Tulane Economics Working Paper Series

A Replication of 'Do Voters Affect or Elect Policies? Evidence from the U.S. House' (The Quarterly Journal of Economics, 2004)

Patrick Button
Department of Economics
Tulane University
pbutton@tulane.edu

Working Paper 1518
September 2015

Abstract

I replicate Lee, Moretti, and Butler (2004) "Do Voters Affect or Elect Policies? Evidence from the US House." *Quarterly Journal of Economics*, 119(3), 807–59, using new advances in regression discontinuity design methodology. Specifically, I use local linear regression with optimal bandwidths (Imbens and Kalyanaraman, 2012) and I follow advice on polynomial modelling in Lee and Lemieux (2010). I also run McCrary (2008)'s density test as an additional robustness check to investigate sorting around the treatment cut-off. I investigate the sensitivity of estimates to polynomial order, bandwidth, and to the inclusion of covariates. The conclusion of Lee, Moretti, and Butler (2004) that voters "elect" rather than "affect" policies still holds under this more rigorous scrutiny.

A Replication of 'Do Voters Affect or Elect Policies? Evidence from the U.S. House' (The Quarterly Journal of Economics, 2004)

Patrick Button[1]

Assistant Professor

Department of Economics

Tulane University

September 5, 2015

ABSTRACT

I replicate Lee, Moretti, and Butler (2004) "Do Voters Affect or Elect Policies? Evidence from the US House." *Quarterly Journal of Economics*, 119(3), 807-59, using new advances in regression discontinuity design methodology. Specifically, I use local linear regression with optimal bandwidths (Imbens and Kalyanaraman, 2012) and I follow advice on polynomial modelling in Lee and Lemieux (2010). I also run McCrary (2008)'s density test as an additional robustness check to investigate sorting around the treatment cut-off. I investigate the sensitivity of estimates to polynomial order, bandwidth, and to the inclusion of covariates. The conclusion of Lee, Moretti, and Butler (2004) that voters "elect" rather than "affect" policies still holds under this more rigorous scrutiny.

---

[1] Email: pbutton@tulane.edu, Phone: 504-862-8353. I thank David Brownstone for many helpful comments and discussions.

1. INTRODUCTION

Lee, Moretti, and Butler (2004), henceforth referred to as LMB, is a highly cited[2] paper in public choice. LMB investigate two theories of how political competition affects policy outcomes. The first theory is that political competition leads to “partial policy convergence”: politicians respond to the political preferences of voters by moderating, to some extent, their policies. This is the theory advocated by Calvert (1985) and is a partial result of the median voter theory of Downs (1957). The second theory is “full policy divergence”: politicians cannot credibly promise to moderate policies, and hence simply adopt their preferred policies after election. Alesina (1988) developed this theory using optimal strategy under a finite game. Under “full policy divergence”, voters merely “elect” policies (in the language of LMB), rather than in “partial policy convergence” where voters are able to “affect" policies by encouraging moderation of political platforms towards that of the median voter.

To determine which theory is more applicable to U. S. House of Representatives elections, LMB employ a regression discontinuity (RD) design. The election structure provides a quasi-experiment whereby districts with extremely close election results are “as good as randomly assigned” either Democrat or Republican representatives. This quasi-random assignment leads to an exogenous electoral benefit to that party in that district in the next election due to the incumbency effect. If the incumbency effect then shifts the preferences of the median voter, and then politicians respond to this exogenous shift in the median voter’s preferences in the next election, then the “partial policy convergence” theory is appropriate (voters “affect” policies). If politicians do not change their policies in response to this exogenous shift in the median voter’s preferences, the “full policy divergence” theory is more applicable (voters “elect” policies).

[2] 470 citations on Google Scholar as of August 30, 2015.

LMB's research to determine which theory is more applicable is important because "partial policy convergence", through responses to the median voter, is an assumption behind many public choice models that predict how goods and services are allocated through a political system. Given the importance of LMB's findings, it is important to ensure that they are robust to recent advances in RD design methodology. I re-do the analysis in LMB using local linear regression and the optimal bandwidth suggested by Imbens and Kalyanaraman (2012), which is considered current best practice (Lee and Lemieux, 2010). I also conduct a more thorough analysis using parametric methods, following advice in Lee and Lemieux (2010) and others. I also run McCrary (2008)'s density test, which is a now standard way to investigate the assumption of quasi-random assignment in the RD design. My replication results are similar, and thus the conclusions in LMB hold up to this additional scrutiny.

2. REPLICATING LMB's MAIN RESULTS

I start by replicating LMB's main results. I relegate my replication of their robustness check where they use different roll-call voting measures (or liberalness) measures to Appendix 1 and I relegate the replication of their heterogeneity analysis to Appendix 2. Before presenting my replication of LMB's main results I summarize the data sources and LMB's methodology.

DATA

The election data used by LMB is U.S. House of Representatives election results, at the district level, from 1946 to 1996[3]. For each election in each district, LMB calculate the Democrat share of the two-party vote. This is calculated as the number of votes for the Democrat divided but the number of votes for both the Democrat and the Republican. Thus votes for third party candidates are ignored. These election results are merged with district-level demographic and income variables (see Table 9) from the U.S. Census and

[3] This data is available on Enrico Moretti's website: http://eml.berkeley.edu//~moretti/data3.html (last accessed August 27, 2015)

with various roll call voting scores (see Appendix 1), the main of which is Americans for Democratic Action (ADA) scores. All of the main estimations in this replication and in LMB use these ADA scores[4]. More information on these data, and their sources, is available in LMB.

INTUITIVE EXPLANATION OF THEIR THEORETICAL MODEL

Here I summarize the theoretical model that LMB construct, and then test empirically. A far more complete summary of this model is in the original LMB paper. Here I present a more intuitive explanation. There are two time periods in LMB's model: time t and time t+1. Time t starts with the election at time t, and includes the following legislative session at time t. Time t+1 starts with the next election, at time t+1, and includes the following legislative session at time t+1. Figure 1 shows this diagrammatically.

LMB use a sample of "close" elections that were narrowly won or lost in the time t election. In these close elections it is "as good as random" if the Democrat or Republican candidate won the election. There are three components that need to be estimated in order for them to separate the "elect" component from the "affect" component. These is the total effect (which LMB label γ), the effect of party affiliation on roll-call voting (π_1), and the incumbency effect ($P^D_{t+1} - P^R_{t+1}$). The elect component is calculated as $\pi_1(P^D_{t+1} - P^R_{t+1})$ and the affect component is what is left over of the total effect, or $\gamma - \pi_1(P^D_{t+1} - P^R_{t+1})$. I present an intuitive explanation of what these factors are and how they are calculated.

Comparing the average voting records of Democrats and Republicans immediately after (time t) their "as good as random" elections (also time t) gives the estimate of the effect of party affiliation on roll-call voting. This estimate is labelled π_1 in their model and in Figure 1. This captures how different Democrats and Republicans are politically, on average.

[4] Appendix 1 presents alternative results using different roll call voting scores or voting measures as a robustness check. The results hold regardless of the measure used.

The incumbency effect ($P^D_{t+1} – P^R_{t+1}$) is calculated by looking at election results in districts at time t+1 which had narrowly decided elections at time t. In these districts, it is "as good as random" if the incumbent is a Democrat or a Republican.

Given that representatives were "as good as randomly assigned" in some districts at time t then looking at roll-call voting records in these same districts at time t+1 gives the total effect, which is labelled as γ in Figure 1. This total effect captures the change in roll-call voting in that district if a Democrat was "as good as randomly assigned" in the election prior to the most recent election. Part of the total effect, γ, will be due to the fact that if a Democrat was "as good as randomly assigned" at time t, there is more likely to be a Democrat in office at t+1 due to the incumbency effect ($P^D_{t+1} – P^R_{t+1}$). And since Democrats vote differently than Republicans (π_1), this partially explains the total change in roll-call voting (γ). This is what LMB deem the "elect" component. It is calculated as $\pi_1(P^D_{t+1} – P^R_{t+1})$. Anything left over of the total effect, after accounting for the elect component, is what LMB call the "affect" component. Since the "as good as randomly assigned" incumbency effect is an exogenous shock to the support of one party or another at time t+1 in the district, candidates may react to this exogenous shock by changing their voting pattern. This is reflected in the "affect" component. Thus to test if there is partial policy convergence (voters "affect" policies) or full policy divergence (voters only "elect" policies), LMB test if the "affect" component is positive and statistically significant. The "affect" component will be positive and statistically significant if the total effect is not entirely explained by the "elect" component.

LMB's METHODOLOGY

LMB's main estimates in their tables are simple differences using a "close" elections sample, where the means of barely Democrat districts (50-52% Democrat share of the two-party vote) are compared to the means of barely Republican districts (48-50%). They also graphically present their estimates of the three components that make up the "affect" and "elect" estimates: the total effect (γ), the effect of party affiliation on roll call voting record (π_1), and the incumbency effect ($P^D_{t+1} - P^R_{t+1}$). Their figures are created by averaging all the data into one percentage point bins of the Democratic vote share, and plotting these 100 points. This presents a non-parametric picture of how out the outcome variable varies with the democratic vote share. Over top of this, they place the following polynomial curve

$$Y = \alpha + \tau D + \sum_{n=1}^{4} \beta_n (X - 0.5)^n + \sum_{m=1}^{4} \beta_n D (X - 0.5)^m + \varepsilon \qquad [1]$$

where X is the democratic vote share, Y is the outcome variable, D is an indicator variable for the democrat winning (i.e. $X \geq 0.5$). This regression is a fourth order polynomial on either side of the treatment cut-off (0.5 Democrat vote share), plus a jump at this cut-off (τ). LMB then add 95% confidence intervals.

REPLICATION RESULTS

Table I presents my replication of the main results. My estimates are in black while LMB's estimates are below those in red italics. To calculate the standard error for the affect and elect components, I had to combine my estimates of each of the three components. To do this, I used a seemingly unrelated regression (SUR) model to estimate the required covariances. It is unclear, based on LMB and their earlier working paper (Lee, Moretti, and Butler, 2001), how they estimated their standard errors for the affect and elect estimates.

However, my replication estimates are very similar to those in LMB. For the total effect, γ, LMB estimate 21.2 (standard error of 1.9) while I estimate 21.28 (1.95). For the effect of party affiliation on roll-call voting, π_1, LMB estimate 47.6 (1.3) and I estimate 47.71 (1.36). For the incumbency effect, ($P^D_{t+1} - P^R_{t+1}$), LMB estimate 0.48 (0.02) and I estimate 0.4843 (0.0289). Since our estimates of these three components are similar, our elect and effect component estimates are also similar. For the elect component, LMB estimate 22.84 (2.2) and I estimate 23.11 (1.50). For the affect component, LMB estimate -1.64 (2.0) and I estimate -1.82 (1.11). Thus, we both find that the affect component is not statistically significantly different from zero. This indicates that the "Full Divergence" theory (voters "elect" policies) best fits this data as US House of Representative candidates do not change their roll call voting behavior in response to exogenous changes in the median voter.

Figures 2, 3, and 4 show my replications of their figures. Instead of using a fourth order polynomial, I follow Ferreira and Gyourko (2009) and others and use a third degree polynomial[5], although this leads to nearly identical figures.

3. REPLICATION FOLLOW REGRESSION DISCONTINUITY DESIGN BEST PRACTICES

The RD design literature has improved significantly since publication of LMB. LMB used a difference in means estimate for districts within two percentage points on either side of the cut-off. This is sub-optimal based on the recent literature. I discuss here the current best practices, starting with the preferred non-parametric method: local linear regression, and then I discuss complementary parametric methods.

NON-PARAMETRIC METHODS

Recent RD design research argues that the best estimation method is local linear regression, using a triangular kernel. A local linear regression is of the form

$$Y = \alpha_l + \beta_l(X - c) + \varepsilon, \qquad \text{where } \mathrm{c} - \mathrm{h} \le \mathrm{X} < c \qquad [2a]$$

$$Y = \alpha_r + \beta_r(X - c) + \varepsilon, \qquad \text{where } \mathrm{c} \le \mathrm{X} \le \mathrm{c} + \mathrm{h} \qquad [2b]$$

where the subscript l refers to the fact that the regression fits the left side of the cut-off and similarly the subscript r refers to the right side. h is the bandwidth of data that are used. Pooling these results regressions gives

[5] Lee and Lemieux (2010) suggest that a low-order polynomial is used to fit the data in figures such as those in this paper.

$$Y = \alpha_l + \tau D + \beta_l(X - c) + (\beta_r - \beta_l)D(X - c) + \varepsilon, \qquad \text{where } c - h \leq X \leq c + h \qquad [3]$$

where D is an indicator equal to one if X ≥ c, zero otherwise, and τ is the estimated treatment effect. Fan and Gijbels (1996) argue that a local linear regression, with a triangular kernel, is optimal for fitting a boundary point.[6] They also argue that local linear regression has better bias properties than other techniques, and Porter (1993) shows that it is rate optimal. The triangular kernel is intuitively appealing over a rectangular kernel since the triangular kernel weights observations more heavily the closer they are to the discontinuity. This makes sense, as observations are more informative the closer they are to the cut-off. However, Lee and Lemieux (2010) argue that results do not differ much based on kernel type. For this reason, Lee and Lemieux (2010) and Imbens and Lemieux (2007) suggest that a rectangular kernel should be used, as it is easier to use and interpret. In this replication I use both types of kernels, which ends up not affecting the results.

Crucially, a bandwidth, h, must also be chosen. There is a tradeoff as using data farther from the cut-off increases bias, but this additional data increases precision. If the goal, as in most estimators, is to minimize the mean squared error, then a bandwidth selected based on Imbens and Kalyanaraman (2012), henceforth called the IK bandwidth, is ideal. The estimation of a discontinuity in an RD design context requires estimating two boundary points (one on either side of the cut-off). Imbens and Kalyanaraman (2012) develop a method to select an optimal bandwidth for local linear regression with a triangular kernel. This method minimizes the mean squared error of the estimate at the boundary points. It is intuitively appealing as one is only interested in fitting these two boundary points. There is no reason to select a bandwidth that minimizes the mean squared error, or any other criteria, over the full support of

[6] In the RD design case, there are two boundary points: one barely on the left side of the cut-off, and one barely on the right side.

the data. Conventional cross-validation methods are less desirable for this reason. Hence, the IK bandwidth is considered the standard.

PARAMETRIC METHODS

Lee and Lemieux (2010) emphasize that non-parametric methods are not a substitute for parametric methods, even though non-parametric methods have more support in the literature. Lee and Lemieux (2010) argue that they are complements, and both should be used to assess the robustness of results to changes in specification. Since RD design estimates can be biased under an incorrect functional form, it is important to use parametric methods as a robustness check to the non-parametric results. If the two methods lead to different results, then this casts some doubts on the estimates as they may be simply a function of a particular specification.

Polynomials are the most common parametric specification. They are of the form

$$Y = \alpha + \tau D + \sum_{n=1}^{p} \beta_n (X - c)^n + \varepsilon, \qquad \text{where } \mathrm{c - h \leq X \leq c + h} \tag{4}$$

where p is the polynomial order.

The difficulty with polynomials is choosing the optimal order, which entails a trade-off between bias and variance. This tradeoff is discussed briefly here, with a more thorough discussion in Button (2015). The optimal order of the polynomial depends on the bandwidth and the curvature of the underlying function, f(x). A larger bandwidth or a curvier function require a larger polynomial. Smaller bandwidths or flatter functions are better served with lower order polynomials. A polynomial of an order too large will over-fit the data, leading to a loss of precision, while a polynomial order too low will cause bias[7].

In addition to choosing the optimal polynomial order, a bandwidth must also be chosen. In practice, the bandwidth is selected first, and then the polynomial order is chosen, although it is in reality a joint selection problem. As for which bandwidth to use, the IK bandwidth is again a possibility here. But since the IK bandwidth is specifically for local linear regression, it may not be ideal. Lee and Lemieux (2010) discuss using cross-validation methods to select a bandwidth, which I employ and discuss with my results.

To investigate the sensitivity of the polynomial estimates to the bandwidth, Lee and Lemieux (2010) suggest using a range of bandwidths. For each bandwidth, they suggest running regressions with polynomials of various degrees (they used 0^{th} to 4^{th}). To assess how well the polynomial order fits the data, given the bandwidth, Lee and Lemieux (2010) suggest typical pretesting, such as using the Akaike Information Criterion (AIC). I create tables of regression estimates following this advice.

ESTIMATION OF MAIN RESULTS USING NON-PARAMETRIC METHODS

I start by presenting results from the best methodology first, which is local linear regression with the IK bandwidth. Table 2 shows estimates for the three parts that make up the affect and elect estimates. I present results using both the preferred triangular kernel, but also a rectangular kernel. As expected, kernel choice does not matter much. My results are similar to LMB's results.

[7] Another suggestion made by Lee and Lemieux (2010) is to use a separate polynomial to fit each side of the cut-off. Using one polynomial, rather than one polynomial for each side of the cut-off, implicitly adds the restriction that the slope of the underlying function is the same on both side of the discontinuity. If this assumption does not hold, then this will lead to bias. However, if it does hold, then estimating a separate polynomial for each side of the cut-off would reduce precision since the extra parameters are not useful. This is discussed in more detail by Button (2015). In this paper I follow Lee and Lemieux (2010) more closely and use only one polynomial for the entire span of the data. However, Button (2015) shows that the incumbency effect estimates in LMB and Lee (2008) are robust to using various polynomial models that use polynomials of different orders on either side of the cut-off.

Since the Imbens and Kalyanaraman (2012) data-dependent bandwidth selection algorithm selects different bandwidths for the three estimates required to separate the affect and elect components, it is difficult to choose a particular bandwidth to combine these estimates into the affect and elect estimates. Combining these three estimates under this methodology is also computationally difficult[8]. However, using the results as-is in Table 2 allows point estimates to be calculated without standard errors. For the triangular kernel, the point estimate of the affect component is -0.47 and the point estimate for the elect component is 21.2. For the rectangular kernel, these point estimates are 0.77 and 20.4 respectively. These point estimates are again similar LMB's. It is possible to combine the three components and generate estimates of the affect and elect components, with standard errors, using a rectangular kernel rather than the preferred triangular kernel. This is presented in Figure 5 and is discussed later.

Sensitivity to Bandwidth

Deviations from the IK bandwidth are suboptimal as any other bandwidth does not minimize the mean squared error. However, it is important to investigate if the estimation results are simply a function of a specified bandwidth. McCrary (2008) and Nichols (2011) suggest re-estimating the local linear results with half and double the IK bandwidth; however, there is no rigorous standard that suggests the range of bandwidths to use to determine robustness.

Table 3 shows local linear regression estimates of the three components, using a triangular kernel, for bandwidths ranging from 10% to 500% of the optimal IK bandwidths. As expected the estimates with wider bandwidths are generally higher and have a smaller standard error. This reflects the fact that using more data to estimate the discontinuity increases precision, but increases bias. The results are opposite

[8] The standard approach to estimate a RD design using a local linear regression with the IK bandwidth and a triangular kernel is to use the "rd" program created by Austin Nichols (see Nichols, 2007). Since this creates estimates individually, it is not possible to combine them using this program.

for the smaller bandwidths. These extremely large changes in the bandwidth do not have a significant impact on the estimates.

Following Card, Dobkin and Maestas (2009), Figure 5 plots local linear regression estimates, using a rectangular kernel, for various bandwidths. The dotted red lines show the local linear estimates using a triangular kernel and the IK bandwidth (calculated from the three components presented in Table 2). The estimates for the affect component are similar for all bandwidths, but they become negatively statistically significant from about 0.2 to 0.45[9]. This range is outside the optimal bandwidth, so these estimates are not preferred. The optimal bandwidth, although unknown, lies somewhere between the component bandwidths presented in Table 2 for the rectangular kernel, which range from about 0.05 to 0.15. Over this range the affect component is not statistically significant.

ESTIMATION OF MAIN RESULTS USING PARAMETRIC METHODS

While the literature argues that non-parametric methods are preferred, Lee and Lemieux (2010) recommend using parametric methods as a robustness check to ensure that the results are not sensitive to the specification chosen. For this reason, I present results using various polynomial models. Table 4 presents estimates for the total effect (γ) using a range of bandwidths and polynomials of orders zero to four, mirroring similar tables presented in Lee and Lemieux (2010). In addition, I present the optimal polynomial order according to both the AIC and BIC for each bandwidth.

[9] However, even if we were to put weight on these negative and statistically significant estimates of the affect component, this wouldn't really change LMB's conclusion. A negative affect estimate means that politicians react in the opposite way as expected. So while an exogenous boost to the Democratic support in a district would suggest that candidates vote more left-wing (higher ADA scores) in response, the negative estimate suggest a counter-intuitive movement to the right instead (lower ADA scores). The partial policy convergence theory suggests that candidates react to the median voter by following her preferences, which would imply a positive estimate of the affect component. So even a statistically significant negative estimate of the affect component helps disprove this theory.

The estimates in Table 4 range from 9.23 to 37.59, with almost all estimates being statistically significant. While this is a wide range of estimates, some of the estimates are more realistic than others. Higher polynomial orders are necessary as the bandwidth increases to reduce bias, while for small bandwidths, there is less scope for bias, so higher polynomial orders can over-fit and lead to a reduction in precision. The AIC and BIC tests generally indicate that higher polynomial orders are required for larger bandwidths. If we look just at the models preferred by the BIC, the estimates only range from 16.02 to 23.97, with all estimates statistically significant. LMB's estimates and my local linear regression estimates fall in the middle of this range. According to cross-validation bandwidth selection tests, following Lee and Lemieux (2010), the optimal bandwidth, h, is 0.0537. Since the estimates do not vary much by bandwidth for the more realistic models, the fact that the optimal bandwidth is 0.0537 is less crucial.

Table 5 presents an identical analysis but for the effect of party affiliation on roll-call voting (π_1). These are much more stable, with estimates ranging from 39.40 to 56.90 (or to 47.30 to 50.14 for the BIC-preferred models), all statistically significant. These estimates are again similar to their corresponding LMB and local linear regressions. The optimal bandwidth selected by cross-validation is 0.0612, but again the estimates under preferred models do not vary much by bandwidth.

Table 6 presents the incumbency effect estimates ($P^D_{t+1} - P^R_{t+1}$). These vary the most, with a range of -0.3727 (but statistically insignificant) to 0.8168. While the huge range is worrisome, the outlier estimates are from less preferred models. The preferred estimates under BIC are much more stable, ranging from 0.3905 to 0.5307, all statistically significant. The optimal bandwidth selected by cross-validation is 0.0425, but again the estimates under preferred models do not vary much by bandwidth. In addition, Button (2015) estimates how the incumbency effect estimate varies across even more plausible models and finds that the estimate is also fairly stable.

Across all three tables, the parametric estimates are fairly stable over bandwidth and polynomial order and match the local linear estimates and LMB's estimates. But what is more important is to look at the combined "affect" and "elect" component estimates. Table 7 presents these estimates along with the estimates of the three components for each bandwidth under the model preferred by the BIC. This is shown visually in Figure 6. The estimate of the affect component varies from -3.71 to 2.97. Out of the 11 estimates presented, seven are statistically insignificant and four are negative and statistically significant. The occasional negative and statistically significant estimate somewhat mirrors Figure 5 where the affect component estimates were negative and statistically significant for the bandwidth range of 0.2 to 0.45. The negative and statistically significant estimates do not occur for the preferred bandwidth of about 0.05[10] and as discussed earlier (see footnote 9) negative statistically significant affect estimates do not change LMB's conclusion.

INVESTIGATING THE VALIDITY OF THE QUASI-EXPERIMENTAL DESIGN

It is important in research that uses an RD design to investigate the assumption of quasi-random assignment. I conduct several tests to investigate this. First, I use the McCrary (2008) density test to investigate possible sorting at the cut-off that could suggest that the quasi-random assumption is violated. Second, I conduct a falsification test where I investigate if there is a similar jump at the treatment cut-off for a previously determined variable, where there should not be one. Third, I check if covariates trend smoothly at the cut-off, as unbalanced covariates would suggest that the "barely Democrat" and "barely Republican" subsamples are not identical. As the McCrary (2008) density test was not available at the time of LMB's study, they did not employ it. However, they did to the other two sets of tests I mentioned. So

[10] The preferred bandwidth is around 0.05. This is based on the optimal bandwidths selected by cross-validation for each of the component parts: the total effect (0.0537), the effect of party affiliation on roll-call voting (0.0612) and the incumbency effect (0.0425).

these later tests are not novel, although I do conduct them using the updated RD design methodology that I used earlier.

Density of the Assignment Variable

One test that provides evidence that there was not manipulation of the assignment variable, and thus the “treatment” and “control” groups are likely balanced, is the McCrary (2008) density test. This test measures the change in density of the frequency distribution of the assignment variable at the cut-off. If the distribution is continuous at the cut-off, it provides evidence against manipulation of the assignment variable. The density would not be continuous if, for example, the cut-off point was a failing grade in a course and professors had a tendency to mercy pass students. This would lead to a large mass point just to the right of the cut-off.

The McCrary (2008) test provides in an estimated discontinuity of 0.049, with a standard error of 0.048. Therefore, the null hypothesis of continuity is not rejected, suggesting that there is no jump in the density of the frequency distribution at the cut-off. Figure 7 shows this test visually and confirms that there is no jump.

Falsification Test Based on Previously Determined Scores

Another check on the quasi-random assignment assumption is to see if there is a discontinuity in a predetermined variable. Figure 8 shows Americans for Democratic Action (ADA) scores for the previous election, plotted against the Democratic vote share during the current election. As expected, there is no discernible discontinuity. If there were one, this could indicate that the barely Democrat and barely Republican districts are not similar. Results for a local linear regression, with a triangular kernel and the IK bandwidth, confirm that there is no discontinuity. This estimation gives an estimate of 3.62 for the discontinuity, with a standard error of 3.68. This is similar to LMB’s estimate.

Continuity of Covariates

Similar to for predetermined variables, district covariates should also be continuous. If there is a jump in a district covariate at the cut-off, this indicates that barely Democrat and barely Republicans differ in that covariate. This violates the quasi-random assignment assumption. Estimates will pick up bias from the effect of covariates. Figures 9 and 10 show that various district covariates are continuous in the assignment variable, and therefore, there is not visual evidence that barely Democrat and barely Republican districts differ in these observables.

Table 8 thoroughly tests the continuity of the district covariates. The first three columns show estimates of the discontinuity at the cut-off using local linear regression, with a triangular kernel, and 50%, 100%, and 200% of the IK bandwidth. The fourth and fifth columns are a difference in means using the close elections sample. The fourth column is my results, and the fifth column is LMB's results.

The results show that most of the district covariates are continuous; however, log income and percentage urban are discontinuous. Barely Democrat districts seem to be slightly richer and more urban than barely Republican districts. The ADA score is somewhat positively correlated with income and moderately positively correlated with percentage urban. So some of the estimates of the total effect (γ) and the effect of party affiliation (π_1) may reflect a mild, but likely trivial, effect from an increase in the percentage urban and an increase in income. But this is not really concerning. Even if all covariates were smooth, by the nature of this statistical testing we would expect 5% of the estimates to be statistically significant under 95% confidence intervals. So one or two estimates being statistically significant isn't damning evidence suggesting unbalanced covariates. Since all the other covariates are smooth, including the frequency (as tested following McCrary, 2008), and since the quasi-random assignment assumptions make intuitive sense in this application, there is negligible concern that the quasi-random assignment assumption doesn't hold.

Since the covariates are generally smooth at the cut-off, the inclusion or exclusion of covariates likely has no effect on the results. However, including covariates can reduce some small sample biases in the specification, and more importantly, improve the precision of estimates by soaking up some of the residual variation[11]. However, covariates are not necessary to provide consistent estimates (Imbens and Lemieux, 2007). I re-estimate the main results including them as another robustness check[12]. Table 9 provides these results for the simple difference-in-means estimate, employed by LMB. The first row shows the original results from Table 1 and the second row adds the same covariates from Table 8. Not surprisingly, the estimates barely change.

4. CONCLUSION

Lee, Moretti, and Butler (2004) find that politicians in U.S. House of Representatives elections do not respond to leftward or rightward shifts in the median voter through the mechanism of quasi-randomly assigned incumbency effects, which shift the median voter by providing benefits for the incumbent. This provides support for the full policy divergence theory. Thus voters "elect" policies rather than "affect" policies. However, the estimation strategy used by LMB is different from the approach suggested by the literature, much of which was published after LMB. However, when I replicate LMB using both their methods and the updated methods, I find the same results. The results are robust to using either a parametric or a non-parametric estimation strategy (and to many versions of both) and the fundamental quasi-randomization assumption seems to hold even under additional testing.

LMB's results pass both my additional barrage of tests, plus the ones they included in their original paper. However, other researchers who used a RD design but employed similarly outdated methodology may not be so lucky. Readers who interpret previous research using RD design designs, but with an

[11] However, Nichols (2011) mentions the possibility that poorly measured or endogenous covariates could actually decrease precision and increase bias. Regardless, inclusion of covariates provides a valuable specification check.
[12] LMB indicate that their results do not change upon inclusion of the covariates, but they do not report these results.

outdated methodology, should be wary of the results. Hopefully as replication becomes more standard in economics, we can identify whether previous studies hold up to improvements in methodology.

5. REFERENCES

Alesina, A. 1988. "Credibility and Policy Convergence in a 2-Party System with Rational Voters." *American Economic Review*, 78(4), 796-805.

Angrist, J. D.; G. W. Imbens and D. B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434), 444-55.

Button, Patrick. 2015. "Model Averaging and Model Uncertainty in Regression Discontinuity Designs." Forthcoming in the *Journal of Econometric Methods*.

Calvert, R. L. 1985. "Robustness of the Multidimensional Voting Model - Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science*, 29(1), 69-95.

Card, David; Carlos Dobkin and Nicole Maestas. 2009. "Does Medicare Save Lives?" *Quarterly Journal of Economics*, 124(2), 597-636.

Downs, A. 1957. "An Economic-Theory of Political-Action in a Democracy." *Journal of Political Economy*, 65(2), 135-50.

Fan, Jianqing and Irene Gijbels. 1996. *Local Polynomial Modelling and Its Applications*. London, New York, and Melbourne: Chapman and Hall.

Ferreira, Fernando and Joseph Gyourko. 2009. "Do Political Parties Matter? Evidence from U. S. Cities." *Quarterly Journal of Economics*, 124(1), 399-422.

Imbens, G. W. and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2), 615-35.

Imbens, Guido W. and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic* Studies, xx: 1-27.

Lee, D. S. and T. Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2), 281-355.

Lee, D. S.; E. Moretti and M. J. Butler. 2004. "Do Voters Affect or Elect Policies? Evidence from the US House." *Quarterly Journal of Economics*, 119(3), 807-59.

Lee, D. S.; E. Moretti and M. J. Butler. 2001. "Credibility and Policy Convergence: Evidence from US Roll Call Voting Records." *National Bureau of Economic Research Working Paper 9315*.

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in Us House Elections." *Journal of Econometrics*, 142(2), 675-97.

McCarty, Nolan M.; Keith T. Poole and Howard Rosenthal. 1997. *Income Redistribution and the Realignment of American Politics*. Washington, DC: The ARI Press.

McCrary, J. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2), 698-714.

Nichols, Austin. 2011. "Rd 2.0: Revised Stata Module for Regression Discontinuity Estimation," http://ideas.repec.org/c/boc/bocode/s456888.html:

Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. Oxford, UK: Oxford University Press.

Porter, Jack. 2003. "Estimation in the Regression Discontinuity Model" Unpublished Material

FIGURE 1 – DIAGRAM OF TIME PERIODS AND EVENTS IN THE LMB MODEL

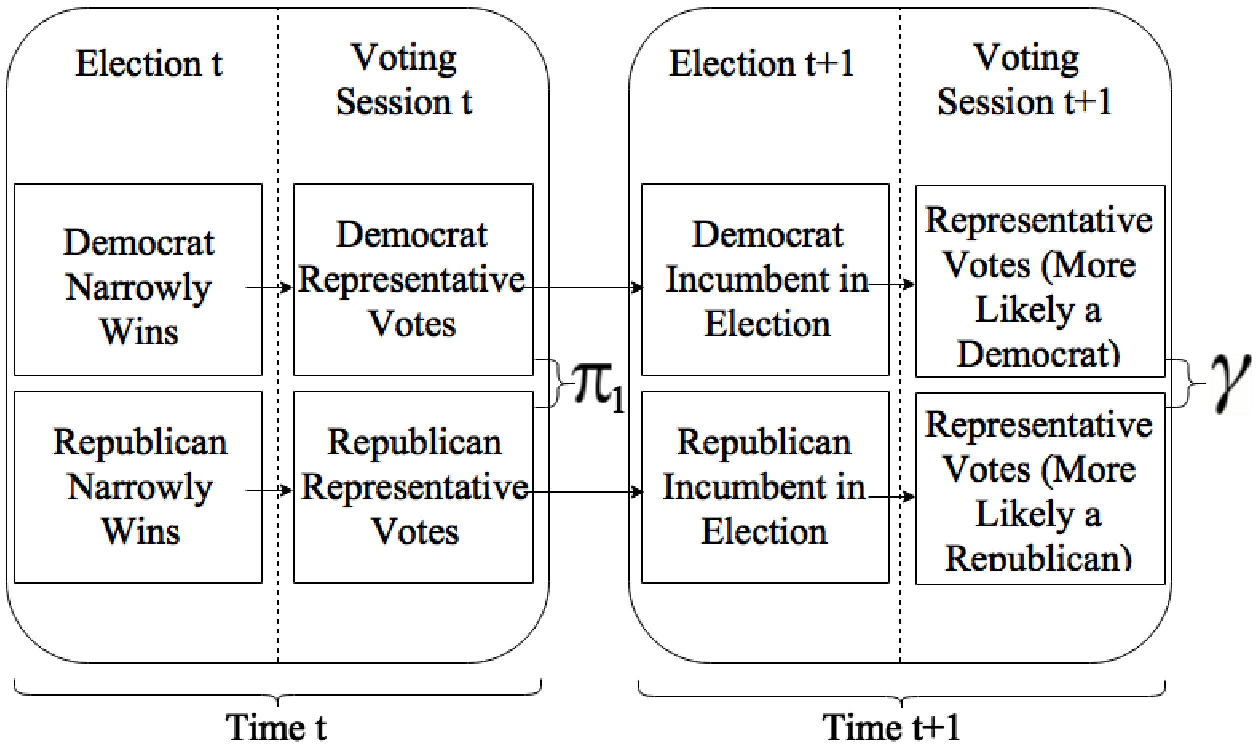
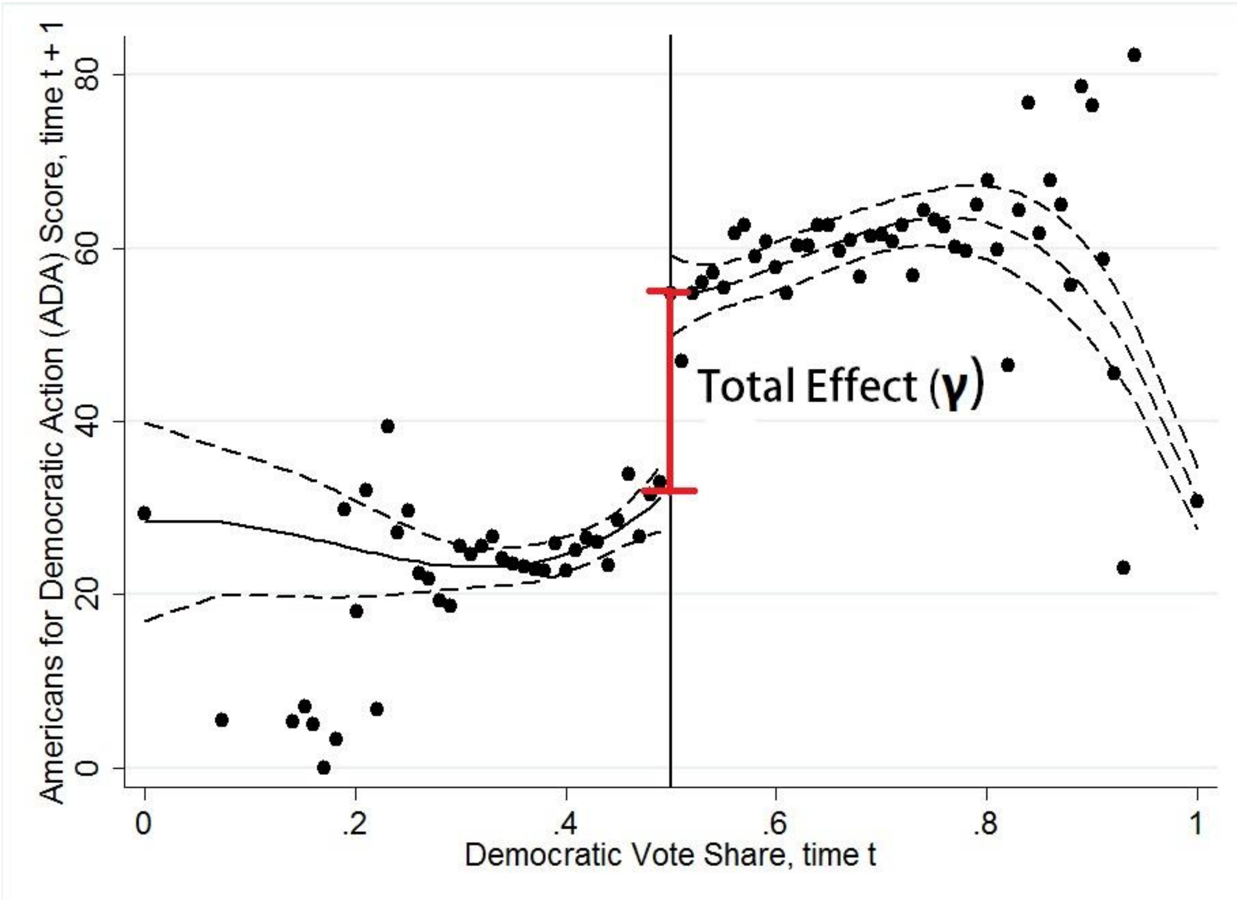
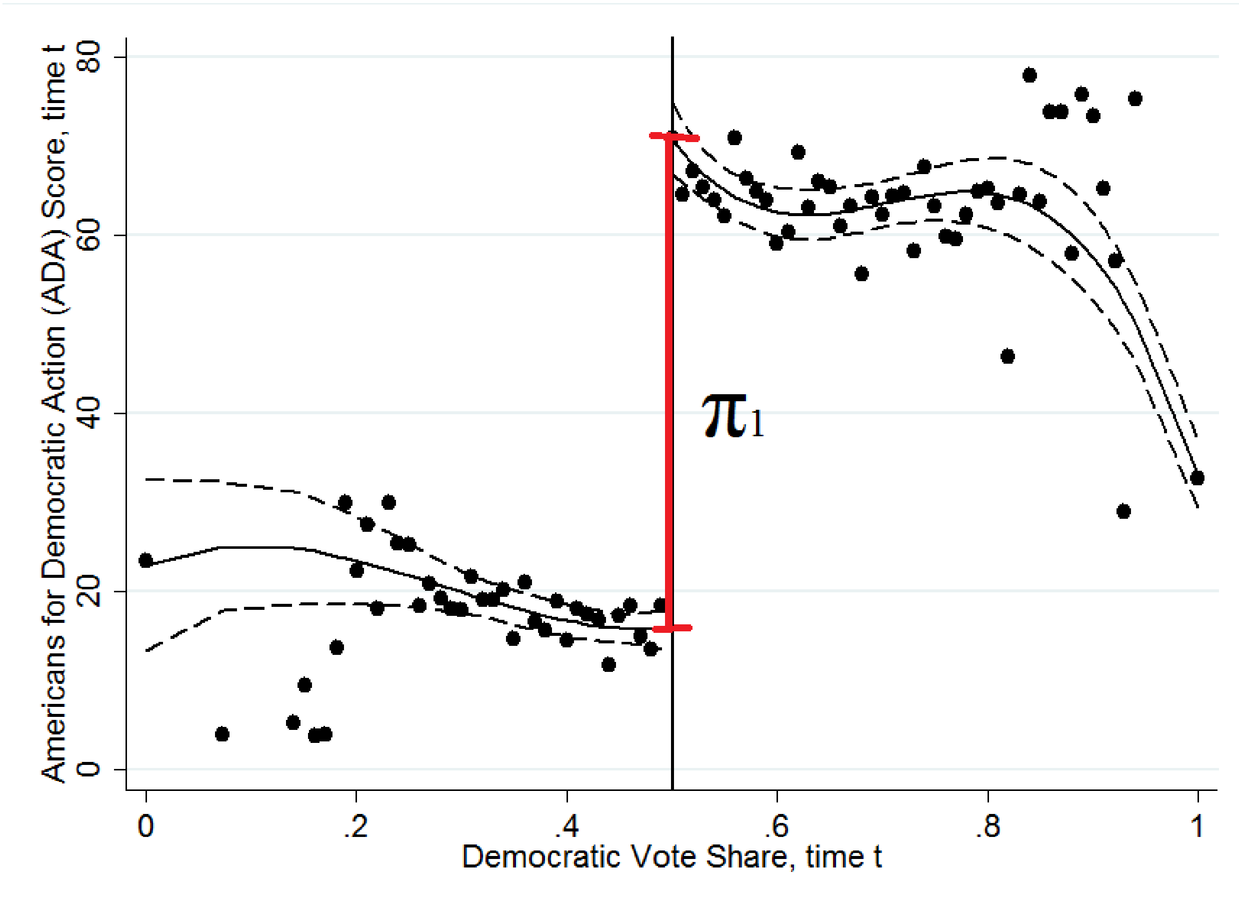


FIGURE 2 – TOTAL EFFECT OF INITIAL WIN ON FUTURE ADA SCORES: γ



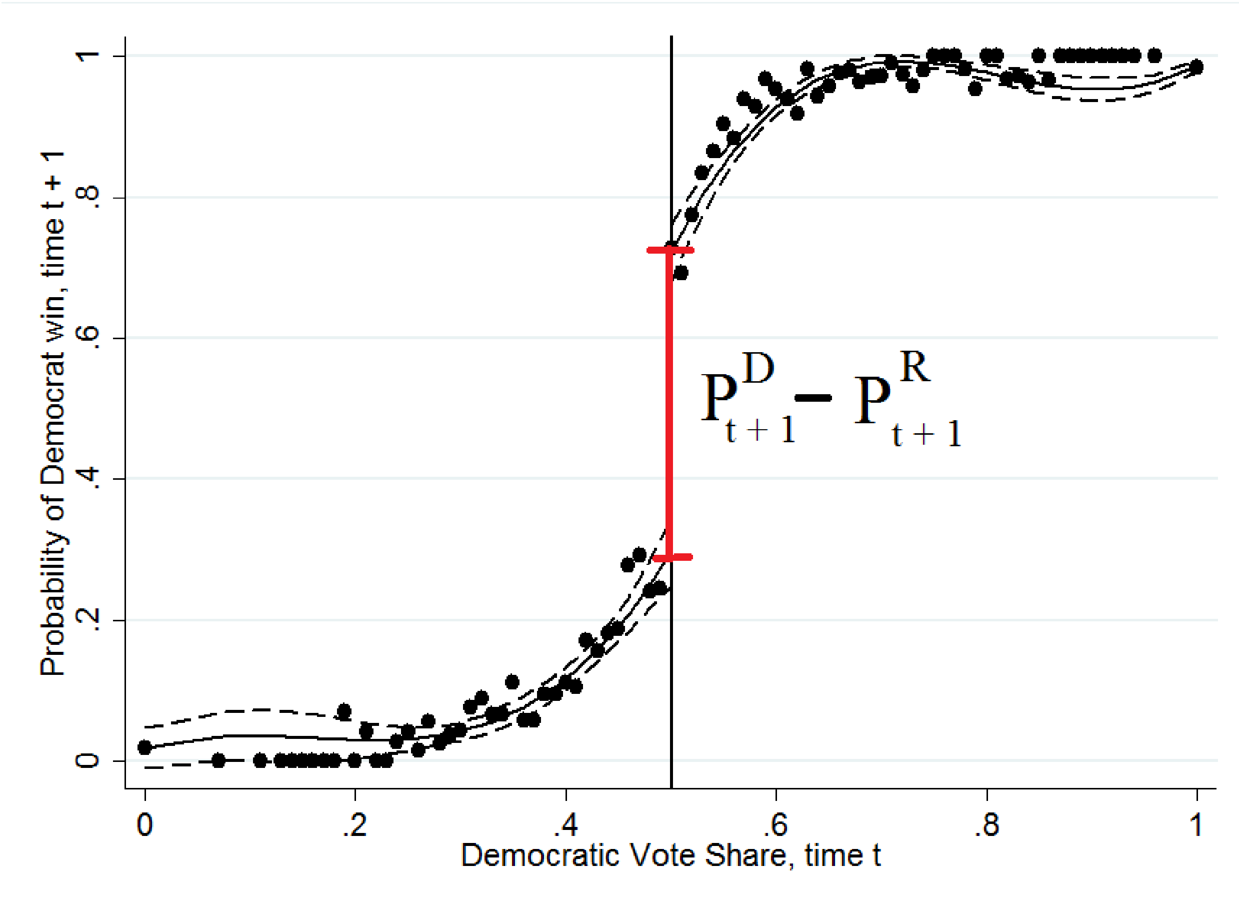
Notes: The scatterplot points represent the mean value of the ADA Score at time t+1 for each one percentage point bin of the Democratic Vote Share at time t. The curved line is from a 3rd order polynomial that separately fits each side of the cut-off (0.5). Dashed lines are 95% confidence intervals.

FIGURE 3 – EFFECT OF PARTY AFFILIATION: π_1



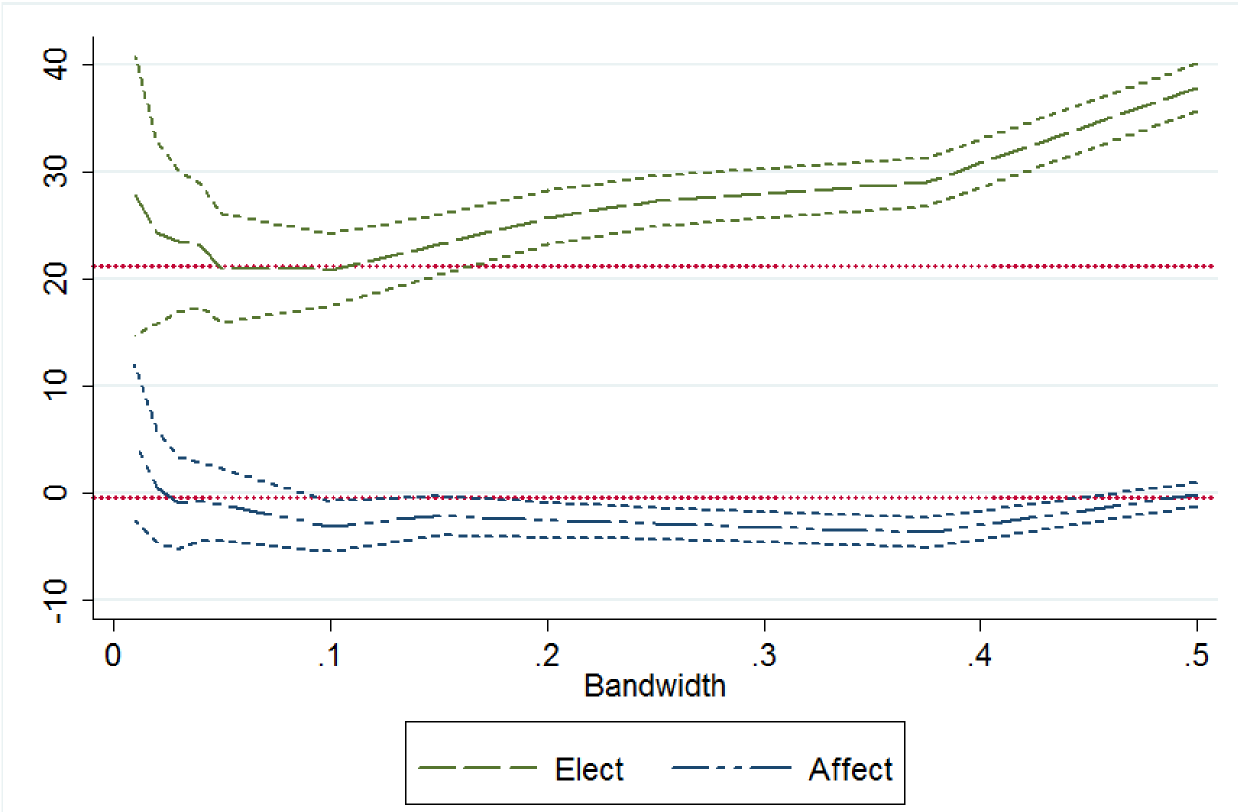
Notes: See the notes to Figure 2.

FIGURE 4 – EFFECT OF INITIAL WIN ON WINNING NEXT ELECTION: ($P^D_{t+1} – P^R_{t+1}$)



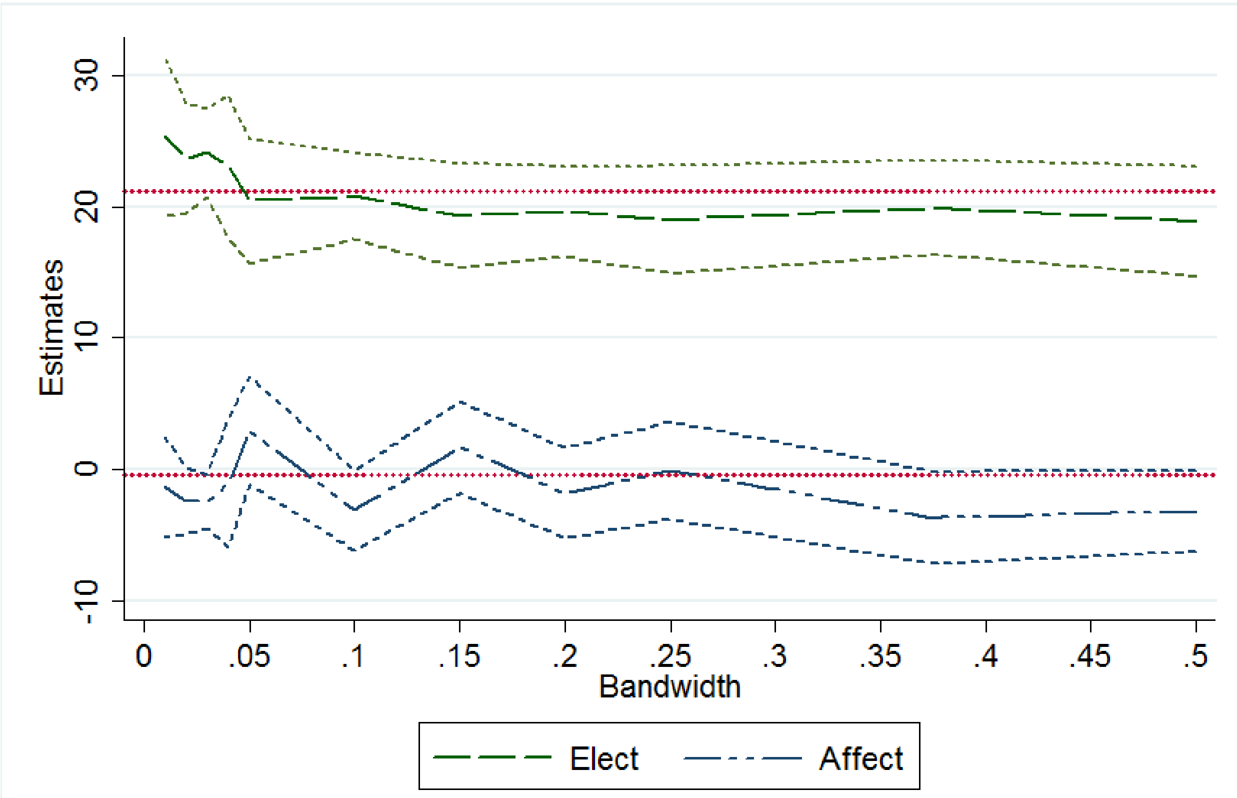
Notes: See the notes to Figure 2.

FIGURE 5 - LOCAL LINEAR REGRESSION ESTIMATES FOR VARIOUS BANDWIDTHS



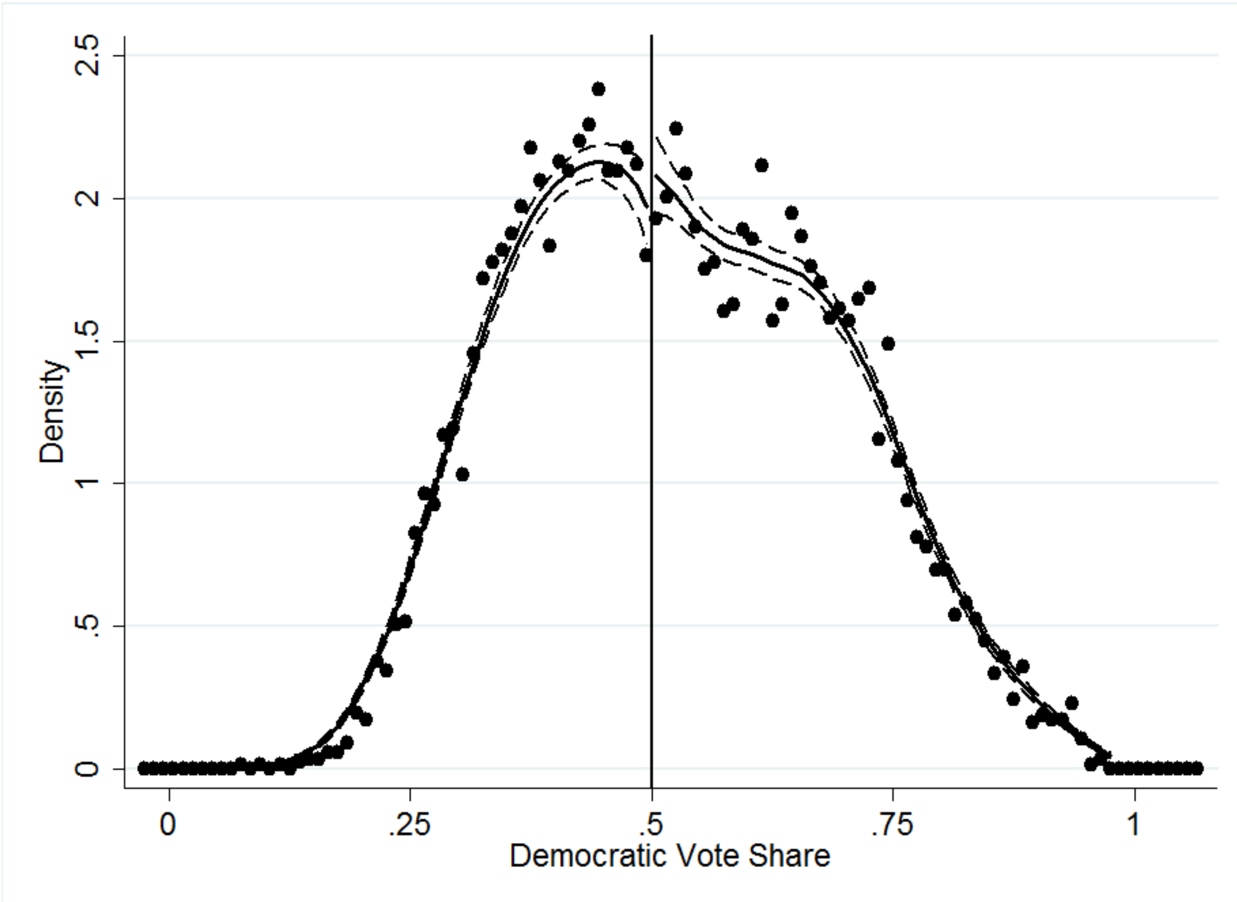
Notes: Estimates are based on local linear regressions with a rectangular kernel. The dashed lines are 95% confidence intervals from standard errors clustered at the district-decade level. Dotted red lines represent the elect (21.2) and affect (-0.47) point estimates calculated from the estimates in Table 2, under a triangular kernel and IK bandwidth.

FIGURE 6 – ELECT AND AFFECT COMPONENT ESTIMATES UNDER POLYNOMIAL MODELS FOR VARIOUS BANDWIDTHS



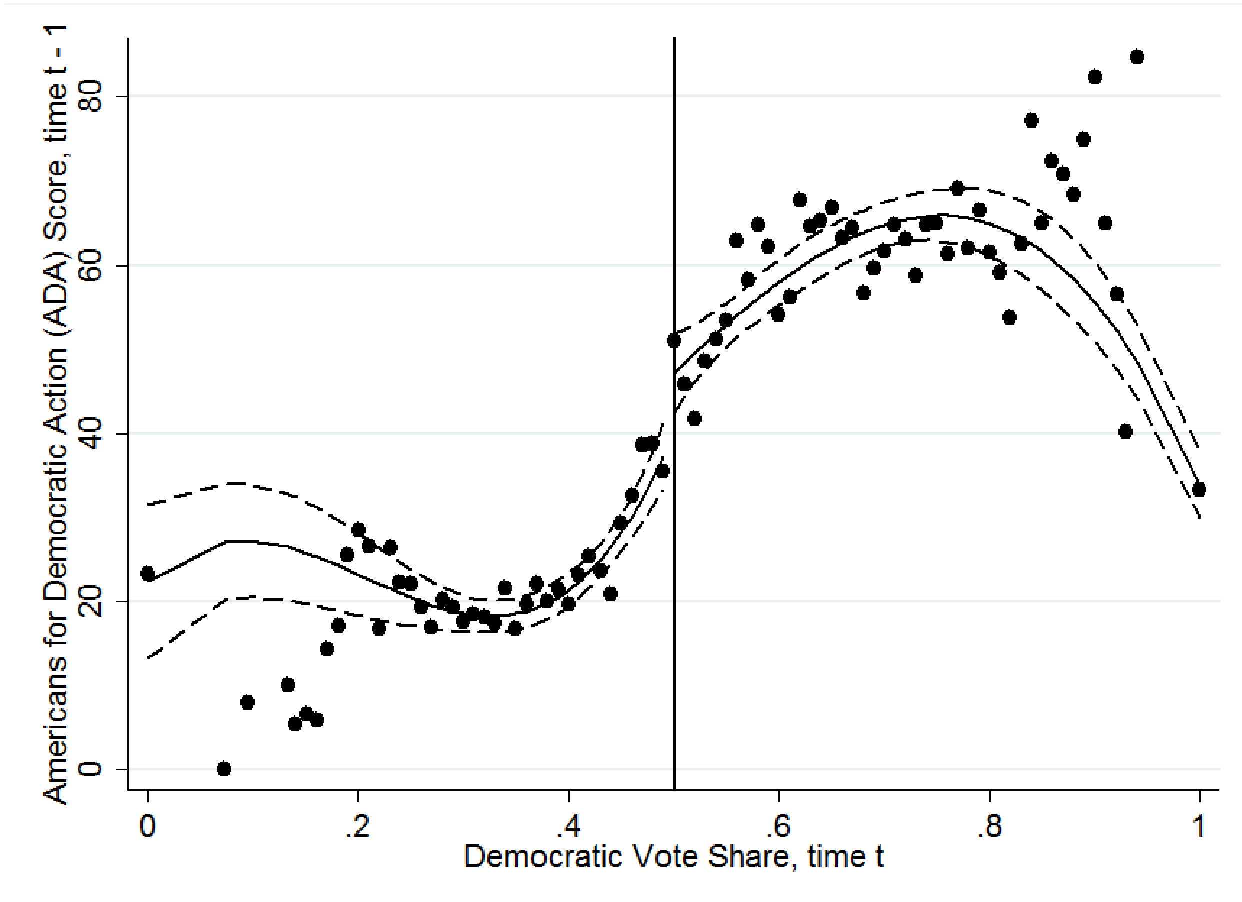
Notes: See the notes to Figure 5. Estimates are based on polynomial models, where the polynomial model used is determined by BIC for each bandwidth.

FIGURE 7 - McCRARY (2008)'S DENSITY TEST



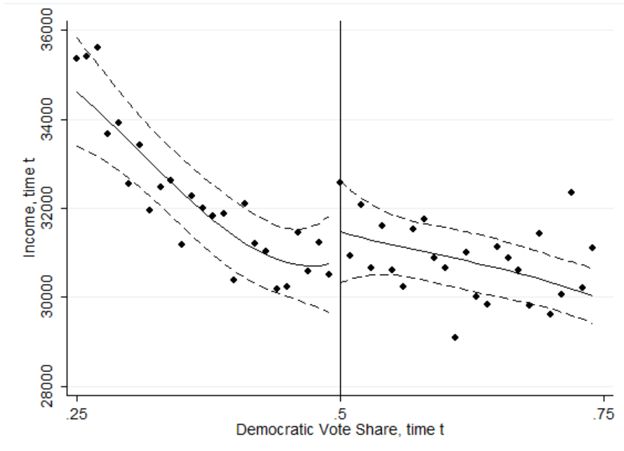
Notes: See the notes to Figure 2. The solid line is instead estimated non-parametrically. This figure excludes elections where there was no major party challenger (Democratic Vote Share of 0 or 1) for better presentation.

FIGURE 8 – TESTING THE CONTINUITY OF THE LAGGED ADA SCORE

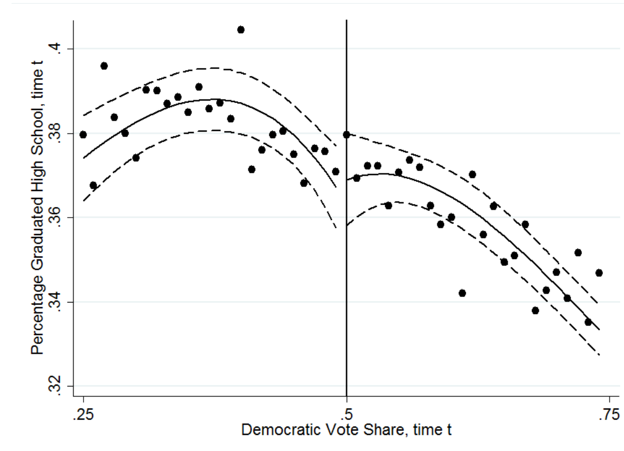


Notes: See the notes to Figure 2.

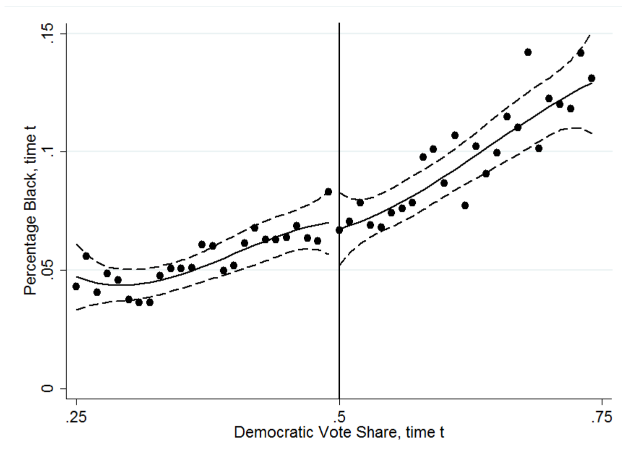
FIGURE 9 – DISTRICT COVARIATES BY DEMOCRAT VOTE SHARE (Part 1)



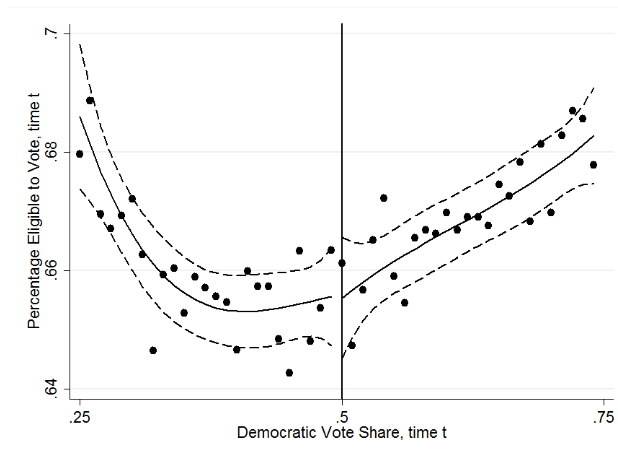
(a) Income



(b) Percent Graduated High School



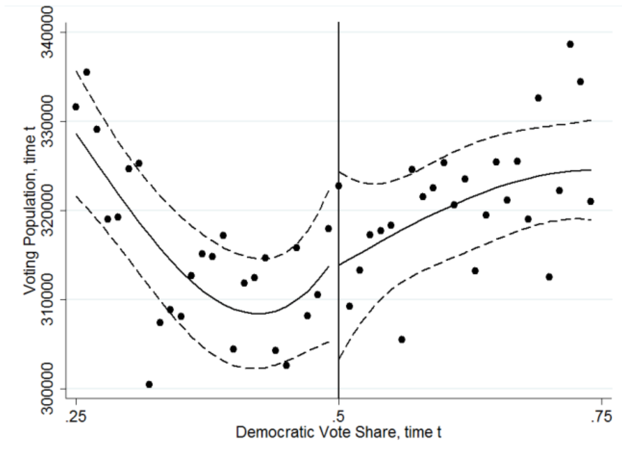
(c) Percent Black



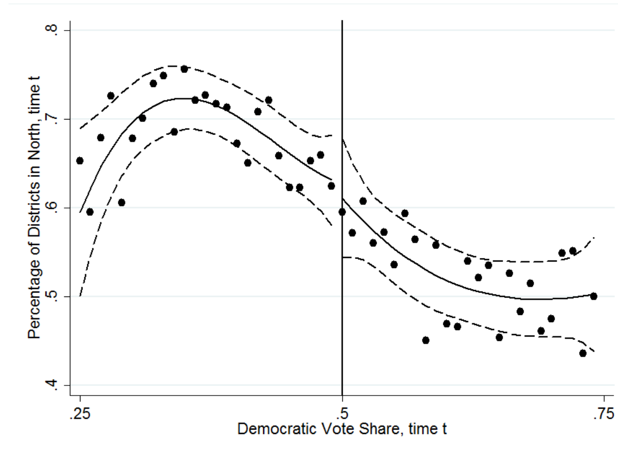
(d) Percent Eligible to Vote

Notes: See the notes to Figure 2.

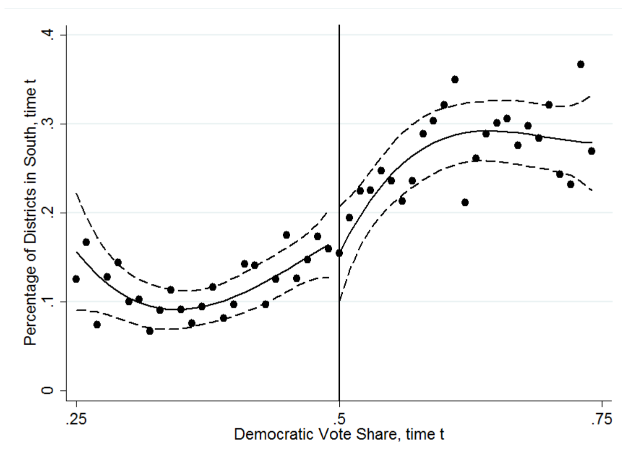
FIGURE 10 – DISTRICT COVARIATES BY DEMOCRAT VOTE SHARE (Part 2)



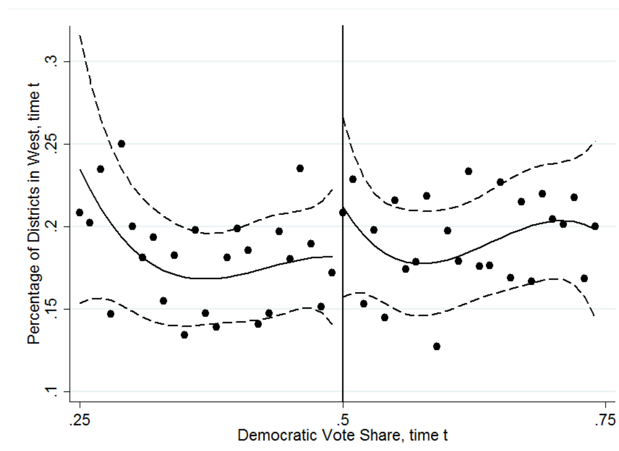
(a) Voting Population



(b) Percent of Districts in the North



(c) Percent of Districts in the South



(d) Percent of Districts in the West

Notes: See the notes to Figure 2.

TABLE 1 - REPLICATION OF ESTIMATES FROM LMB (2004): DIFFERENCES IN MEANS USING CLOSE ELECTIONS SAMPLE

Total Effect (γ) (1)	π_1 (2)	$P^D_{t+1} - P^R_{t+1}$ (3)	Elect Component (col.(2)*col.(3))	Affect Component (col.(1) – col.(4))
21.28	47.71	0.4843	23.11	-1.82
(1.95)	(1.36)	(0.0289)	(1.50)	(1.11)
21.2	*47.6*	*0.48*	*22.84*	*-1.64*
(1.9)	*(1.3)*	*(0.02)*	*(2.2)*	*(2.0)*

Notes: Standard errors, in parentheses, are clustered at the district-decade level. The unit of observation is the district-congressional session. Sample size is 915. My estimates are in black while Lee, Moretti, and Butler's (2004) estimates are below in red italics.

TABLE 2 – LOCAL LINEAR REGRESSION RESULTS BY KERNEL TYPE, USING THE IK BANDWIDTH

	Total Effect (γ)	π_1	$P^D_{t+1} - P^R_{t+1}$
Triangular	20.75	52.10	0.4073
	(3.61)	(2.20)	(0.0403)
IK Bandwidth	0.1680	0.1906	0.0712
Rectangular	21.15	51.65	0.3870
	(4.03)	(2.34)	(0.0414)
IK Bandwidth	0.1320	0.1497	0.0559

Notes: See the notes to Table 1.

TABLE 3 – SENSITIVITY OF THE LOCAL LINEAR REGRESSION RESULTS TO DEVIATIONS FROM THE IK BANDWIDTH

	Bandwidth							
	10%	20%	50%	100%	200%	300%	400%	500%
Total Effect (γ)	41.34	25.369	19.95	20.75	23.5	24.62	31.14	32.65
	(11.69)	(8.14)	(4.32)	(3.61)	(2.87)	(3.43)	(3.94)	(3.40)
π_1	47.64	50.7	51.76	52.1	51.78	53.97	55.53	56.04
	(8.23)	(5.33)	(2.98)	(2.20)	(1.80)	(2.09)	(2.23)	(2.00)
$P^D_{t+1} - P^R_{t+1}$	0.4529	0.5651	0.4697	0.4073	0.4254	0.4774	0.516	0.5394
	(0.1739)	(0.1054)	(0.0594)	(0.0403)	(0.0368)	(0.0321)	(0.0273)	(0.0243)

Notes: See the notes to Table 1. Results used a triangular kernel.

TABLE 4 – TOTAL EFFECT (γ) ESTIMATES BY BANDWIDTH AND POLYNOMIAL ORDER

	Bandwidth										
Polynomial of order:	All Data	0.375	0.25	0.2	0.15	0.1	0.05	0.04	0.03	0.02	0.01
Zero	31.43	37.16	35.99	34.44	32.60	30.01	23.41	21.97	21.66	21.28	23.97
	(1.00)	(0.97)	(0.99)	(1.03)	(1.11)	(1.27)	(1.63)	(1.85)	(2.13)	(2.66)	(3.89)
One	37.59	25.26	24.43	23.31	21.44	17.63	19.58	22.36	22.54	24.04	31.36
	(1.36)	(1.51)	(1.60)	(1.71)	(1.90)	(2.28)	(3.29)	(3.60)	(4.28)	(5.23)	(7.27)
Two	18.22	21.53	19.27	18.10	16.91	19.18	25.67	23.84	27.56	35.81	30.17
	(1.77)	(2.01)	(2.23)	(2.45)	(2.80)	(3.49)	(4.80)	(5.41)	(6.31)	(7.59)	(11.30)
Three	21.56	16.56	17.32	16.62	18.71	24.70	26.07	31.88	34.72	28.08	9.23
	(2.31)	(2.55)	(2.95)	(3.28)	(3.81)	(4.55)	(6.46)	(7.28)	(8.34)	(10.67)	(16.73)
Four	16.02	18.62	17.02	21.42	26.13	28.00	34.94	35.74	33.47	20.66	24.12
	(2.85)	(3.13)	(3.71)	(4.12)	(4.64)	(5.65)	(8.14)	(9.09)	(10.92)	(14.00)	(21.14)
Optimal order by AIC	5	7	7	6	5	6	4	3	5	6	3
Optimal order by BIC	4	3	2	2	1	1	0	0	0	0	0
Observations	13983	11914	10686	9140	7228	4936	2523	2017	1505	950	445

Notes: See the notes to Table 1.

TABLE 5 - EFFECT OF PARTY AFFILIATION (π_1) ESTIMATES BY BANDWIDTH AND POLYNOMIAL ORDER

	Bandwidth										
Polynomial of order:	All Data	0.375	0.25	0.2	0.15	0.1	0.05	0.04	0.03	0.02	0.01
Zero	42.59	48.82	48.71	48.59	49.01	49.04	48.95	49.40	50.14	48.90	47.76
	(0.88)	(0.81)	(0.90)	(0.93)	(0.97)	(1.08)	(1.27)	(1.39)	(1.55)	(2.01)	(2.97)
One	56.90	48.84	49.75	49.66	49.22	48.82	48.74	48.58	46.70	45.26	45.04
	(1.05)	(1.32)	(1.18)	(1.20)	(1.30)	(1.47)	(2.05)	(2.24)	(2.73)	(3.31)	(4.91)
Two	46.26	50.00	49.24	49.10	48.98	48.55	46.09	43.91	44.77	45.68	39.40
	(1.33)	(1.40)	(1.47)	(1.60)	(1.74)	(2.16)	(3.11)	(3.57)	(4.18)	(5.04)	(7.71)
Three	51.12	47.28	48.48	47.75	47.67	47.09	43.37	45.49	43.86	39.55	35.97
	(1.58)	(1.66)	(1.88)	(2.05)	(2.38)	(2.99)	(4.37)	(4.85)	(5.60)	(7.11)	(10.59)
Four	47.30	50.08	47.62	48.87	47.28	44.81	44.79	42.92	42.12	39.40	33.59
	(2.20)	(2.01)	(2.32)	(2.59)	(3.07)	(3.77)	(5.45)	(6.06)	(7.24)	(9.30)	(13.03)
Optimal order by AIC	6	5	2	0	0	3	2	2	1	1	0
Optimal order by BIC	4	1	0	0	0	0	0	0	0	0	0
Observations	18810	15972	14293	12220	9566	6484	3286	2641	1974	1257	597

Notes: See the notes to Table 1.

TABLE 6 – Incumbency Effect ($P^D_{t+1} - P^R_{t+1}$) ESTIMATES BY BANDWIDTH AND POLYNOMIAL ORDER

	Bandwidth										
Polynomial of order:	All Data	0.375	0.25	0.2	0.15	0.1	0.05	0.04	0.03	0.02	0.01
Zero	0.8168	0.7986	0.7816	0.7553	0.7170	0.6481	0.5298	0.4931	0.4813	0.4843	0.5307
	(0.0078)	(0.0085)	(0.0092)	(0.0104)	(0.0123)	(0.0158)	(0.0232)	(0.0261)	(0.0306)	(0.0383)	(0.0552)
One	0.6656	0.5882	0.5513	0.5166	0.4702	0.4244	0.4176	0.4667	0.4571	0.4634	0.5764
	(0.0145)	(0.0180)	(0.0203)	(0.0228)	(0.0264)	(0.0328)	(0.0473)	(0.0534)	(0.0632)	(0.0803)	(0.1149)
Two	0.5111	0.4584	0.4190	0.4042	0.3943	0.3999	0.5025	0.5059	0.5212	0.6479	0.5184
	(0.0223)	(0.0269)	(0.0309)	(0.0349)	(0.0401)	(0.0496)	(0.0728)	(0.0824)	(0.0983)	(0.1236)	(0.1833)
Three	0.4225	0.4083	0.3905	0.3745	0.4120	0.4548	0.5240	0.5452	1.7371	0.4998	0.1622
	(0.0308)	(0.0358)	(0.0416)	(0.0464)	(0.0542)	(0.0675)	(0.1010)	(0.1156)	(1.0043)	(0.1693)	(0.2609)
Four	0.3996	0.3697	0.3884	0.4271	0.4560	0.5441	0.5861	0.6461	-0.3727	0.3798	0.2259
	(0.0394)	(0.0441)	(0.0525)	(0.0590)	(0.0683)	(0.0868)	(0.1303)	(0.1472)	(4.8861)	(0.2175)	(0.3064)
Optimal order by AIC	6	7	7	6	5	4	3	2	5	6	5
Optimal order by BIC	4	3	3	2	2	1	1	1	0	0	0
Observations	14963	12708	11386	9724	7684	5226	2688	2148	1602	1018	484

Notes: See the notes to Table 1.

TABLE 7 – COMBINING ESTIMATES FROM TABLES II, III, and IV to ESTIMATE "ELECT" and "AFFECT" COMPONENTS

	Bandwidth										
	All Data	0.375	0.25	0.2	0.15	0.1	0.05	0.04	0.03	0.02	0.01
Total Effect	15.70	16.23	18.92	17.80	20.99	17.66	23.41	21.97	21.66	21.28	23.97
(γ)	(2.90)	(2.59)	(2.26)	(2.50)	(1.92)	(2.32)	(1.63)	(1.85)	(2.13)	(2.66)	(3.89)
	4	3	2	2	1	1	0	0	0	0	0
π_1	47.30	48.84	48.71	48.59	49.01	49.04	48.95	49.40	50.14	48.90	47.76
	(2.20)	(1.32)	(0.90)	(0.93)	(0.97)	(1.08)	(1.27)	(1.39)	(1.55)	(2.01)	(2.97)
	4	1	0	0	0	0	0	0	0	0	0
$P^D_{t+1} - P^R_{t+1}$	0.3996	0.4083	0.3905	0.4042	0.3943	0.4244	0.4176	0.4667	0.4813	0.4843	0.5307
	(0.0394)	(0.0358)	(0.0416)	(0.0349)	(0.0401)	(0.0328)	(0.0473)	(0.0534)	(0.0306)	(0.0383)	(0.0552)
	4	3	3	2	2	1	1	1	0	0	0
Elect	18.90	19.94	19.02	19.64	19.32	20.81	20.44	23.06	24.13	23.68	25.35
Component	(2.09)	(1.80)	(2.05)	(1.72)	(1.99)	(1.65)	(2.37)	(2.72)	(1.68)	(2.08)	(3.06)
Affect	-3.20	-3.71	-0.10	-1.84	1.67	-3.15	2.97	-1.08	-2.47	-2.40	-1.38
Component	(1.54)	(1.76)	(1.86)	(1.72)	(1.75)	(1.53)	(2.07)	(2.45)	(1.02)	(1.29)	(1.88)

Notes: See the notes to Table 1. The order of the polynomial used, based on the BIC, is shown below the standard errors.

TABLE 8 - ESTIMATED DISCONTINUITY OF DISTRICT COVARIATES

	50%	100%	200%	Close Elections	Close Elections (LMB's Result)
North	**-0.045**	-0.023	**-0.043**	-0.060	*-0.059*
	(0.022)	(0.024)	**(0.022)**	(0.037)	*(0.036)*
South	**-0.211**	0.001	0.019	0.008	*0.009*
	(0.029)	(0.022)	(0.020)	(0.028)	*(0.036)*
West	0.037	0.012	0.017	0.058	*0.001*
	(0.035)	(0.022)	(0.014)	(0.030)	*(0.036)*
Log Income	0.048	**0.046**	**0.046**	0.030	*0.052*
	(0.030)	**(0.022)**	**(0.018)**	(0.026)	*(0.033)*
% High-School Graduates	0.004	0.007	0.002	0.001	*0.008*
	(0.009)	(0.006)	(0.006)	(0.008)	*(0.008)*
% Urban	**0.066**	**0.055**	**0.039**	**0.057**	***0.056***
	(0.029)	**(0.012)**	**(0.016)**	**(0.023)**	***(0.023)***
% Black	-0.004	-0.007	-0.010	-0.003	*-0.004*
	(0.012)	(0.008)	(0.008)	(0.009)	*(0.009)*
Manufacturing Employment	0.006	0.0063	**0.0066**	0.004	*0.005*
	(0.005)	(0.0034)	**(0.0028)**	(0.004)	*(0.004)*
Total Population	36965	55521	39240	7354	*8640*
	(23847)	(32417)	(21555)	(8552)	*(8428)*
% Eligible to Vote	-0.007	-0.001	0.004	-0.004	*-0.003*
	(0.007)	(0.005)	(0.005)	(0.006)	*(0.006)*
Open Seats	0.065	0.023	-0.003	0.027	***0.056***
	(0.055)	(0.041)	(0.028)	(0.034)	***(0.023)***

Notes: See the notes to Table 1. Columns 1 to 3 are results from a local linear regression using 50%, 100% and 200% of the IK Bandwidth. Column 4 is a difference in means for the close elections sample (48-52%). Column 5 is Lee, Moretti, and Butler (2004)'s results using the same methodology as in column 4. Bolded estimates indicate significance at the 95% level.

TABLE 9 - SENSITIVITY OF RESULTS TO THE INCLUSION OF COVARIATES

	Total Effect (γ)	π_1	$P^D_{t+1} - P^R_{t+1}$	Elect Component	Affect Component
No Covariates	21.28	47.71	0.4843	23.11	-1.82
	(1.95)	(1.36)	(0.0289)	(1.50)	(1.11)
All Covariates From Table XI	24.12	49.83	0.54	27.06	-2.93
	(3.21)	(2.34)	(0.05)	(2.57)	(1.66)

Notes: See the notes to Table 1. Sample size is 547 with covariates, and 915 without.

APPENDIX 1: SENSITIVITY OF RESULTS TO ALTERNATIVE MEASURES OF ROLL CALL VOTING

LMB chose Americans for Democratic Action (ADA) scores as the primary measure of "liberalness" or "conservativeness" of politicians, due to the prevalence use of these scores in the literature. This section re-estimates the affect and elect components under various other scores.

DW-NOMINATE SCORE

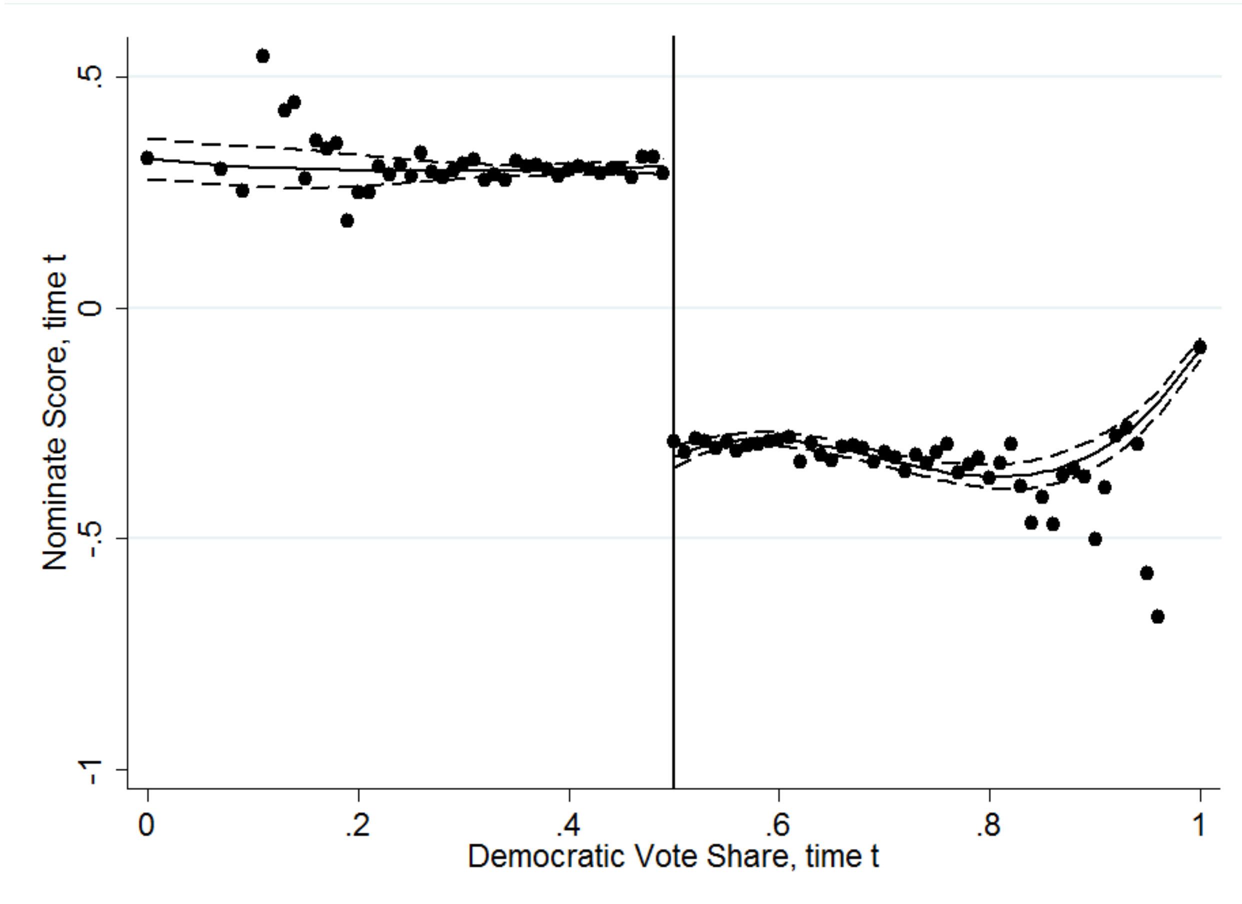
DW-NOMINATE score, developed by McCarty, Poole, and Rosenthal (1997), is another measure of liberalness that is well documented in the literature. The score has two dimensions: one measuring differences in economic ideology (DW-NOMINATE 1) and the other measuring differences in race-relations policies (DW-NOMINATE 2). The authors use the first score as the second score becomes irrelevant by about the mid 70s. (Poole and Rosenthal, 1997) For the first score, higher scores are less liberal. Appendix Table 1 estimates the affect and elect components for the DW-NOMINATE 1 score, using the close elections sample methodology of table I. LMB's estimates are in red italics below mine. The results are very similar to the ADA score results. Appendix Figure 1 displays the estimate of π_1.

APPENDIX TABLE 1 - RESULTS BASED ON DW-NOMINATE SCORES – CLOSE ELECTIONS SAMPLE

Total Effect (γ) (1)	π_1 (2)	$P^D_{t+1} - P^R_{t+1}$ (3)	Elect Component (col.(2)*col.(3))	Affect Component (col.(1) – col.(4))
-0.25	-0.59	0.46	-0.27	0.02
(0.03)	(0.02)	(0.04)	(0.02)	(0.02)
-0.36	*-0.58*	*0.62*	*-0.34*	*-0.02*
(0.03)	*(0.02)*	*(0.04)*	*(0.04)*	*(0.04)*

Notes. See the notes to Table 1.

APPENDIX FIGURE 1 – Estimate of π_1 using DW-NOMINATE 1 SCORES



Notes: See the notes to Figure 2.

PERCENT VOTED WITH DEMOCRAT LEADER SCORE

The authors created their own measure of "liberalness", which is the percent of representatives that voted the same way as the Democrat party leader. One would expect that Democrat representatives would be more likely than Republican representatives to vote with their leader. Since voting is less whipped in the United States than in other countries, particularly those following the Westminster model, not all Democrats will vote with their leader, and some Republicans will vote with the Democratic leader. Thus, this is another attractive measure that can be used to assess the robustness of estimates to various scores. Appendix Table 2 estimates the affect and elect components for this score. The results reach the same conclusion as those earlier. Appendix Figure 2 displays the estimate of π_1.

APPENDIX TABLE 2 - RESULTS BASED ON PERCENT VOTED WITH DEMOCRAT LEADER – CLOSE ELECTIONS SAMPLE

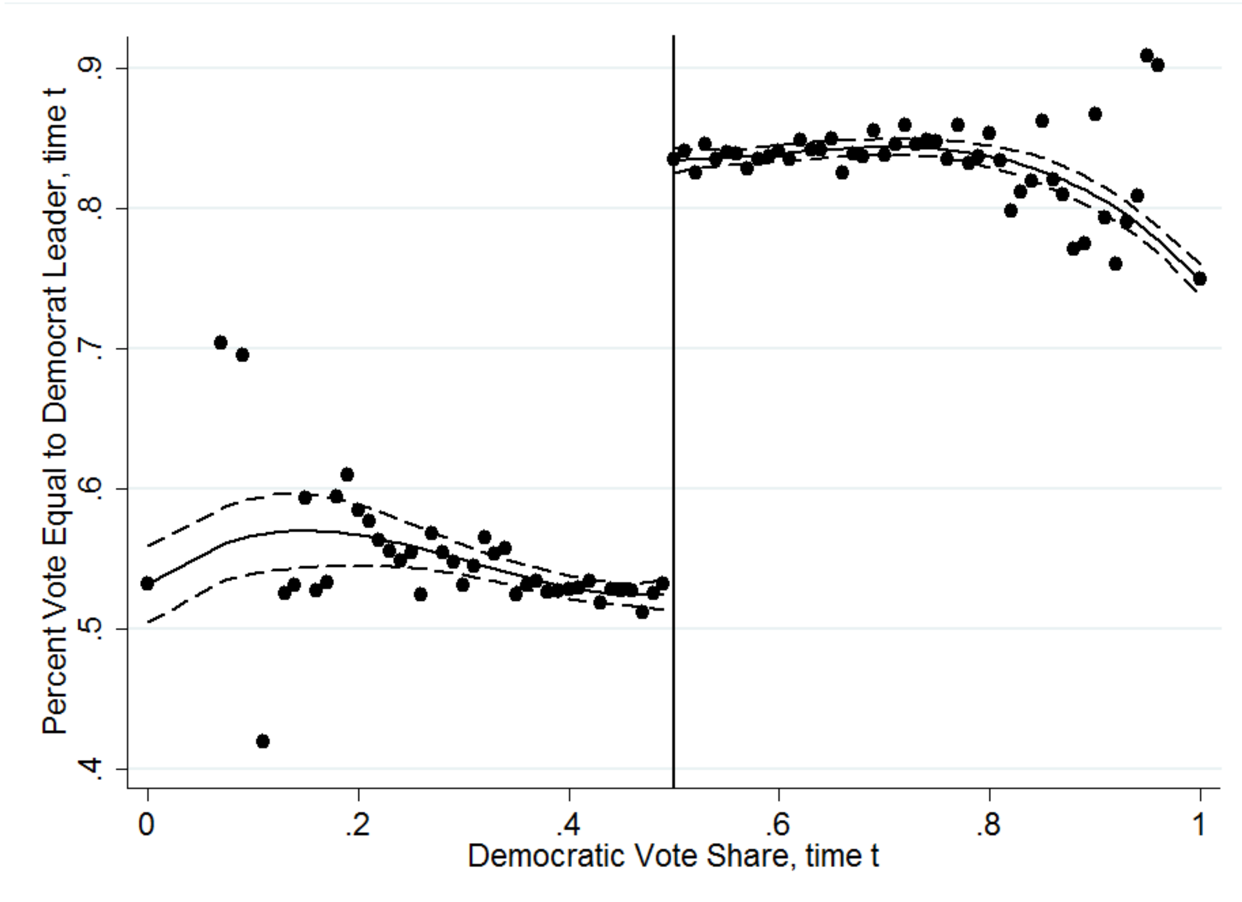
Total Effect (γ) (1)	π_1 (2)	$P^D_{t+1} - P^R_{t+1}$ (3)	Elect Component (col.(2)*col.(3))	Affect Component (col.(1) – col.(4))
0.14	0.30	0.47	0.14	-0.004
(0.02)	(0.01)	(0.04)	(0.01)	(0.009)
0.13	*0.29*	*0.46*	*0.13*	*0.00*
(0.01)	*(0.006)*	*(0.02)*	*(0.02)*	*(0.02)*

Standard Errors, in parentheses, are clustered at the district-decade level. The unit of observation is the district-congressional session. Standard errors, in parentheses, are clustered at the district-decade level. The unit of observation is the district-congressional session. Sample size is 1009. My estimates are in black while Lee, Moretti, and Butler's (2004) estimates are below in red italics.

OTHER INTEREST GROUP SCORES

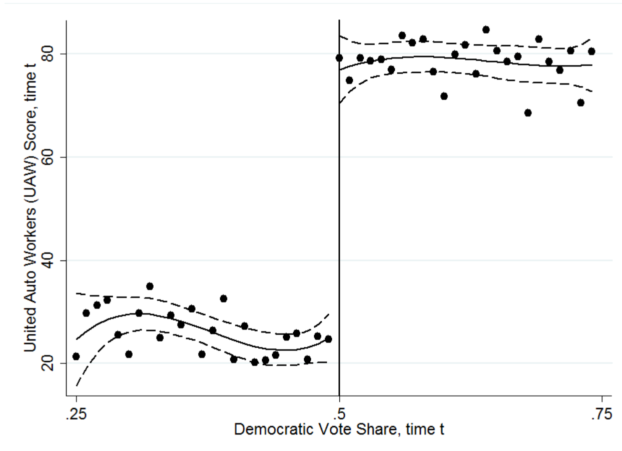
To prove that the earlier affect and elect component estimates are not simply from the use of Americans for Democratic Action (ADA) scores, the authors estimate the results using scores from other interest groups. This includes both "liberal" and "conservative" groups. The liberal groups are defined as those where higher scores indicate voting more liberal. Appendix Figures 3 and 4 show the estimates of π_1 using these liberal scores. Conservative group scores are defined where higher scores are less liberal. Appendix Figures 5 and 6 show the estimates of π_1 using these conservative scores.

APPENDIX FIGURE 2 – ESTIMATE OF π_1 USING PERCENT VOTED WITH DEMOCRAT LEADER

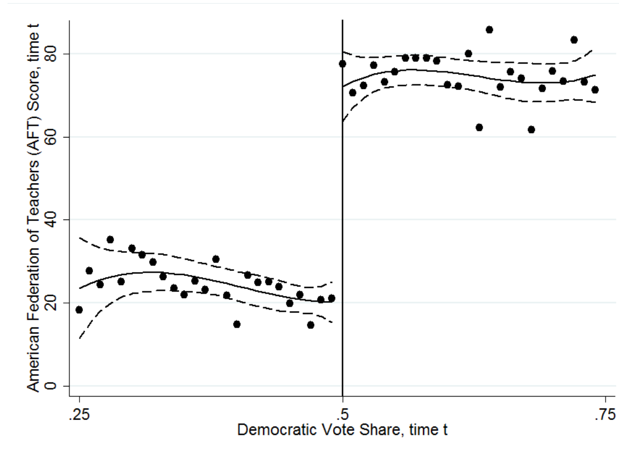


Notes: See the notes to Figure 2.

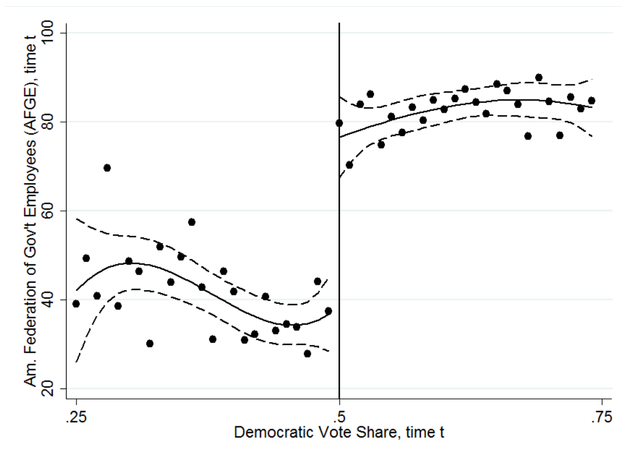
APPENDIX FIGURE 3 – ESTIMATES OF π_1 USING LIBERAL INTEREST GROUP SCORES (PART 1)



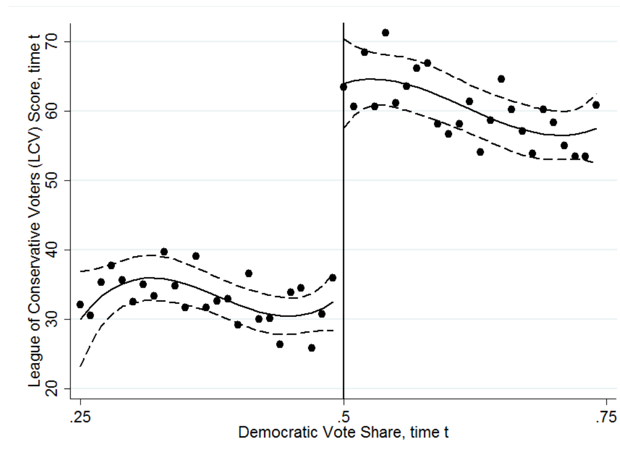
(a) United Auto Workers



(b) American Federation of Teachers



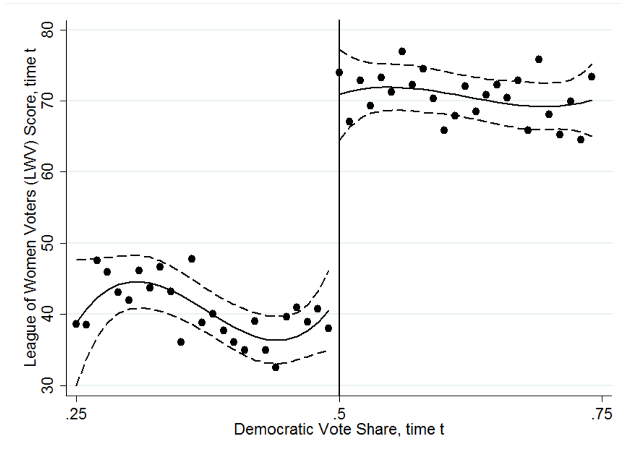
(c) American Federation of Government Employees



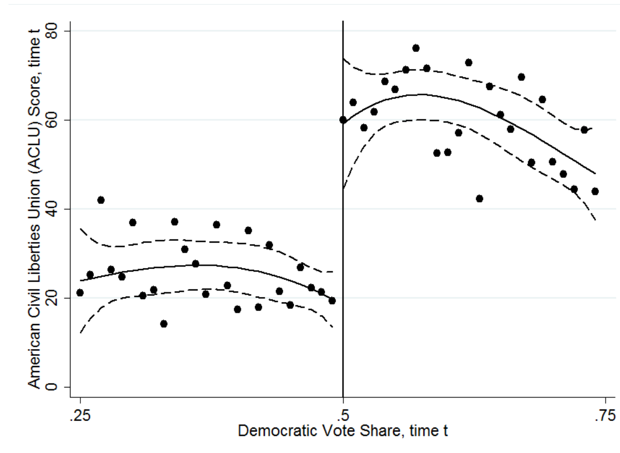
(d) League of Conservative Voters

Notes: See the notes to Figure 2. "Liberal" interest groups are defined as those where higher scores are associated with more liberal policies.

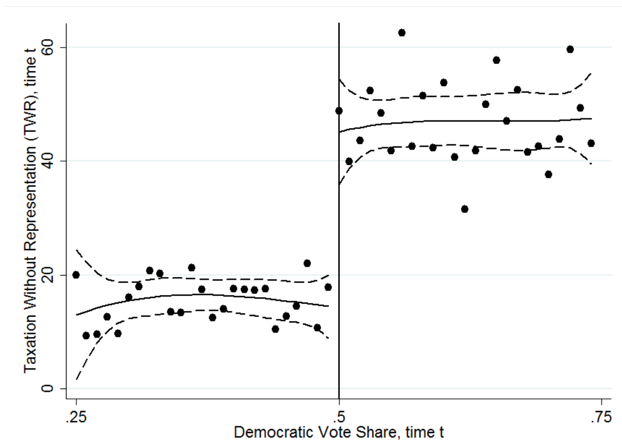
APPENDIX FIGURE 4 – ESTIMATES OF π_1 USING LIBERAL INTEREST GROUP SCORES (PART 2)



(a) League of Women Voters



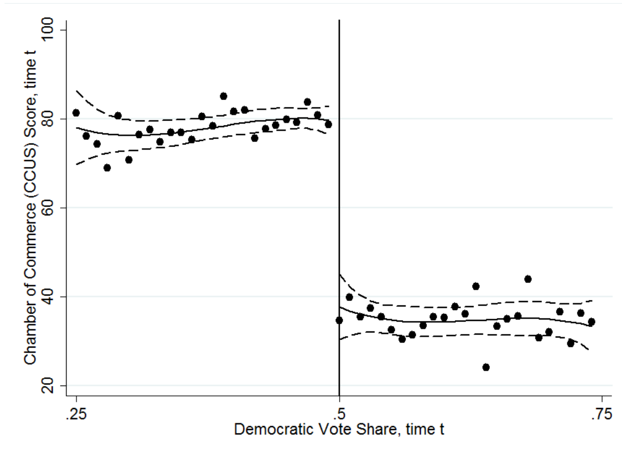
(b) American Civil Liberties Union



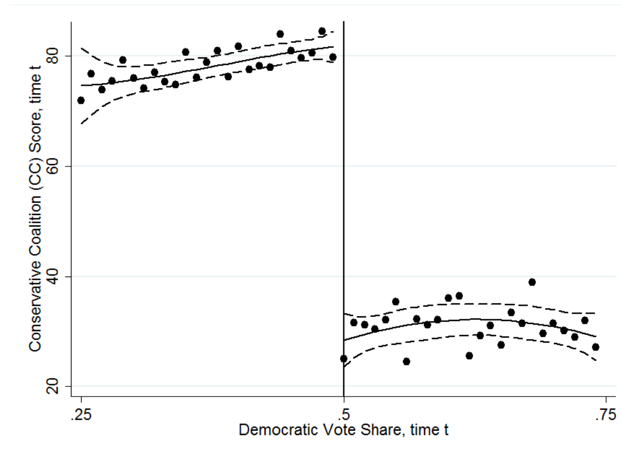
(c) Taxation without Representation

Notes: See the notes to Figure 2 and Appendix Figure 3.

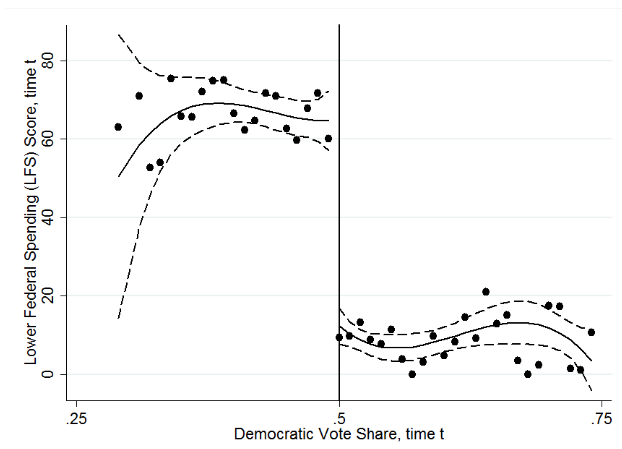
APPENDIX FIGURE 5 – ESTIMATES OF π_1 USING CONSERAVTIVE INTEREST GROUP SCORES (PART 1)



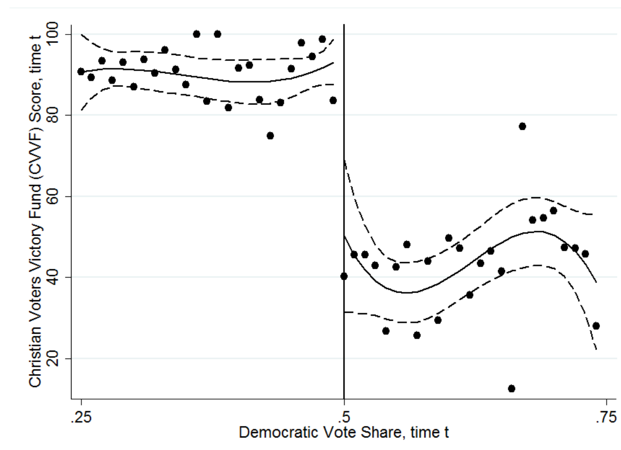
(a) Chamber of Commerce



(b) Conservative Coalition



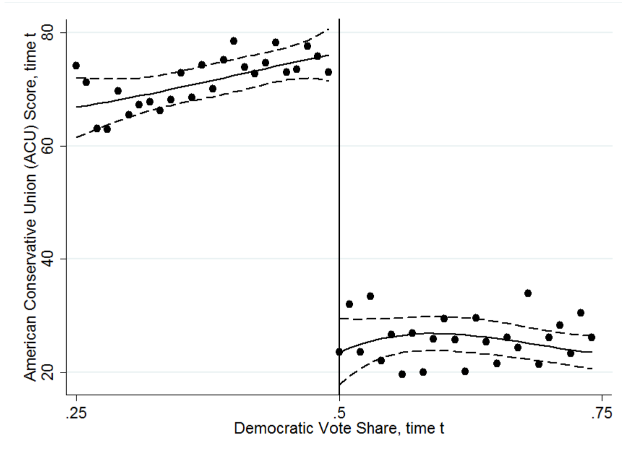
(c) Lower Federal Spending



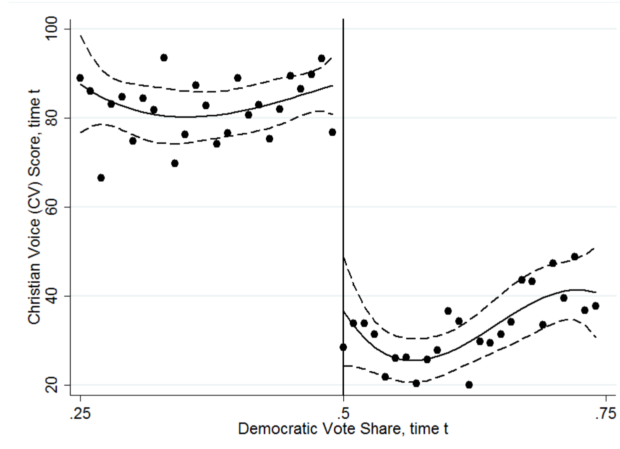
(d) Christian Voters Victory Fund

Notes: See the notes to Figure 2. “Conservative” interest groups are defined as those where higher scores are associated with less liberal policies.

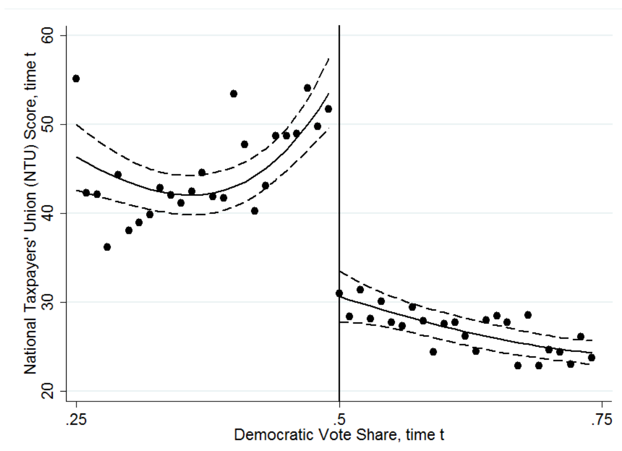
APPENDIX FIGURE 6 – ESTIMATES OF π_1 USING CONSERVATIVE INTEREST GROUP SCORES (PART 2)



(a) American Conservative Union



(b) Christian Voice

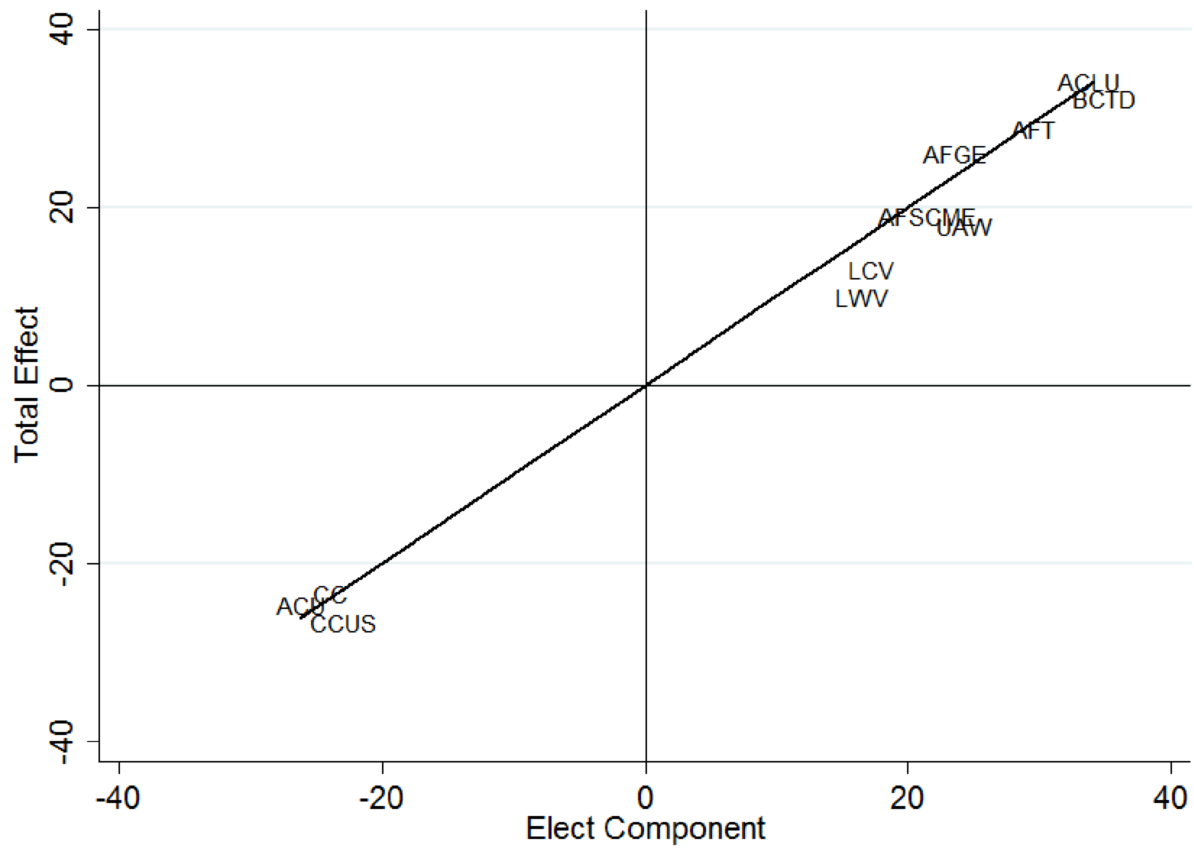


(c) National Taxpayers Union

Notes: See the notes to Figure 2 and Appendix Figure 5.

Appendix Figure 7 combines all of the estimates of the total effect and elect components from these various scores. These are estimated using the close sample methodology of Table 1. The 45 degree line represents when total effect and the elect components are identical, indicating that there is no affect component. All the scores lie near the line, indicating that LMB's results are robust to other roll call voting scores. My replication of this analysis again matches the analysis of LMB.

APPENDIX FIGURE 7 – TOTAL EFFECT (γ) VERSUS ELECT COMPONENT FOR ALTERNATIVE INTEREST GROUP RATINGS



Notes: ACLU = American Civil Liberties Union; ACU = American Conservative Union; AFGE = American Federation of Government Employees; AFSCME = American Federation of State, County, Municipal Employees; AFT = American Federation of Teachers; BCTD = AFL-CIO Building and Construction; CC = Conservative Coalition; CCUS = Chamber of Commerce; LCV = League of Conservative Voters; LWV = League of Women Voters; UAW = United Auto Workers

APPENDIX 2 - HETEROGENEITY

As noted by LMB, candidates' preferred policies could differ over time and across districts. The implicit assumption made in the earlier analysis was that the difference in policy positions between Democrat and Republican candidates was constant over time and across districts. As noted by LMB, this assumption is violated if, for example, the gap in intended policies between Democrats and Republicans from Alabama is different from the gap between Democrats and Republicans from Massachusetts. The assumption would also be violated if Democrat and Republican candidates were more similar in some elections, and were more distinct in others.

HETEROGENEITY OVER DISTRICTS

LMB show that it is possible to relax the assumption that the difference between parties' positions across districts is the same. LMB estimate the affect and elect components for three different groups: the "top" group, the "middle" group, and the "bottom" group. These distinctions are based loosely on groups discussed in Angrist, Imbens, and Rubin (1996): the "always takers", the "compliers", and the "never takers". The "always takers" are those districts that would have been won by a Democrat at time t + 1, regardless of receiving the treatment (Democrat) in time t, or not. Similarly, the "never takers" are those that would have been won by a Republican at time t + 1, regardless of treatment assignment at time t. Compliers are districts that were won by Democrats at time t + 1, only because they were quasi-randomly assigned a Democrat at time t. It is possible for the econometrician to determine ex post which group each district falls into.

Appendix Table 3 shows the transition matrix for the close elections sample, with LMB's results presented in red italics to the right of my results. Districts tend to keep the same party in power that was quasi-randomly assigned to them in the previous election, due to the incumbency effect. Those that were assigned Democrats only switch to a Republican 27.4% of the time. Those that were assigned a Republican only switch 24.2% of the time.

The top, middle, and bottom groups are constructed using the close elections sample discussed earlier, so only elections won or lost by two percentage points are included in any of these three groups. The top group is constructed by including the districts in the top right corner of Appendix Table 3 with the districts that had the largest Democrat vote share at time t + 1 and had a Democrat at time t. There are 224 districts in this group. Similarly, the bottom group includes districts from the bottom left corner of Appendix Table 3, and districts with the largest Republican vote share at time t + 1, that had a Republican at time t. There are 250 of these districts. The middle group includes the rest of the districts, of which there are 441. For a more detailed discussion of how these groups were determined, see LMB, appendix 3.

Appendix Table 4 shows the average ADA scores for each type of district, by their treatment status in the previous election. Column 4 shows LMB's result. For the top and bottom groups, any difference in the average ADA scores due to treatment status represents only the affect component. There is no elect component in this case, since the treatment had no effect on incumbency status. The results indicate no affect component for these districts. As expected, for the middle group, which includes compliers and districts that were the more marginally won, there is a large difference in average ADA scores. This is because the treatment was mostly effective. So there is no evidence that competition moderates policies even for districts that defy treatment.

HETEROGENEITY OVER TIME

The earlier estimates of the affect and elect components could differ over time, so the assumption that these estimates are constant over time should be relaxed. Appendix Table 5 shows separate estimates of these components for four different time periods, along with LMB's results. Although the estimates seem to change slightly over time, the elect component is always statistically significant, while the affect component is not.

APPENDIX TABLE 3 - TRANSITION MATRIX FOR CLOSE ELECTIONS SAMPLE

	% Democrat, time t	% Republican, time t
% Democrat, time t+1	72.6 *(72.6)*	24.2 *(24.1)*
% Republican, time t+1	27.4 *(27.4)*	75.8 *(75.9)*
Total	100	100

Notes: To the left of my estimates are LMB's estimates in parenthesis.

APPENDIX TABLE 4 - HETEROGENEITY ESTIMATES BY GROUP TYPE, CLOSE ELECTIONS SAMPLE

	Average ADA_{t+1} in previously Dem. districts	Average ADA_{t+1} in previously Rep. districts	Difference (col.(2) - col. (1))	LMB's Difference
Top Group	67.7	66.0	-1.7	*-1.7*
	(3.0)	(2.9)	(3.1)	*(3.0)*
Middle Group	18.7	66.1	47.4	*47.4*
	(1.7)	(1.8)	(1.9)	*(1.8)*
Bottom Group	21.1	16.5	-4.6	*-4.6*
	(2.2)	(2.1)	(2.4)	*(2.3)*

Notes: See the notes to Table 1. Sample sizes are 224 for top group, 441 for middle group, and 250 for bottom group. Lee, Moretti, and Butler (2004)'s results are in column 4.

APPENDIX TABLE 5 - RESULTS BASED ON ADA SCORES, BY DECADE – CLOSE ELECTIONS SAMPLE

	Total Effect (γ)	π_1	$P^D_{t+1} - P^R_{t+1}$	Elect Component	Affect Component
1946-1958	14.24	41.80	0.41	17.34	-3.10
	(4.48)	(3.36)	(0.07)	(3.09)	(2.58)
	14.2	*41.7*	*0.41*	*17.0*	*-2.8*
	(3.2)	*(2.3)*	*(0.05)*	*(4.8)*	*(4.0)*
1960-1968	23.59	49.51	0.51	25.26	-1.67
	(4.87)	(4.15)	(0.07)	(3.71)	(2.17)
	23.5	*49.5*	*0.51*	*25.2*	*-1.7*
	(3.5)	*(2.7)*	*(0.05)*	*(4.9)*	*(4.1)*
1970-1978	11.55	46.63	0.40	18.83	-7.28
	(6.22)	(4.36)	(0.09)	(4.26)	(3.99)
	11.5	*46.6*	*0.40*	*18.6*	*-7.1*
	(4.7)	*(3.1)*	*(0.06)*	*(5.1)*	*(5.1)*
1980-1996	46.84	56.70	0.77	43.47	3.37
	(5.02)	(3.79)	(0.07)	(5.67)	(2.21)
	46.8	*56.6*	*0.76*	*43.0*	*3.8*
	(3.7)	*(2.8)*	*(0.05)*	*(4.9)*	*(4.5)*

Notes: See the notes to Table 1. Sample sizes, in chronological order, are 322, 246, 183, and 164. Lee, Moretti, and Butler (2004)'s results are in red italics below mine.