable.

10 Robustness

In this section, we subject our findings to various falsification tests. If our assumption of local random assignment at the threshold is valid, our incumbency estimates should be somewhat insensitive to the inclusion of predetermined covariates. Just like in a randomized experiment, their inclusion should only increase the precision of our incumbency estimates, because the covariates will soak up some of the variance from the error term.

Tables 5 and 6 demonstrate that this is the case for the German data. In fact, for both parties, the causal effect of party incumbency on vote share is remarkably robust across different specifications.[29] Since the findings in both tables are substantively similar, they can be discussed at the same time.

In both tables, column 1 presents the baseline estimates of the party incumbency effect based on our polynomial without any additional covariates. In column 2, we add each party's lagged vote share to the specification. It enters highly significant. More important, for both parties the magnitude of the incumbency effect remains almost identical, while precision is slightly increased. In column 3 we add lagged turnout and a squared term for lagged vote share; the incumbency estimates remain substantively unaffected. In column 4 we enter a full set of district level fixed effects.[30] In column 5 we

[29]Similar tables are available for the probability of winning upon request.

[30]Prior studies of incumbency advantage have used district or state fixed effects to cap-

add a full set of year effects[31] and in column 6 we add the full set of covariates. Again, the estimated incumbency advantage for both parties is remarkably robust, i.e. the estimate stays well within sampling variability of the baseline estimate while precision is increased. Finally, the estimated effects even stay rather stable when we control for pre-determined characteristics using a very different method. In column 7 we first regress each party's current vote share on the full set of covariates and district and year fixed effects, and then estimate the discontinuity gap using the residuals of the first stage regression as the response variable. Again, the estimates stays well within sampling variability of the baseline estimate. This is what we would expect if treatment is conditionally independent of all pre-determined covariates at the threshold. If the average of the predetermined covariates is continuous through the threshold, a linear function of those covariates should be smooth through the threshold as well (Lee, forthcoming).

Tables 7 and 8 display similar falsification tests for the British case.[32] For both the Tories and the Labour party, the incumbency effect is strikingly stable across different specifications. In each model, the point estimate stays well within the sampling variability of the baseline estimate, while we gain

ture the normal vote (Levitt and Wolfram, 1997; Ansolabehere and Snyder, 2004). Apart from the normal vote, district fixed effects will account for any unobserved heterogeneity across districts in all time-constant factors that may affect vote share.

[31]In the incumbency literature year fixed effects are often included to account for common trends such as partisan tides (national swings towards a party). Note that even with district and year fixed effects conventional (non-quasi-experimental) estimates may still be biased due to district specific, transient shocks that may affect vote shares. These should not, however, bias our RDD estimates given that local random assignment holds.

[32]Similar tables are available for the probability of winning upon request.

efficiency.[33]

Overall, these robustness tests greatly increase our confidence in our estimates and the assumption of conditional independence at the threshold underlying our model. Our results lead us to conclude that in both Germany and the UK, party incumbency has a clear causal effect on parties' fortunes in close district elections. While prior estimates of the incumbency advantage have been plagued by potential biases, our analysis shows that the party incumbency effect is real, even if the pervasive problems of non-random selection into treatment are accounted for.

11 Conclusion

It is well known that estimating causal effects from observational studies provides researches with a formidable challenge. There is an increasing awareness across all the social sciences that conventional regression models are ill-suited for this task, given the pervasive self-selection to treatment based on unobservables. As Sobel puts it in a recent review: "rarely is the state of knowledge in the social sciences adequate for a researcher to feel confident that he or she has measured all of the relevant covariates. The possibility of hidden bias (relevant unobserved covariates) is great." (Sobel, 2000, 649).

The literature on incumbency advantage in the U.S. House of Representatives has begun to recognize these shortcomings. Recent scholarship is

[33]Kernel regression estimates will be presented in a later version of this paper.

increasingly turning towards quasi-experimental frameworks that allow for more reliable causal estimates under a weaker set of assumptions. Our paper applies this idea to British and German elections. Following recent work in economics, we use a Regression Discontinuity Design that exploits the local random assignment to treatment that takes place in close district races. Based on the assumption that candidates do not exert perfect control over their vote shares, candidates that barely won or lost a race should be similar in their distribution of observed and unobserved confounders, thus enabling us to estimate the causal effects of party incumbency. In both British and German federal elections, we find that party incumbency has a significant positive impact on vote shares and the probability of winning. This stands in contrast to previous more ambiguous findings. Future research should rely on similar quasi-experimental research designs to discriminate between the individual sources of party incumbency advantage, for the cases considered here and others.

References

Alford, J. R. and Hibbing, J. R. (1981), 'Inreased incumbency advantage in the house', *Journal of Politics* **43**, 1042–1061.

Angrist, J. and Krueger, A. (1999), Amsterdam: Elsevier Science, chapter Empirical Strategies in Labor Economics.

Angrist, J. and Lavy, V. (1999), 'Using maimonides rule to estimate the effect of class size on scholastic achievement', *Quaterly Journal of Economics* **114**, 533–567.

Ansolabehere, S., Snowberg, E. C. and Snyder, J. M. (2004), Television and the incumbency advantage in u.s. elections, Technical report.

Ansolabehere, S. and Snyder, J. M. (2004), 'Using term limits to estimate incumbency advantages when officeholders retire strategically', *Legislative Studies Quarterly* **XXIX**, 487–515.

Ansolabehere, S., Snyder, J. M. and Stewart, C. (2000), 'Old voters, new voters, and the personal vote: Using redistricting to measure the incumbency advantage', *American Journal of Political Science* **44**, 17–34.

Battistin, E. and Rettore, E. (2002), 'Testing for programme effects in a regression discontinuity design with imperfect compliance', *Journal of the Royal Statistical Society* **Series A, 165**(1), 39–57.

Battistin, E. and Rettore, E. (2003), 'Another look at the regression discontinuity design', *Manuscript. Institute for Fiscal Studies, London.* .

Bawn, K. (1999), 'Voter responses to electoral complexity: Ticket splitting, rational voters and representation in the federal republic of germany', *British Journal of Political Science* **28**, 487–505.

Berk, R. and de Leeuw, J. (1999), 'An evaluation of californias inmate classification system using a generalized regression discontinuity design', *Journal of the American Statistical Association* **94**(448), 1045–1052.

Black, S. (1999), 'Do better schools matter? parental valuation of elementary education', *Quarterly Journal of Economics* **144**, 577–599.

Born, R. (1979), 'Generational replacement and the growth of incumbent reelection margins in the u.s. house', *American Political Science Review* **73**, 811–817.

Butler, D. M. and Butler, M. J. (2005), 'Splitting the difference: What explains split-party delegations in the us senate?', *Manuscript Stanford University, Department of Political Science* **January 21**.

Cain, B. E., Ferejohn, J. A. and Fiorina, M. P. (1984), 'The constituency service basis of the personal vote for u.s. representatives and british members of parliament', *American Political Science Review* **78**, 110–125.

Cappelleri, J. C. (1991), *Cutoff-based designs in comparison and combination with randomized clinical trials*, Unpublished doctoral disertation, Cornell University.

Caramani, D. (2000), *Elections in Western Europe since 1815: Electoral Results by Constituencies*, Macmillan.

Collie, M. P. (1981), 'Incumbency, electoral safety, and turnover in the house of representatives', *American Political Science Review* **75**, 119–131.

Cover, A. D. (1977), 'One good term deserves another: The advantages of incumbency in congressional elections', *American Journal of Political Science* **21**, 523–541.

Cover, A. D. and Mayhew, D. R. (1977), *Congress Reconsidered*, Praeger, chapter Congressional Dynamics and the Decline of Competitive Congressional Elections.

Cox, D. R. (1958), *Planning of Experiments*, New York: Wiley.

Cox, G. W. and Katz, J. N. (2002), *Elbridge Gerry's Salamander*, Cambridge University Press.

Cox, K. E. and Schoppa, L. J. (2002), 'Interaction effects in mixed-member electoral systems', *Comparative Political Studies* **35**, 1027–1053.

Dehejia, R. . and Wahba, S. (1999), 'Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs', *Journal of the American Statistical Association* **94**, 1053–1062.

DiNardo, J. and Lee, D. S. (2004), 'Economic impacts of new unionization on private sector employers: 1984-2001', *Quarterly Journal of Economics* **119**, 1383–1442.

Erikson, R. S. (1972), 'Malapportionment, gerrymandering, and party fortunes in congressional elections', *American Political Science Review* **66**, 1234–1255.

Fan, J. (1992), 'Design adaptive nonparametrix regression', *Journal of the American Statistical Associateion* **87**, 999–1004.

Gaines, B. J. (1998), 'The impersonal vote? constituency service and incumbency advantage in british elections, 1950-92', *Legislative Studies Quarterly* **23**, 167–195.

Gelman, A. and Huang, Z. (2004), Estimating incumbency advantage and its variation, as an example of a before-after study, Technical report.

Gelman, A. and King, G. (1990), 'Estimating incumbency advantage without bias', *American Journal of Political Science* **34**, 1142–64.

Goldberger, A. S. (1972*a*), 'Selection bias in evaluating treatment effects: Some formal illustrations', *Discussion Paper No. 123. Madison: University of Wisconsin, Institute for Research on Poverty* .

Goldberger, A. S. (1972*b*), 'Selection bias in evaluating treatment effects: The case of interaction', *Discussion Paper. Madison: University of Wisconsin, Institute for Research on Poverty* .

Haerdle, W. (1990), *Applied Nonparametric Regression*, Cambridge University Press, New York.

Hahn, J., Todd, P. and van der Klaauw, W. (2001), 'Identification and estimation of treatment effects with a regression-discontinuity design', *Econometrica* **69**, 201–209.

Holland, P. W. (1986), 'Statistics and causal inference', *Journal of the American Statistical Association* **81**(396), 945960.

Honaker, J., Katz, J. N. and King, G. (2002), 'A fast, easy, and efficient estimator for multiparty electoral data', *Political Analysis* **10**, 84–100.

Jacobson, G. C. (1987), 'The marginals never vanished: Incumbency and competition in elections to the u.s. house of representatives, 1952-82', *American Journal of Political Science* **31**, 126–141.

Jacobson, G. C. and Dimock, M. A. (1994), 'Checking out: The effects of bank overdrafts on the 1992 house elections', *American Journal of Political Science* **38**, 601–624.

Katz, J. and King, G. (1999), 'A statistical model for multiparty electoral data', *American Political Science Review* **93**(1), 15–32.

King, G. and Gelman, A. (1991), 'Systemic consequences of incumbency advantage in u.s. house elections', *American Journal of Political Science* **35**, 110–138.

King, G., Honaker, J., Joseph, A. and Scheve, K. (2001), 'Analyzing incomplete political science data: An alternative algorithm for multiple imputation', *American Political Science Review* **95**(1), 49–69.

Knack, S. (1994), 'Does rain help the republicans? theory and evidence on turnout and the vote', *Public Choice* **79**, 187–209.

Krashinsky, M. and Milne, W. J. (1993), 'The effects of incumbency in u.s. congressional elections, 1950-1988', *Legislative Studies Quarterly* **18**, 321–344.

Krehbiel, K. and Wright, J. R. (1983), 'The incumbency effect in congressional elections: A test of two explanations', *American Journal of Political Science* **27**, 140–157.

LaLonde, R. J. (1986), 'Evaluating the econometric evaluations of training programs with experimental data', *American Economic Review* **76**, 604–620.

Lancaster, T. D. (1998), *Stability and change in German elections: how electorates merge, converge, or collide*, Praeger, chapter Candidate Characteristics and Electoral Performance: A Long-Term Analysis of the German Bundestag, pp. 281–300.

Lee, D., Moretti, E. and Butler, M. J. (2004), 'Do voters affect or elect policies? evidence from the u.s. house', *Quarterly Journal of Economics* **119**(3), 807–859.

Lee, D. S. (2001), 'The electoral advantage to incumbency and voters' valuation of politicians' experience: A regression discontinuity analysis of elections to the u.s.', *NBER Working Paper No. 8441* **August**.

Lee, D. S. (forthcoming), 'Randomized experiments from non-random selection in u.s. house elections', *Journal of Econometrics* .

Lee, D. S. and Card, D. (2004), 'Regression discontinuity inference with specification error', *UNIVERSIY OF CALIFORNIA, BERKELEY WORKING PAPER NO. 74* **June**.

Levitt, S. D. and Wolfram, C. D. (1997), 'Decomposing the sources of incumbency advantage in the u.s. house', *Legislative Studies Quarterly* **XXII**, 45–60.

Linden, L. (2004), 'Are incumbents really advantaged? the preference for non-incumbents in indian national elections', *Columbia University Manuscript* **January**.

Martorell, F. (2004), 'Do graduation exams matter? a regression-discontinuity analysis of the impact of failing the exit exam on high school and post-high school outcomes', *UC Berkely Manuscript* **September**.

Matsudaira, J. D. (2004), 'Sinking or swimming? evaluating the impact of english immersion vs. bilingual education on student achievement', *University of Michigan manuscript* **October**.

Mayhew, D. R. (1974), 'Congressional elections: The case of the vanishing marginals', *Polity* **6**, 295–317.

Miguel, E. and Zaidi, F. (2003), 'Do politicians reward their supporters? public spending and incumbency advantage in ghana', *UC Berkeley Mimeo* .

Payne, J. L. (1980), 'The personal electoral advantage of house incumbents', *American Politics Quarterly* **8**, 375–398.

Porter, J. (2002), 'Asymptotic bias and optimal convergence rates for semi-parametric kernel estimators in the regression discontinuity model', *Harvard Institute of Economic Research, Discussion Paper Number 1989* **December**.

Rosenbaum, P. R. (2002), *Observational Studies*, New York: Springer-Verlag 2nd edition.

Rubin, D. B. (1974), 'Estimating causal effects of treatments in randomized and nonrandomized studies', *Journal of Educational Psychology* **66**, 688–701.

Rubin, D. B. (1977), 'Assignment to treatment group on the basis of a covariate', *Journal of Educational Statistics* **2**(1), 1–26.

Rubin, D. B. (1978), ‘Bayesian inference for causal effects: The role of randomization’, *Annals of Statistics* **6**(1), 34–58.

Rubin, D. B. (1990), ‘Comment: Neyman (1923) and causal inference in experiments and observational studies’, *Statistical Science* **5**(4), 472480.

Shadish, W., Cook, T. and Campbell, D. (2002), *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Houghton-Mifflin.

Sobel, M. E. (2000), ‘What do randomized studies of housing mobility reveal? causal inference in the face of interference’, *University of Columbia, Mansucript* .

Thistlethwaite, D. and Campbell, D. (1960), ‘Regression-discontinuity analysis: An alternative to the ex post facto experiment. journal of educational psychology’, *Journal of Educational Psychology* **51**, 309–17.

Trochim, W. (1984), *Research Design for Program Evaluation: The Regression Discontinuity Approach*, Sage Publications, Beverly Hills.

Figures

Figure 1: The Party Incumbency Effect for the CDU: Outcome Vote Share in Next Election

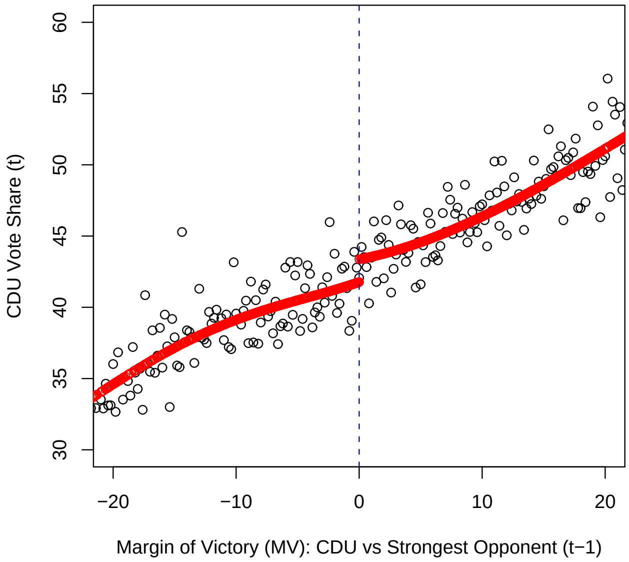


Figure 2: The Party Incumbency Effect for the SPD: Outcome Vote Share in Next Election

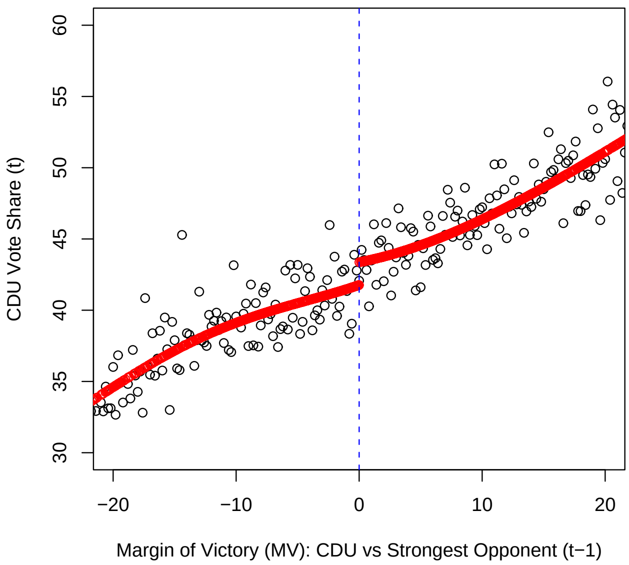


Figure 3: The Party Incumbency Effect for the Tories: Outcome Vote Share in Next Election

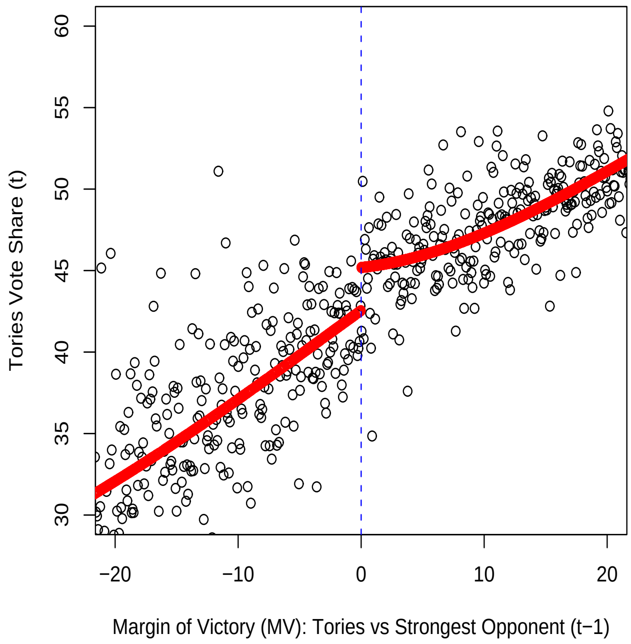


Figure 4: The Party Incumbency Effect for the Labour Party: Outcome Vote Share in Next Election

Figure 5: Validity Check for Local Random Assignment at Discontinuity - The Party Incumbency Effect for the CDU: Outcome Vote Share in Current Election

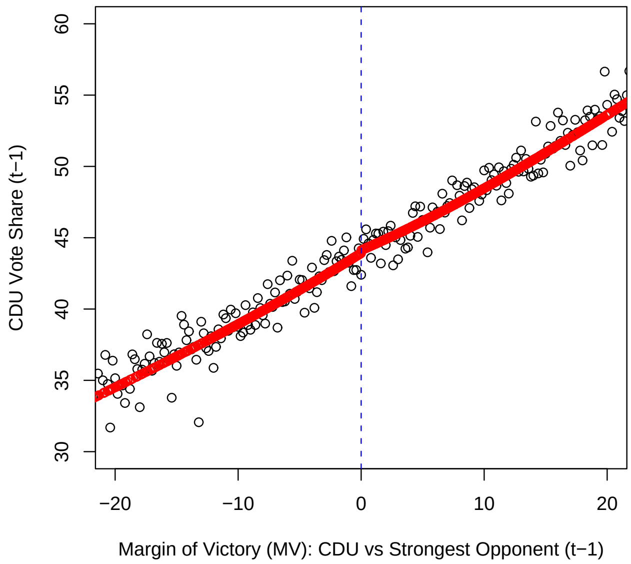


Figure 6: Validity Check for Local Random Assignment at Discontinuity - The Party Incumbency Effect for the SPD: Outcome Vote Share in Current Election

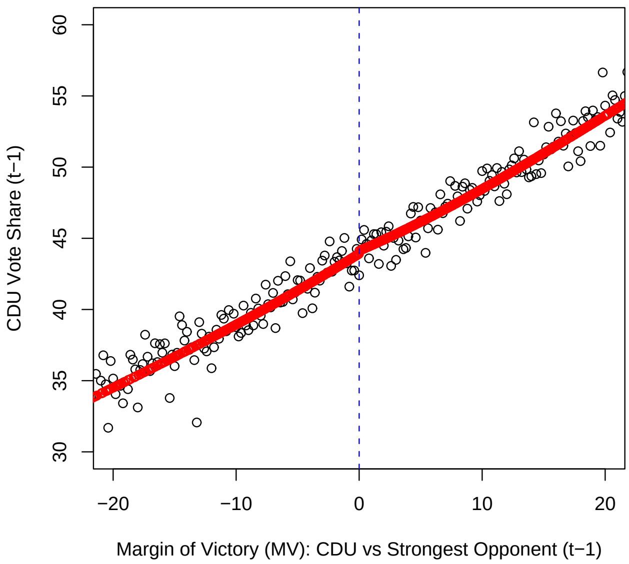


Figure 7: Validity Check for Local Random Assignment at Discontinuity - The Party Incumbency Effect for the Tories: Outcome Vote Share in Current Election

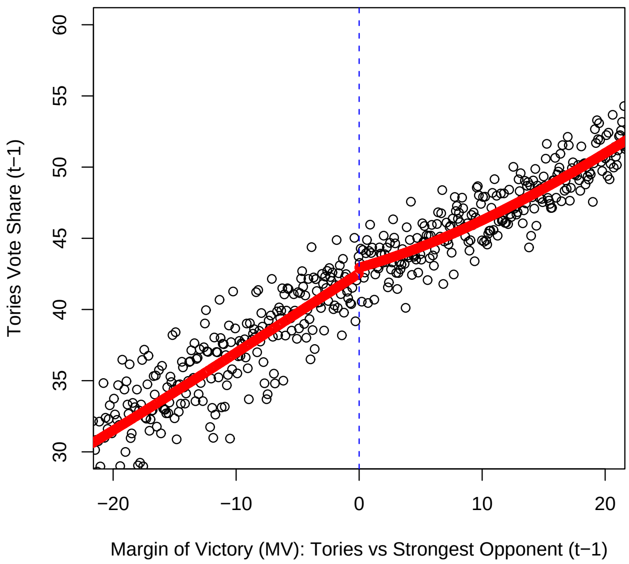
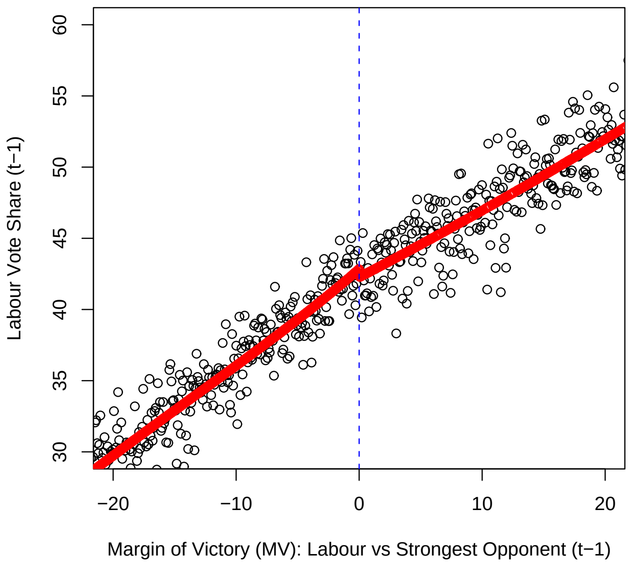


Figure 8: Validity Check for Local Random Assignment at Discontinuity - The Party Incumbency Effect for the Labour Party: Outcome Vote Share in Current Election



Tables

Table 1: The Causal Effect of Party Incumbency in Germany (1961-1998)

Response Variable	Incumbent	Non.Incumbent	ATE	.90 LB	.90 UB
Vote Share SPD	46.17	44.70	1.48	0.28	2.66
	(0.57)	(0.42)	(0.74)		
Vote Share CDU	43.37	41.77	1.60	0.27	2.94
	(0.48)	(0.65)	(0.79)		
PR(Win) SPD	0.77	0.57	0.20	0.00	0.38
	(0.09)	(0.07)	(0.11)		
PR(Win) CDU	0.42	0.21	0.21	0.05	0.37
	(0.07)	(0.07)	(0.10)		

All estimates at the threshold based on multiplicative fourth-order polynomial fit to both sides of the threshold. Robust standard errors in parenthesis. $N = 1.972$ in all estimations.

Table 2: The Causal Effect of Party Incumbency in the UK (1959-1992)

Response Variable	Incumbent	Non.Incumbent	ATE	.90 LB	.90 UB
Vote Share labour	43.16	41.30	1.86	0.47	3.23
	(0.60)	(0.58)	(0.81)		
Vote Share Tories	45.19	42.55	2.64	1.44	3.82
	(0.51)	(0.53)	(0.71)		
PR(Win) labour	0.63	0.45	0.18	0.02	0.33
	(0.06)	(0.07)	(0.09)		
PR(Win) tories	0.60	0.37	0.23	0.07	0.39
	(0.07)	(0.07)	(0.10)		

All estimates at the threshold based on multiplicative fourth-order polynomial fit to both sides of the threshold. Robust standard errors in parenthesis. $N = 3.470$ in all estimations.

Table 3: Random Assignment Checks for Predetermined Covariates in Germany

		CDU			SPD	
Covariate	Inc	Non-Inc	Diff	Inc	Non-Inc	Diff
$Voteshare$	44.13	43.92	0.21	43.90	44.21	−0.31
	(0.25)	(0.31)	(0.40)	(0.31)	(0.25)	(0.40)
$VShare^2$	1954.89	1933.27	21.61	1933.34	1958.71	−25.37
	(22.68)	(25.89)	(34.42)	(27.24)	(19.88)	(33.72)
$Turnout$	85.21	84.80	0.41	84.79	85.14	−0.34
	(0.45)	(0.63)	(0.77)	(0.63)	(0.44)	(0.77)
$Turnout^2$	7283.16	7209.08	74.08	7207.98	7270.53	−62.55
	(75.00)	(104.22)	(128.41)	(104.31)	(74.04)	(127.92)
$VShare$*$Turnout$	3757.36	3727.11	30.25	3728.53	3780.31	−51.79
	(36.27)	(44.24)	(57.21)	(45.66)	(33.89)	(56.86)
$VShare^2$*$Turnout^2$	1418.23	1398.25	19.98	1403.45	1443.74	−40.30
	(29.20)	(31.68)	(43.09)	(34.99)	(23.88)	(42.36)
State1	0.05	0.08	−0.03	0.08	0.05	0.03
	(0.02)	(0.03)	(0.04)	(0.03)	(0.02)	(0.04)
State2	0.03	0.04	−0.01	0.04	0.03	0.01
	(0.02)	(0.03)	(0.04)	(0.03)	(0.02)	(0.04)
State3	0.20	0.07	0.12	0.07	0.19	−0.11
	(0.04)	(0.04)	(0.06)	(0.04)	(0.04)	(0.06)
State4	0.00	0.01	−0.01	0.01	0.00	0.01
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
State5	0.23	0.32	−0.09	0.32	0.23	0.09
	(0.05)	(0.07)	(0.08)	(0.07)	(0.04)	(0.08)
State6	0.19	0.17	0.02	0.17	0.19	−0.02
	(0.04)	(0.05)	(0.06)	(0.05)	(0.04)	(0.06)
State7	0.04	0.08	−0.05	0.08	0.04	0.05
	(0.02)	(0.03)	(0.04)	(0.03)	(0.02)	(0.04)
State8	0.08	0.11	−0.03	0.11	0.07	0.04
	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)	(0.05)
State9	0.15	0.10	0.05	0.10	0.15	−0.06
	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)	(0.05)
State10	0.05	0.01	0.03	0.01	0.05	−0.03
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)

All estimates at the threshold based on multiplicative fourth-order polynomial fit to both sides of the threshold. Robust standard errors in parenthesis. $N = 1.972$ in all estimations. The unit of interaction of squared terms is scaled down (by 1000) for presentational purposes.

Table 4: Random Assignment Checks for Predetermined Covariates in the UK

	Labour			Tories		
Covariates	Inc	Non-Inc	Diff	Inc	Non-Inc	Diff
$VShare_{t-1}$	42.26	42.84	-0.57	42.94	42.53	0.41
	(0.29)	(0.27)	(0.40)	(0.27)	(0.29)	(0.39)
$VShare^2_{t-1}$	1796.97	1825.84	-28.86	1851.10	1815.59	35.51
	(25.30)	(20.48)	(32.55)	(24.41)	(22.80)	(33.40)

All estimates at the threshold based on multiplicative fourth-order polynomial fit to both sides of the threshold. Robust standard errors in parenthesis. $N = 3.470$ in all estimations. The unit of interaction of squared terms is scaled down (by 1000) for presentational purposes.

Table 5: Robustness Checks for CDU Incumbency Effect

Response Variable	VShare	VShare	VShare	VShare	VShare	VShare	Residuals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Incumbency	1.598	1.405	1.377	1.615	1.015	1.104	0.925
	(0.786)	(0.693)	(0.699)	(0.941)	(0.514)	(0.479)	(0.373)
$VShare_{t-1}$		0.927	0.506			0.162	
		(0.042)	(0.213)			(0.204)	
$VShare^2_{t-1}$			0.004			0.004	
			(0.002)			(0.002)	
$Turnout_{t-1}$			0.087			0.274	
			(0.022)			(0.049)	
District FEs				X	X	X	
Year FEs					X	X	

All estimations include a fourth-order polynomial in the margin of victory in previous election plus all interactions with incumbency (coefficients not shown here). Robust standard errors in parenthesis. $N = 1.972$ in all estimations. In Model 7 the dependent variable is the residuals from a regression of vote share in the current election on all predetermined covariates plus all FEs. See text for details.

Table 6: Robustness Checks for SPD Incumbency Effect

Response Variable	VShare	VShare	VShare	VShare	VShare	VShare	Residuals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Incumbency	1.475	1.615	1.607	1.755	0.908	0.926	0.795
	(0.735)	(0.693)	(0.69)	(0.864)	(0.522)	(0.468)	(0.375)
$VShare_{t-1}$		0.534	0.435			0.245	
		(0.032)	(0.122)			(0.12)	
$VShare^2_{t-1}$			0.003			0.004	
			(0.002)			(0.001)	
$Turnout_{t-1}$			−0.122			−0.162	
			(0.019)			(0.034)	
District FEs				X	X	X	
Year FEs					X	X	

All estimations include a fourth-order polynomial in the margin of victory in previous election plus all interactions with incumbency (coefficients not shown here). Robust standard errors in parenthesis. $N = 1.972$ in all estimations. In Model 7 the dependent variable is the residuals from a regression of vote share in the current election on all predetermined covariates plus all FEs. See text for details.

Table 7: Robustness Checks for Tories Incumbency Effect

Response Variable	VShare	VShare	VShare	VShare	VShare	VShare	Residuals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Incumbency	2.635	2.356	2.37	2.325	2.009	1.961	1.552
	(0.711)	(0.69)	(0.693)	(.709)	(.501)	(.471)	(0.387)
$VShare_{t-1}$		0.573	1.358			0.754	
		(0.034)	(0.153)			(0.155)	
$VShare^2_{t-1}$			−0.009			−0.002	
			(0.002)			(0.002)	
District FEs				X	X	X	
Year FEs					X	X	

All estimations include a fourth-order polynomial in the margin of victory in previous election plus all interactions with incumbency (coefficients not shown here). Robust standard errors in parenthesis. $N = 3.470$ in all estimations. In Model 7 the dependent variable is the residuals from a regression of vote share in the current election on all predetermined covariates plus all FEs. See text for details.

Table 8: Robustness Checks for Labour Incumbency Effect

Response Variable	VShare	VShare	VShare	VShare	VShare	VShare	Residuals
Model No	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Incumbency	1.861	1.979	2.079	2.486	1.668	1.624	1.261
	(0.81)	(0.824)	(0.798)	(1.002)	(0.612)	(0.595)	(0.444)
$VShare_{t-1}$		0.207	1.449			1.022	
		(0.034)	(0.069)			(0.071)	
$VShare_{t-1}^2$			(0.001)			(0.001)	
District FEs				X	X	X	
Year FEs					X	X	

All estimations include a fourth-order polynomial in the margin of victory in previous election plus all interactions with incumbency (coefficients not shown here). Robust standard errors in parenthesis. $N = 3.470$ in all estimations. In Model 7 the dependent variable is the residuals from a regression of vote share in the current election on all predetermined covariates plus all FEs. See text for details.