DETECTING MANIPULATION IN U.S. HOUSE ELECTIONS*

Jason Snyder

First Version: August 2003

This Version: January 2005

Abstract

A fundamental tenent of representative democracy is that votes should be counted in a fair and unbiased manner. Contrary to this premise I find that in U.S. House elections this is not the case. I employ a novel approach to detect electoral manipulation by looking at extremely close elections where the outcome between any two candidates should be random. I find that in extremely close elections the incumbent wins markedly more often than would be expected. This suggests that the vote counting process is biased in favor of incumbents in a manner that is independent of the underlying incumbency advantage.

I Introduction

One of the fundamental tenets of democracy is that electoral outcomes represent the voter's will. An immediate corollary is that votes are recorded and tabulated accurately. When complete accuracy is impossible, the counting of the votes should at least be unbiased.

*This paper was written while the author was a graduate student at the Haas School of Business, UC Berkeley. The author can be contacted at snyder@haas.berkeley.edu. This paper was written while at the Haas School of Business, UC Berkeley. Meghan Busse was especially helpful with the execution of this project and her dedication of time and effort to others. I would also like to thank Henry Brady, Ernesto Dal Bo, Rui de Figueiredo, David Lee, David Levine, Siona Listokin, Lamar Pierce, Jasjeet Sekhon, Timothy Simcoe, Daniel Snow, Pablo Spiller, Oliver Williamson, Justin Wolfers, Catherine Wolfram, the participants at the Berkeley Positive Political Theory seminar and the wonderful seminar participants at the Haas School of Business Ph.D. Student Seminar for thoughtful comments and encouragement. Rui de Figueiredo and Gary King both graciously provided data for this project. I am grateful to the Bradley Foundation and the Institute of Innovation, Management, and Organization at the Haas School of Business at UC Berkeley for support. All mistakes are mine alone. This paper is dedicated to the memory of Eric "Eazy - E" Wright.

Historical evidence suggests that candidates sometimes manipulate the outcome of elections by influencing the vote counting process. My paper develops a novel test for ballot box manipulation. I examine the results of extremely close U.S. House elections, and find that incumbents win significantly more often than one would expect.

There are two major difficulties in detecting electoral tampering. The first is that it is hard to disentangle the well known incumbency advantage from manipulation of the election process. Second, since manipulation is by definition clandestine it is hard to gather evidence that the outcomes are biased. Acts such as coercion of election workers, destroying ballots, and placing friendly individuals within election commissions are difficult to verify since they are done under the veil of secrecy.

The solution I propose is to search for indirect evidence of systematic manipulation. I argue that the outcome of an extremely close election should resemble a random process. This is because the amount of noise in the voting system is greater than the handful of votes that make the difference in the outcome. So, if the ballot box is not tampered with, incumbents will win approximately half of these very close elections. On the other hand, if cheating is occurring, then incumbents may win these close elections more often than challengers. Since incumbents have more experience, greater resources, and are more insulated within the local political infrastructure than their opponents, they would be more likely to succeed in tampering with the tabulation process. I suggest that the likelihood of manipulation is highest in precisely those races that are the closest, since the outcomes hinge on a very small number of votes. This discontinuity in the returns to manipulation generates a natural experiment to detect electoral fraud.

In a sub-sample of U.S. House elections from 1898 to 1992, where the difference between winning and losing is on average 155 votes with the average number of ballots cast in these contests exceeds $110,000$, incumbents win approximately 58% of the elections. Statistical tests reveal that this is significantly different from the null hypothesis of a random outcome. I also derive a series of robustness tests that show this effect is in fact detecting bias in the vote counting process and not the underlying incumbency advantage. I interpret this as evidence of manipulation of outcomes in favor of incumbents.

This work contributes to the burgeoning empirical literature on corruption. This literature has introduced a set of tools to enable researchers to detect the existence of corruption in a variety of settings. A fundamental technique that these papers use is to examine the distributions of outcomes that are a result of a supposedly random data generation process and look for anomalies in the distribution of data that suggest manipulations. For example Duggan and Levitt [2002] make use of a similar discontinuity design to identify the existence of fraud in Sumo wrestling while Wolfers [2003] looks at the prevalence of corruption in college basketball. Degeorge, Patel, and Zeckhauser [1999] use discontinuities in thresholds to identify the prevaliance of earnings management by U.S. firms.

There has been a significant amount of work done on the impact of incumbency on electoral outcomes. Much of this work has focused on identifying the causal impact of incumbency on subsequent electoral outcomes. Gelman and King [1990], Levitt and Wolfram [1997], and Ansolabehere, Snyder, and Stewart [2000] are some representative papers in this tradition. Additionally this work is very closely related to Lee [2003]. In this important contribution Lee recognizes the usefulness of using the institution of majority rule voting to create a compelling counterfactual to estimate the incumbency advantage in elections. He looks at close elections in period t and then looks at the outcomes in the same district in period $t+1$. He finds that if a candidate barely won an election in period t their margin of victory in period $t+1$ is significantly higher. This paper differs from Lee in that I don't look at subsequent election $t+1$, but rather the current election where conditional on the election being extremely close incumbency advantage should be washed away absent manipulation.

This paper proceeds as follows: First I discuss some of the anecdotal evidence in favor of electoral fraud. I will then develop my identification strategy and formal empirical methodology. Finally I discuss my results, offer some conclusions and suggest avenues for future research.

II Biasing Elections

Conceptually I separate biasing elections into two categories: *ex-ante* to the election day and *ex-post* after the election day. *Ex-ante* manipulations consist of voter intimidation, vote buying, or any activities that would deter or encourage voters using extra-legal means. *Ex-post* manipulation occurs after the ballots have been cast, prime examples include "finding" new voters, destroying ballots, and strategic tabulation. To further illustrate how these mechanisms are used in American politics, I consider two cases from American political history.

II.1 The 1948 Texas Democratic Senate Primary

One instance of both *ex-ante* and *ex-post* manipulative behavior comes from the Texas Democratic primary for Senator held in 1948 between Coke Stevenson and Lyndon Baines Johnson. In a poll released one week before the primary the challenger Johnson was trailing the incumbent Stevenson by a 48% to 41% margin with 11% undecided. Given these numbers it would have been improbable for Johnson to win the primary without either a phenomenal turn of events or the use of extra-legal means to influence the outcome of the election. One of Johnson's closest associates Edward Clark stated that "Campaigning was no good anymore... We had to pick up some votes." [Caro 1982] Johnson and his allies thus poured money into the hands of judges, local bosses, and other public officials to provide additional inducement to motivate the voters of Texas. In one instance, Johnson gave the Streets Commissioner of San Antonio "a thousand dollars in one-dollar bills for expenses of poll-watchers." In short, Johnson was using *ex-ante* manipulations of the democratic process to influence the electoral outcomes.

The Johnson campaign did not stop with these *ex-ante* mechanisms; it augmented these strategies with *ex-post* attempts to influence the process by which votes were tabulated. When the Texas Elections Bureau closed on election day, Lyndon Johnson was trailing Stevenson by 854 votes out of nearly a million votes cast. Given the importance of the election it was not surprising that there were both opportunities and incentives to manipulate the vote counting process. A regional newspaper described the process of calculating the election totals as one where error was present "in counting, copying and tabulating, the votes pass through the hands of eight different groups, between the voter and the final declared result... With a million votes running the gauntlet of 'the Human Element' eight different times, there will always be mistakes regardless of the honesty and good intentions of the humans involved." [*The State Observer*] As Caro writes:

> "Few persons familiar with Texas politics, though, were confident of the universality of the 'honest and good intentions of the humans involved'; there was common knowledge in the upper levels of Texas politics of the precincts that were for sale, the 'boxes' in which the County Judge wouldn't 'bring the box' (report the precinct totals to the Election Bureau) until the man who paid him told him what he wanted that total to be. In close elections, precinct results were altered all through the state." [Caro 1982]

Four years after Lyndon Johnson's death in 1977 the election judge of Duval County Luis Salas admitted to out-right fraud in the process of tabulating the election results. "If they [the votes] were not for our party, I made them for our party" [Caro 1982]. As a result Johnson was able to overcome the electoral day deficit of 854 votes and win the primary by a mere 87 votes out of nearly a million ballots through the widely acknowledge use of both *ex-ante* and *ex-post* strategic manipulations. As a consequence of this electoral manipulation Johnson easily won the general election and went on to have a significant impact on the landscape of American politics. Although in this example Johnson was in fact the challenger, it is a useful example because it illustrates some of the mechanisms by which *ex-post* electoral manipulation can occur.

II.2 The Florida Presidential Election in 2000

As perhaps the most famous close election in American history, this case provides stark evidence of many of the *ex-post* and *ex-ante* activities that biased the voting process and tabulation. The 2000 presidential election was one of the most highly contested in American history. After the election on November 8th, 2000 the votes for George Bush and Al Gore were so close that the officials tabulating the results could not come to conclusive result as to who won the election in the state. Given that Florida was the deciding state in the Presidential election, an unprecedented amount of scrutiny was paid to this election. This example clearly contrasts with the Johnson case. In the Johnson case the behavior was clearly rooted in nefarious means. This case on the other hand illustrates how other non-voting institutions can influence electoral outcomes after all the votes have been cast and is certainly not meant to suggest that this election was in any way stolen.

Numerous *ex-ante* events skewed the voting, which led many observers to question the fairness of the process. Perhaps the most famous of these *ex-ante* mechanisms was the design of the ballots in the county of West Palm Beach, Florida. In this county Pat Buchanan won a surprising .8% of the vote when he was expected to win at most .3% of the vote. Upon further inspection, the design of the ballots in West Palm Beach was deemed by many to be confusing and misleading, leading some Gore voters to choose Buchanan or double vote for both Buchanan and Gore.[1] Although it is impossible to be conclusive, some empirical analysis has suggested that it is extremely likely that Gore would have won more than half of those potential mistakes [Brady 2000], and as a consequence the election. In addition to misleading ballots, other problems such as antiquated voting machines and poorly trained poll watchers were also said to contribute to the failure to count all of the ballots [New York Times, Nov. 12th, 2001].

As highlighted by the Florida election, the *ex-post* attempts to manipulate the manner and method by which the votes were tabulated by the two candidates played an important role in determining who won the election. Since the tabulation process was large, complex, and fraught with many unforeseen contingencies it is easy to see why both parties tried to bias the chaos in their favor. A year-long study conducted by the New York Times found that the election could have easily been swayed by the definition of a valid vote. For instance

[1] Although this was an *ex-ante* manipulation it is doubtful that it occur as a result of a deliberate strategy.

if only full punches counted on punch card ballots Gore would have won by 134 votes [New York Times Nov. 12th, 2001]. Had ballots with chads that were detached at three corners counted, George W. Bush would have won by 2 votes. This is just one of the many possible counting scenarios where the determination of the process of *ex-post* counting process would influence the final outcome.

As a consequence of the ambiguity in the counting process, both camps engaged in strategic behavior in order to bias the tabulation process. One of the key elements of Gore's strategy was to lobby the courts for selective recounts in counties that he felt would favor him. In the aftermath of the New York Times recount "the numbers reveal the flaws in Mr. Gore's post-election tactics and, in retrospect, why the Bush strategy of resisting county-by-county recounts was ultimately successful" [New York Times Nov. 12th, 2001]. One of the strategies central to Bush's campaign was to push for flawed ballots from overseas servicemen to be accepted, the majority of which were thought to be in favor of Bush. While Bush was aggressive in pushing for the acceptance of those ballots, Gore was hesitant to fight against their acceptance for fear that he would isolate the military community. These and other fights that took place in the judicial courts and the court of public opinion clearly demonstrate the importance of *ex-post* political strategic behavior in trying to influence the final outcome.

III Identification Strategy

The fundamental problem with identifying the existence of electoral manipulation is that by its very nature it is a clandestine activity. Disentangling it from other factors that go into the final electoral outcome is a problematic task. One approach to this problem is to look at the distinction between *ex-ante* and *ex-post* manipulations. If it were possible to control for all *ex-ante* factors, legitimate or otherwise, and identify a group that was more likely to be able to engage in *ex-post* manipulations of electoral outcomes, then it would be possible to detect the existence of corruption. Ideally if the *ex-ante* votes were randomly assigned then it would be possible to test for the existence of *ex-post* manipulations by looking to see if the vote totals deviate from the random assignment process. The key assumption is that a fair *ex-post* tabulation process should unbiasedly filter the *ex-ante* voting process to final outcomes.

To replicate this ideal experiment with non-experimental data, I take two critical steps to identify the existence of bias. The first step I take is to use incumbency to proxy for variation in the relative ability of candidates to have elections biased in their favor. Conditional on the random assignment of votes on election day, if the incumbent received significantly more votes than expected from this assignment process, this would present evidence in favor of manipulation. Once election day has occurred, candidate characteristics should be irrelevant and the tabulation process should filter the votes in an unbiased manner. If elections were not biased in favor of incumbents, the empirical results would show the random assignment process unbiasedly filtering through the *ex-post* tabulation. Recall from the above examples that incumbency is not the only dimension along which one can have power to exert influence over electoral outcomes. In the Johnson case, deep pockets substituted for political connections in turning the Senate primary in Texas in his favor. Thus incumbency is only a crude proxy and any finding of bias from the data will occur despite measurement error.

The second step needed to identify the presence of corruption is to replicate the random assignment process in a non-experimental setting. Since votes are deliberate choices made by the electorate based on candidate qualities, assuming an exogenous voting process is implausible. However, within the voting process it is widely acknowledged that there is substantial randomness. Voter turnout and decisions are often influenced by events beyond the control of candidates. Changes in the weather, a recently broken heart, or traffic jams are just a few reasons why we might expect randomness in the voting process. Given this logic we would expect that for a small enough part of the distribution, a uniform distribution would provide a good approximation of the underlying distribution. Figure I.A illustrates this point nicely. Here one can see that even though the density function is sloping upwards local approximations to the density function can be made using uniform distributions A and B. Conceptually as the narrowness of A and B increases they become better and better approximations.[2]

The next step is to look for a part of the distribution where it would be expected that the incentives to perform *ex-post* manipulations would be the highest. The majority rule voting process used in House elections suggests that the incentives should be highest at the 50% vote share, the boundary where a candidate could just barely win or just barely lose. If one were to look at a very small range close to the 50% point, the distribution of incumbent vote shares should be reasonably approximated by a uniform distribution. If going into the election the final vote is going to be very close then in expectation the final winner should be random since the noise in the process will overwhelm any inherent incumbency advantage. Fifty percent of the time the incumbent candidate should win and fifty percent of the time the challenger should win. *Ex-ante* both candidates have very imprecise information of exactly who will vote for them and how many people will turn out. However *ex-post* the information about the close electoral outcomes rapidly comes into view as the process of tabulating the votes takes place. If *ex-post* manipulations were occurring, then a discontinuity in the distribution will occur at the 50% cutoff. One would see that the final vote shares received by incumbents would cluster to the right of the 50% point, the difference between just barely winning instead of just barely losing supporting the hypothesis that votes are *ex-post* being changed to favor the incumbent. Figures I.B and I.C demonstrate these ideas. In Ib it is easy to see at the discontinuity the area under the two boxes is substancially different. Furthermore Ic shows that relative to other points on the density function the jump and the discontinuity is larger.

IV Methodology

Although this paper deals with the case of electoral manipulation, this methodology provides a general approach to detecting discontinuities in functions over a complete domain. In the case of electoral manipulation, the relevant density function is given in figure 1, which is the

[2] For expositional purposes I made the uniform approximates markedly larger in these figures than in the actual application.

distribution of vote shares received by incumbents over all House elections from 1898-1992. As is evident from this graph there is an upward trend in this density function, so to argue that there is a discontinuity at the majority rule point it is necessary to show that one is not simply picking up the underlying trend in the density. My approach proceeds in two stages. First I derive a simple test for a discontinuity under the counterfactual assumption that density function is uniformly distributed close to the suspected discontinuity. This stage is useful in that it provides a sense of how the data behaves and what the magnitude of the discontinuity is. In the second stage I relax the assumption of uniformity of the distribution by employing techniques of non-parametric density estimation following Porter [2003] and McCrary [2004]. These techniques essentially allow for the estimation of the discontinuity taking into account the trends elsewhere in the distribution.

IV.1 The Intuitive Approach

To derive the relevant test I start with benign assumptions about the form of the distribution function and use this to develop a refutable test statistic. For any given cumulative density function $f(x)$ I define a discontinuity in the following manner:

Definition 1 *There is a jump discontinuity in* $f(x)$ *at* $x = x_m \Leftrightarrow$ *The following conditions hold: For* $\delta > 0$; $\lim_{\delta \to 0} f(x_m - \delta) = K_1$; $\lim_{\delta \to 0} f(x_m + \delta) = K_2$; $K_1 \neq K_2$ *and* $f(x_m) \in \{K_1, K_2\}$

This definition describes a jump discontinuity in a function. As opposed to using a more general definition of a discontinuity, I use this specific class of discontinuities since it is the only relevant class by virtue of the fact that the cumulative distribution function is nowhere decreasing. Furthermore for analytical ease I assume over the range $[x_m - \delta, x_m + \delta]$ that $f'(x) \geq 0$ wherever $f(x)$ is differentiable[3] and that $f(x) > 0$ over the domain. Additionally I will assume that the number of discontinuities in $F(x)$ is finite. Using this I establish the following bounds:

[3] This is done without loss of generality since over a small enough domain $[x_m - \delta, x_m + \delta]$ $f(x)$ will be monotonic. The results hold as well if $f'(x) < 0$. Assuming $f'(x) \geq 0$ is done simply to make the bounds easy to interpret.

(1) $\int_{x_m-\delta}^{x_m} f(x)dx \leq \delta f(x_m)$

(2) $\int_{x_m}^{x_m+\delta} f(x)dx \leq \delta f(x_m+\delta)$

For notational convenience let $A = \delta f(x_m)$ and $B = \delta f(x_m + \delta)$ the upper bounds on 1&2 respectively. This leads to the following theorem.

Theorem 1 $\lim_{\delta\to 0}[\frac{B}{A+B} - \frac{A}{A+B}] = 0 \Rightarrow$ *No Jump Discontinuity*

Proof. To proceed consider the contrapositive of the theorem: *A Jump Discontinuity* $\Rightarrow$ $\lim_{\delta\to 0}[\frac{A}{A+B} - \frac{B}{A+B}] \neq 0$. Trivially computing the limit $\lim_{\delta\to 0}[\frac{B}{A+B} - \frac{A}{A+B}] = \lim_{\delta\to 0}[\frac{f(x_m-\delta)-f(x_m)}{f(x_m+\delta)+f(x_m)}] = \frac{K_2-K_1}{K_2+K_1} \neq 0$. This proves the contrapositive and as a consequence the theorem ■

This result says that for a small enough δ that the relative area under $f(x)$ over $[x_m-\delta, x_m]$ and $[x_m, x_m+\delta]$ should be approximately equal. If, on the other hand, there was a jump discontinuity at x_m then the area to the left of x_m would be relatively smaller than the area to the right of x_m.

Since it is impossible to directly observe the functional form of the distribution, but rather a finite number of draws from the distribution it is necessary to transform the theorem into a test statistic. As a consequence to test for a discontinuity at x_m I must arbitrarily choose a finite number $\delta > 0$ (hopefully small) and empirically observe a finite number n observations drawn from the range $(x_m-\delta, x_m+\delta)$. Unfortunately non-parametric econometric techniques provide little guidance as to what the optimal δ should be so the subsequent choices are arbitrary (see DiNardo and Tobias [2001]). The fundamental trade-off is that as one shrinks δ, the magnitude of the potential bias introduced by picking up underlying trends in $f(x)$ decreases, but so does the power of the subsequent test. Furthermore, l observations fall within $(x_m - \delta, x_m]$ and u fall within $(x_m, x_m + \delta)$ where by definition $l + u \equiv n$. As a consequence of the theorem above if $|\frac{l}{n} - \frac{u}{n}| \to 0$ as $\delta \to 0$ this would imply that $\frac{l}{n} \approx \frac{u}{n}$ and that there is no jump discontinuity present. A natural implication this is that the counterfactual distribution to test for the existence of a jump discontinuity is $l = \frac{1}{2}n$ and $u = \frac{1}{2}n$. That is to say, half of the total draws over $(x_m - \delta, x_m + \delta)$ should fall to the left of x_m and half should fall to the right. *This exactly describes the binomial distribution, which* *is the local approximation of any distribution.* Without loss of generality we can characterize the binomial density for finite n:

(3) $b(u|n,p) = \frac{n!}{u!(n-u)!} p^u (1-p)^{n-u}$

As a consequence of above reasoning if δ was sufficiently small then we should see $p = .5$. To test whether the probability that we would observe u or greater from $b(u|n, \frac{1}{2})$ we simply sum up the associated probabilities (since it is discrete) and arrive at the following test statistic:

(4) $test\ statistic = \sum_{i=u}^{n} \frac{n!}{i!(n-i)!} (\frac{1}{2})^n$

A small test statistic implies that the likelihood that u and l came from a binomial distribution where $p = .5$ is small. Since the binomial distribution was endogenously derived from very mild assumptions on the form of the function, this implies that the true probability is greater than .5 and as a consequence there is a jump discontinuity in the density at x_m.

IV.2 Non-Parametric Discontinuous Density Estimation

The next method involves using a non-parametric density estimation as discussed in Hahn, Todd, and van der Klaauw [2001] and Porter [2003]. Whereas the intuitive approach gives an understanding of what the behavior of the distribution is close to the density, it does rely on the strong identification assumption that the distribution local to the suspected discontinuity is uniform. Using non-parametric density estimation it is possible to relax this assumption by using the data to estimate the exact functional form of the distribution absent parametric assumptions. Essentially this answers the question of whether this is a discontinuous jump or rather just a part of the distribution where the slope is increasing?

Using the approach advocated in McCrary [2004], I construct an estimate of a discontinuous density function by first creating a finely binned unsmoothed histogram and then applying local linear smoothing to the data created from the histogram. Superimposing the local linear smoother on the unsmoothed density function is useful for giving the reader a sense of what the data looks like in relation to smoothing choices made by the researcher. The remainder of the section is a brief outline of the key points of the approach in McCrary [2004].

The first step of this process is to construct a histogram of the density function. This construction is similar to other density functions save one crucial difference: the midpoints of the bins are chosen such that they straddle the suspected discontinuity. In the case at hand for bins of the size 0.05% of the vote, one can construct midpoints for each of these bins as follows: $\{.00025, .00075, .., .49975, .50025, ..., .99975\}$. A graphical representation of the figure consists of a plot of the midpoints with the associated number of elections contained in the bins. For instance a pair (X_j, Y_j) equal to $(.48725, 15)$ means that there are 15 elections contained in the bin $[.487, .4875]$.

Subsequently in the second step, the local linear regression smooths over these constructed bins to provide a non-parametric estimate of the discontinuity in the density function. The "second step smoother" is characterized by the following regression function:

(5) $L(\alpha, \beta, r) = \sum_{j=1}^{J} \{Y_j - \alpha - \beta(X_j - r)\}^2 W_j$

where the weighting function is given by $W_j = K(\frac{(X_j - r)}{h})\{1(X_j \geq x_m)1(r \geq x_m) + 1(X_j < x_m)1(r < x_m)$. Here r is the evaluation point on the domain while α and β are parameters that are *locally* maximized. Recall also that x_m is the point where there is a suspected discontinuity and h is the bandwidth for the kernel. As Fan and Gijbels [1996] show the triangular kernel $K(t) = \frac{3}{4}(1 - |t|)_+$ is boundary optimal which makes it a good candidate for evaluating discontinuities in density functions. This weight means that there are two density functions being estimated, one to the left of the suspected discontinuity and one to the right. If the two density functions are not equal at the suspected discontinuity we can not reject the hypothesis of no discontinuity. On the other hand, if there is no discontinuity in the underlying density this estimator converges to a conventional kernel density estimator.

Using these estimators it can be shown that one can obtain estimates $\hat{f}_L$ for density function and $\hat{\delta}_L$ for the Eicker-White heteroskadastic corrected standard error directly to the left and to the right of the suspected density function. These estimates can be characterized by $\hat{f}_L(c^-)$ and $\hat{f}_L(c^+)$ with associated standard errors $\hat{\delta}_L(c^-)$ and $\hat{\delta}_L(c^+)$ respectively. This leads to a simple computation of the relevant t-statistic:

$$\textbf{(6)}\ \ test\ statistic = \frac{|\hat{f}_L(c^-) - \hat{f}_L(c^+)|}{\sqrt{\hat{\delta}_L(c^-)^2 + \hat{\delta}_L(c^+)^2}}$$

The two critical tuning parameters that must be selected by the researcher are both the initial binsizes used in the histogram and the bandwidths used for the local linear smoother. Unfortunately there is little to tell us to what the optimal bandwidth / binsize parameter is. Asymptotically it is know that $\frac{binsize}{bandwidth} \to 0$ yet all this tells us is that the binsize should be somewhat smaller than the bandwidth.[4] McCrary's suggestion of selecting binsizes that seem to provide the most detail of what the density function looks like close to the suspected discontinuity seems to be a sensible suggestion. Of course, as the binsizes become too fine one rapidly loses any sense of the what the density function looks like. For instance, in this data if the binsize was .000001 one would usually only find zero or one election per bin. Thus there is some art to selecting the binsize. Since the second stage regression will smooth out the coarseness in the data the histogram is useful for visually displaying the data. Even after the binsize has been specified the problem of specifying the bandwidth still remains. Since most selections means are *ad hoc* I show the results using a variety of plausible bandwidth specifications to demonstrate robustness to this choice.

Data & Results

Data

To test for electoral manipulation I look at Congressional House of Representatives election results from 1898 to 1992. This data has been graciously provided to the public by Gary King in ICPSR data set 6311. This data contains information on vote totals, incumbency, party affiliation, and district locations. For more details on this data set see Gelman and King (1990). This initial data consists of 21, 045 elections. After removing elections in which

[4] McCrary also points out some useful rules of thumb which can guide researchers as to what specifications to chose.

there was no incumbent, more than one incumbent, the elections where the Democrats and Republicans weren't the top two vote getters and elections where at least one party received zero votes the sample is reduced to 13,948 elections.[5] I then construct vote shares by taking the votes received by the incumbent candidate and dividing it by the total votes received by the top two candidates. Within this sample the incumbent wins 89.8% of the elections, with an average margin of victory of 62.5%. The average number of votes for the top two candidates is 114,596.

Main Results

Before turning to the main results, a close examination of figure II.B and II.C is warranted. Both of these figures are views of the distribution close to the majority rule cutoff point. Figure IIb is mildly suggestive that there might be a discontinuity in the data; however the bins are quite large and could potentially obscure the important relationships in the data. Figure IIc provides a graphical representation of the density curve with markedly finer bins. Somewhat strikingly, there appears to be a structural break in the data almost exactly at the majority rules point.

To implement the intuitive approach outlined above I chose an initial $\delta = .005$ and select the midpoint $x_m = .5$, the cutoff between where a candidate just barely wins versus just barely loses. Given that the initial choice of δ is arbitrary, all that is necessary is to show that there is a discontinuity at $\delta = .005$. The analysis works for other δ and the results for $\delta = .003$. Elections to the left of the majority rules point are coded as a loss while elections to the right of the majority rule point are coded as a win. They key results are reported at the top of table II & III. On average these very close elections are decided by 155 votes out of 108,025 ballots cast for the top two candidates from either party. In these elections the incumbent actually wins 183 elections out of 315, whereas the incumbent would be expected to only win 157.5 of the elections if the draw was random. The incumbent wins 58.1% of the

[5] This point is critical for much of non-parametric analysis. Leaving in the large number of uncontested elections leads to mass points in the density function. Conventional non-parametric estimators do not deal with these points well and it is hard to argue why there inclusion is relevant for estimation at a distinctly different part of the distribution.

elections. Using test statistic (4) I find that the probability of this observation coming from a random draw is .2%, which would be refuted at any conventional confidence interval. This is a clear piece of evidence that the *ex-post* process does not filter the *ex-ante* votes without bias. This is consistent with the final vote totals being manipulated in favor of incumbents. It is important to note that this 58.1% incumbent win rate merely measures the ability of incumbents relative to challengers to manipulate elections rather than the absolute level of manipulation. For instance, this result would be consistent with the scenario that every candidate tries to bias the system every time the election is close but incumbents are slightly better at it. This result can therefore be thought of as a lower bound on the absolute level of electoral manipulation that is taking place.

Generally the non-parametric smoother returns results that are consistent with the notion of a discontinuity at the majority rules point. Recall that the major objective of this technique is to relax the assumption that local to the discontinuity the distribution should be uniform. One way of thinking about what this non-parametric technique tells us is that not only is there a jump at the majority rules point, it is also a large jump relative to other jumps in the density function. Figures IIIa-IIIf provide a picture of how the non-parametric smoother is behaving close to the majority rules point under various specifications. What these graphs show is that the result is robust to how sensitive the model is specified. In III.A there is a great deal of smoothing involved, yet a discontinuity is still present and when markedly less smoothing is used in III.F there is still a discontinuity present. Table IV also shows the estimated impact of the discontinuity for a variety of binsize / bandwidth specifications. For more interpretation and explication of this table the interested reader is referred to the appendix. I generally find that for reasonable specifications there appears to be a statistically significant jump in the discontinuity, although like most other non-parametric density estimation procedures there is considerable sensitivity to the binsize/bandwidth specifications.

Is Ex-ante Incumbency Advantage Naturally Increasing in Close Elections?

One alternative hypothesis to this analysis is that there is an underlying advantage to incumbency that is even more relevant in close elections. For instance, incumbents might know when they are in a close election and have certain advantages at winning these elections relative to challengers. If an incumbent has more money on election day in a close election that hinges on a few hundred votes then she might be able to spend that money to push herself to victory. This perspective suggests the empirical finding is a result of an incumbent's campaigning advantage in close elections rather than in the *ex-post* tabulation process. This distinction can be thought of in a simple manner: the density function is simply increasing at a very fast rate compared to other parts of the distribution *but is still locally continuous.*

For this alternative hypothesis to be plausible two conditions must be met: 1) The incumbent knows when they are in a close election and 2) Conditional on knowing that they are in a close election, the *ex-ante* actions that an incumbent can take will influence the final vote outcomes. Referring to figure IV one can observe that distribution B would correspond to the case where the *ex-ante* advantage is increasing in close elections. Distribution A is consistent with only *ex-post* advantages being of importance, since the noise in the electoral process overwhelms any underlying incumbency advantage local to the discontinuity. If distribution B were in fact the true distribution then a natural consequence would be that the slope of $f(x)$ would be increasing over a small range directly to the left of the majority rules cutoff. If there was a local incumbency advantage absent *ex-post* manipulation I would expect the following logic to hold: incumbents are more likely to lose by a little than by a lot. If an incumbent were truly skilled at getting out the vote in a close election then the incumbent would be more likely to obtain a vote share around 49.9% as opposed to a vote share close to 49.8%. One logical caveat is in order: looking directly to the right of the majority rules point and observing an increasing slope is theoretically ambiguous. An increasing slope could be evidence of a local incumbency advantage, but it could also be evidence of an incumbent manipulating the final results slightly further away from the majority rules point.

There are good reasons to believe that given a candidate decides to manipulate the results he would change 100 votes as opposed to 10 votes. This conceptual difficulty becomes much less pressing when simply looking to the left of the majority rules point.

Two crucial pieces of evidence suggest that the density is not increasing to the left of the majority rules point. Referring to figures III.D it appears that to the left of the majority rules point the slope of the density function is actually decreasing. Since the bandwidth in this specification is sensitive to changes close to the discontinuity, it demonstrates that directly to the left of the discontinuity the slope does not appear to be increasing. When examining the data I find that there are 137 elections in the bin between $[49, 49.5]$ yet only 132 elections within the bin between $[49.5, .5]$. Figures V.A-V.E points to further evidence that is consistent with this view. When running the non-parametric density estimation at points to the left of the majority rules point, I find little evidence of a discontinuity. As shown in figure Va-Ve running the estimation procedure with the discontinuity specified at 49.5% to 49.9% I do not find evidence of a significant discontinuity. At these points the density function appears to be continuous.[6] Only when I run the procedure at the majority rules point do I find statistically significant evidence of a discontinuity. When I run the procedure to the right of the majority rules point I do find a significant discontinuity, but for the arguments given above the implication on this result is markedly more ambiguous. Overall, a purely *ex-ante* electoral advantage would require that incumbents only have an *ex-ante* advantage in very close elections when the race is almost precisely at the majority rules point.

Furthermore, the previous argument rested in the assumption that argument (1) was true. Anecdotal and empirical evidence from the polling literature suggests that this is not the case. As Leigh and Wolfers [2003] point out polling technology is extremely imprecise. For instance, in the 2001 Australian Parliamentary election John Howard's party won a narrow election getting 50.5% of the votes for the top two parties.[7] ACNielsen, Morgan, and

[6] Naturally, though the difference is not significant, there seems to be some visual evidence of a discontinuity at 49.9%.

[7] As a consequence, since Howard's party won the majority he became Prime Minister. Thanks to Geoff Edwards for clarifying some details on Australia's electoral system for me.

Newspoll (three prominent election forecasting groups) predicted that Howard's party would win 52%, 45.5%, and 53% of the votes respectively. In close elections where the accuracy of the polls matter most, these polling companies were highly inaccurate.[8] Additionally, this election was undoubtedly more important to the electorate than almost any House election. Considering that most of my sample comes from before the birth of modern information technology, it is probable that the polling technology was less accurate than in the Australian case. Since polling technology is fairly imprecise one can conjecture that a candidate will put forth their maximal effort on election day regardless of what the polls indicate. This suggests that again the noise in the election and polling process should overwhelm any local *ex-ante* incumbency advantage that candidates have in very close elections.

Additional Results

It would be of interest to look at how the covariates such as experience and party impact *ex-post* manipulation. However since this is a non-parametric estimation technique, one quickly runs into the curse of dimensionality. Examining the interaction of incumbency and other variables of interest such as experience severely limits the power of the tests. In general, I find that when I cut the sample in half as is required by most of the interesting covariates I lose most of the power to conduct the non-parametric tests outlined above. As such I rely on the intuitive approach used in the methodology and caution the reader against drawing strong conclusions. These results are all contained in tables II and III.

One variable of particular interest is the time trend that exists: is corruption an artifact of the past or is it still prevalent today? Surprisingly we see that the existence of corruption seems to be *increasing* over time. When the data is divided into the following intervals $[1898, 1924], [1926, 1950], [1952, 1974]$, and $[1976, 1992]$ I find that the mean percentage of close elections won by incumbents is increasing by era. In table II I find that the mean number of elections won by incumbents between 1898 and 1924 is 47.2%, which is not rejected by the test statistic. However all of the other time intervals exhibit a significant discontinuity.

[8] Interestingly though Leigh and Wolfers find that betting markets tended to be better predictors than polling.

Testing against the binomial distribution all of the results for close elections are significant for the intervals after 1924. One plausible rationale for this increase in manipulation is that technology such as automobiles, computers, etc. have allowed for greater efficiency and coordination within the vote tabulation process, thus requiring fewer people to tabulate the votes. As Shleifer and Vishny [1993] point out, a key component of corruption is secrecy, so as a consequence there is a decline in the costs of *ex-post* manipulations as fewer people will be needed to turn an election over to an incumbent. Additionally, one could imagine substitution towards *ex-post* manipulation as the costs of *ex-ante* manipulation increase.

Another issue of interest is the role that political experience plays on how incumbents manipulate elections. First I construct an experience variable indicating how many terms the candidate has been in office. The results indicate that in absence of any other covariates, the probability that an incumbent will win a close election increases with experience. When a candidate has a single term of experience his probability of winning a close election is between $52\% - 54\%$ in both bandwidths. However, as I look at when a candidate has three or more terms of experience, his probability of winning the election increases to $62\% - 65\%$ and is statistically significant. These numbers are suggestive of the fact that as a candidate becomes more experienced he is more able to manipulate close elections in his favor.

Also of interest is the role that political parties play in the *ex-post* process of electoral manipulation. To test for this party effect I change the unit of observation from vote share received by incumbents to vote share received by Democrats and proceed with the analysis as presented above. Not surprisingly I find that in elections where there is no incumbent, the percentage of victory is evenly split between Republicans and Democrats. Neither party is more manipulative relative to the other. When Republican incumbents run I find they win approximately 55.9% percent of the the elections, while Democrats win approximately 61.1% of the elections. Though Democrats appear to be more prone to *ex-post* manipulation, testing for equality of the means for both of the samples fails to reveal a significant difference between the two means.

Another set of estimates I ran was to look at the impact of state political conditions variables on close elections.[9] I find no evidence that the party of the governor nor the composition of the upper and lower state houses have any impact on *ex-post* manipulation. One might believe that if a Democrat was running in a state with a Democrat governor and a Democrat upper and lower house that there would be more *ex-post* manipulation regardless of incumbency status. I also find no evidence that in states that are heavily Democratic or Republican there is any systematic bias in the influence the outcomes of close elections vary in any systematic manner. This suggests that incumbency is more important relative to the regional political environment along this dimension. The ability and desire to influence the *ex-post* vote counting process seems to rest in the individual rather than in the regional political machine. Intuitively this is sensible since the marginal returns to the local party of having an individual seat are low relative to the potential costs; the same cannot be said for individual candidates who derive consumption value from being in office.

Conclusions

This paper provides evidence that close elections are systematically biased in favor of incumbents. Using multiple approaches I find an overall pattern of a discontinuity precisely at the majority rules point. This jump is so precise that it is difficult to reconcile this result with the alternative explanation that incumbents have a natural advantage at winning close elections. Given that the process of collecting and tabulating votes is very complex, especially in close elections, it is not surprising that it could be subject to implicit or explicit manipulation. This does not say that the candidate necessarily has a hand in it, but rather that some force is pushing close elections in their favor. Nor is this manipulation necessarily illegal. Rather it merely says that the process that one would like to believe is, if not perfect, at least fair, is biased in favor of incumbents. This evidence suggests that the advantage of incumbency isn't only based on legislative experience and ideological position but also in the ability to have the democratic process biased in their favor. From a policy standpoint this

[9] This data was graciously provided by Rui De Figueiredo. For more information on the construction of this dataset see De Figueiredo (2003)

provides some motivation in favor of increased enforcement of legislation such as the Voting Rights Act of 1965, which was passed to ensure that citizen's votes were properly handled.

Additionally, this paper provides a lower bound on the level of electoral manipulation in Congressional elections. First, I only look at *ex-post* manipulations, rather than the potentially more important *ex-ante* manipulations. I only look at close elections for identification purposes, but that does not rule out that there aren't systematic manipulations in other types of elections. Finally, the observation that 58% of all close elections are won by incumbents is a relative number rather than an absolute one. This number is consistent with the notion that everyone could be cheating, both the challengers and the incumbents, but that incumbents are slightly better at it.

Finally, this paper is a contribution to the growing empirical literature on corruption by showing how these techniques can be fruitfully used to examine issues of voting manipulation. Rather than relying on parametric assumptions this paper uses institutional details to non-parametrically identify the existence of *ex-post* manipulations. Further empirical research on corruption in politics would without a doubt provide many interesting findings that would be of significant relevance. Understanding what determinants other than incumbency impact the ability of politicians to exercise political influence would be a logical next step, as would examining whether this holds in contexts other than House elections. Additionally, looking at other settings would be useful, although the experience from this paper would caution future researchers to restrict their search to datasets with a large number of elections. Given that at times decisions hinge on the choices of one or two elected representatives, the real costs of the types of manipulations that are described in this paper are potentially quite important.

Appendix: The Specification of the Non-Parametric Model

In the non-parametric density the first important step is to choose a binsize to represent the unsmoothed histogram. One *ad hoc* rule of thumb suggestion is that the binsizes should equal to $b = \frac{2\hat{\sigma}}{\sqrt{n}}$ where $\hat{\sigma}$ is the sample standard deviation of incumbent vote share and n is the number of observations. However, one problem with using this rule is that it makes the bins very large and obscures much of the detail around the point of the discontinuity. The difference between figures II.A and II.B illustrates this point nicely. Using the *ad hoc* rule of thumb derived from data regarding the entire distribution leads to a binsize of .0018. This leads to fairly large bins close to the discontinuity approximately ranging in size from 30 to 80 elections. Using a finer binsize of .0005 in figure IIb provides a more detailed picture of the data. I find that the results are fairly robust to binsize selection, so for expositional purposes I use the smaller binsize of .0005 for the graphical presentation presentation of the data.

The next tuning parameter that must be selected is the bandwidth of the kernel density estimator. Again there are no underlying rationals for selecting the optimal bandwidth although the normal reference rule (another rule of thumb) gives an automatic procedure for selecting the bandwidth with $h = 2.576\frac{\hat{\sigma}}{n^{\frac{1}{5}}}$. I find that the results are generally consistent with presence of a discontinuity; however the specification of the bandwidth does influence magnitude and statistical significance in important ways. The easiest way to illustrate this trade-off is to look at figure III.A-III.F. In these results as the specified bandwidth becomes smaller the estimates of the discontinuity becomes relatively more sensitive to the data near the discontinuity. Comparing III.A and III.F illustrates this point clearly. In III.A where the bandwidth is set to .08, the estimated curves don't seem to vary significantly relative to the data close to the discontinuity. On the other hand with a bandwidth size of .0025 in III.F the data local to the discontinuity seems to impact the shape of the curve dramatically. Not knowing the correct specification, I have chosen to present the results from a wide variety of specifications.

Table IV presents evidence for a discontinuity for 121 different specifications which is meant to be a representative sub-sample of the number of different specifications I have used. I started the construction of this table by computing combinations of binsizes and bandwidths.[10] The columns represent the specified bandwidths and the rows represent the specified binsizes. The variables presented are the size of the discontinuity and the associated t-statistics are in parentheses. The variation in the t-statistics in this table is substantial, with the lowest t-statistic equal to 1.530 for the $(.000175, .005)$ binsize bandwidth pair to a t-statistic of 16.540 for the $(.0035, .005)$ pair.[11] Two striking observations are apparent. First, discretion in the binsize and bandwidth choices can lead to substantial differences in the statistical significance of the discontinuity. However, I still observe that the lowest t-statistic still is significant at the 13% confidence level.[12] On average for the entire table the mean t-statistic equals 3.42. This result is driven by outliers that are most likely a misspecification. For instance, the abnormally large results for the $(.0035, .005)$ would seem to be a violation of the asymptotic condition that $\frac{binsize}{bandwidth} \to 0$.

Using the bandwidth and binsize reference rules allows for the construction of a more plausible range of binsizes and bandwidths. The outlined box inside table IV gives the reader an idea of what combinations these rules will result in when computing them across various ranges of the data. For instance, when the entire range of vote shares is used the rules result in a binsize bandwidth combination approximately equal to $(.00175, .004)$. When restricting this computation of the reference rules to the 25% to 75% range, this results in a combination approximately equal to $(.0015, .003)$. In general, when using these rules as the range becomes tighter around the majority rules point, the binsizes and bandwidths become

[10] For bandwidths I used the range $[.005, .001, ..., .1]$ and for binsizes I used the range $[.0005, .00075, ..., .004]$. This is logical since finer bandwidths seemed completely chaotic (see figure IIIf) and finer binsizes tended to have zero values local to the discontinuity. I omitted many of these specifications from table IV for space considerations.

[11] All subsequent calculations are based on all of the combinations that I computed, and not just those shown in the table (see previous note)

[12] Additionally, this observation comes from the choice of a large bin and a very small bandwidth. The density estimate is extremely jagged and probably violates the principle that the binsize should be markedly smaller than the bandwidth.

smaller since they are no longer being enlarged to accommodate sparser observations at the tails. The construction of this box is *ad hoc* and subject to discretion on the part of the researcher. The mean of all of the t-statistics in this box is 2.44 with a standard deviation of .41. The lowest t-statistic in this box is 1.798 which is still significant at approximately the 7% confidence level. On average for all of these specifications the confidence level is markedly lower at approximately the 1.5% confidence level.

Additionally in unreported results I have run the non-parametric discontinuous density estimator on other parts of the range of vote shares. These results suggest that the estimator performsas expected. For reasonable specifications of the binsizes and bandwidths I find that the estimator produces results significant at the 5% confidence level approximately 5% of the time and results significant at the 1% confidence level $1\% - 2\%$ of the time. It does not appear that this technique is inherently biased towards significant results.

References

[1] Ansolabehere, Stephen, James M. Snyder Jr. and Charles I. Stewart III "Old Voters, New Voters, and Personal Vote: Using Redistricting to Measure the Incumbency Advantage" *American Journal of Political Science* January (2000): 17-34

[2] Brady, Henry "What Happened in Palm Beach County?" unpublished letter 2000

[3] Caro, Robert *The Years of Lyndon Johnson: Means of Ascent* (1990) Random House, New York

[4] De Figueiredo, Rui J. P. Jr "Budget Institutions and Political Insulation: Why States Adopt the Item Veto" Forthcoming *Journal of Public Economics*

[5] Degeorge, Francois, Jayendu Patel, and Richard Zeckhauser "Earnings Management to Exceed Thresholds" *Journal of Business* Issue #1 (1999): 1-33

[6] DiNardo, John and Justin Tobias "Non-Parametric Density and Regression Estimation" *Journal of Economic Perspectives* Fall (2001): 11-28

[7] Duggan, Mark and Steven D. Levitt "Winning Isn't Everything: Corruption in Sumo Wrestling" *American Economic Review* December (2002): 1594-1605

[8] Fan, Jianqing and Irene Gijbels *Local Polynomial Modelling and Its Applications* (1996) Chapman and Hall, New York

[9] Gelman, Andrew, and Gary King "Estimating Incumbency Advantage without Bias" *American Journal of Political Science* November (1990): 1142-64

[10] Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design." *Econometrica* January (2001): 201-209

[11] King, Gary "Elections to the United States House of Representatives: 1898-1992" ICPSR dataset 6311

[12] Lee, David "The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the U.S. House" NBER Working Paper 8441 (2003)

[13] Levitt, Steven D. and Catherine D. Wolfram "Decomposing the Sources of Incumbency Advantage in the U.S. House." *Legislative Studies Quarterly* February (1997): 45-60

[14] McCrary, Justin "Estimating a Discontinuous Density Function" working paper (2004)

[15] New York Times "Special Section on the 2000 Election" November 12th, 2001

[16] Porter, Jack "Estimation in the Regression Discontinuity Model" working paper (2003)

[17] Shleifer, Andrei and Robert W. Vishny "Corruption" *The Quarterly Journal of Economics*, Vol. 108, No. 3. (Aug., 1993): pp. 599-617.

[18] Wolfers, Justin "Point Shaving: Corruption in NCAA Basketball" working paper (2003)

Table I: Summary Statistics

Election Characteristics	Observations	Mean Total Votes	SD Total Votes	Difference Mean	Difference SD
Overall for all Elections with Incumbents	13948	114596.1	101668.0	27196.5	38317.7
Centered at 50% With Band Width = .06	178	108510.5	82568.1	88.0	529.6
Centered at 50% with Band Width = .01	315	108025.1	86584.1	155.2	867.6

Table II: Results with Bandwidth = .01

	N	Mean	Prob [Mean > .5]
Main Results	315	0.581	0.002
1898-1925	110	0.473	0.748
1925-1951	106	0.613	0.013
1951-1975	60	0.650	0.014
1975-1992	39	0.692	0.012
Experience > 2	139	0.612	0.005
Experience <= 2	176	0.557	0.076

Table III: Results with Bandwidth = .06

	N	Mean	Prob [Mean > .5]
Main Results	178	0.573	0.030
1898-1925	57	0.439	0.855
1925-1951	61	0.607	0.062
1951-1975	38	0.658	0.036
1975-1992	22	0.682	0.067
Experience = 2	70	0.643	0.011
Experience <= 2	108	0.528	0.315

Table IV: Discontinuity Specifications

Binsize	Bandwidth 0.005	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.050	0.060	0.080
0.00050	**0.506**	**0.725**	**0.761**	**0.708**	**0.596**	**0.481**	**0.411**	**0.374**	**0.345**	**0.305**	**0.329**
	(2.391)	(2.995)	(3.150)	(3.206)	(2.950)	(2.503)	(2.220)	(2.098)	(2.089)	(1.967)	(2.405)
0.00075	**0.294**	**0.616**	**0.714**	**0.675**	**0.580**	**0.473**	**0.401**	**0.366**	**0.341**	**0.301**	**0.327**
	(1.947)	(2.468)	(3.003)	(3.291)	(3.165)	(2.711)	(2.370)	(2.234)	(2.202)	(2.042)	(2.443)
0.00100	**0.447**	**0.672**	**0.743**	**0.689**	**0.586**	**0.473**	**0.403**	**0.368**	**0.343**	**0.302**	**0.328**
	(2.291)	(2.549)	(2.867)	(3.101)	(2.988)	(2.521)	(2.205)	(2.061)	(2.042)	(1.907)	(2.329)
0.00125	**0.492**	**0.724**	**0.764**	**0.691**	**0.591**	**0.477**	**0.406**	**0.372**	**0.345**	**0.303**	**0.329**
	(4.083)	(2.395)	(2.588)	(2.687)	(2.586)	(2.183)	(1.910)	(1.798)	(1.773)	(1.652)	(2.053)
0.00150	**0.357**	**0.656**	**0.746**	**0.681**	**0.578**	**0.474**	**0.402**	**0.366**	**0.342**	**0.303**	**0.329**
	(1.940)	(2.139)	(2.447)	(2.692)	(2.691)	(2.404)	(2.140)	(2.009)	(2.016)	(1.904)	(2.375)
0.00175	**0.345**	**0.631**	**0.704**	**0.651**	**0.571**	**0.473**	**0.411**	**0.379**	**0.352**	**0.310**	**0.334**
	(1.530)	(1.643)	(1.935)	(2.187)	(2.253)	(2.037)	(1.876)	(1.821)	(1.885)	(1.810)	(2.283)
0.00200	**0.339**	**0.649**	**0.698**	**0.645**	**0.552**	**0.454**	**0.393**	**0.361**	**0.343**	**0.301**	**0.329**
	(3.876)	(2.383)	(2.624)	(2.830)	(2.751)	(2.313)	(2.049)	(1.926)	(1.945)	(1.802)	(2.261)
0.00250	**0.670**	**0.765**	**0.782**	**0.675**	**0.580**	**0.467**	**0.399**	**0.368**	**0.346**	**0.303**	**0.331**
	(5.370)	(2.535)	(2.420)	(2.460)	(2.489)	(2.117)	(1.858)	(1.751)	(1.725)	(1.590)	(1.976)
0.00300	**0.540**	**0.668**	**0.807**	**0.689**	**0.580**	**0.476**	**0.419**	**0.377**	**0.348**	**0.309**	**0.331**
	(8.634)	(7.742)	(4.370)	(4.615)	(4.218)	(3.096)	(2.523)	(2.168)	(1.960)	(1.777)	(2.152)
0.00350	**0.853**	**0.893**	**0.826**	**0.714**	**0.602**	**0.485**	**0.415**	**0.388**	**0.362**	**0.322**	**0.343**
	(16.540)	(13.443)	(7.115)	(5.736)	(4.066)	(2.803)	(2.296)	(2.193)	(2.176)	(2.031)	(2.601)
0.00400	**0.728**	**0.704**	**0.765**	**0.693**	**0.577**	**0.473**	**0.412**	**0.374**	**0.352**	**0.315**	**0.341**
	(8.967)	(7.331)	(4.804)	(4.319)	(3.983)	(2.934)	(2.395)	(2.171)	(2.060)	(1.901)	(2.396)

Estimated size of the discontinuity and associated t-statistic are in parenthesis.
The bandwidth and binsize parameters are used to calibrate the non-parametric estimation procedure outlined in the text.
The area inside the shaded box roughly indicates the areas where the reference rules for binsize and bandwidth estimates

Figure I.A

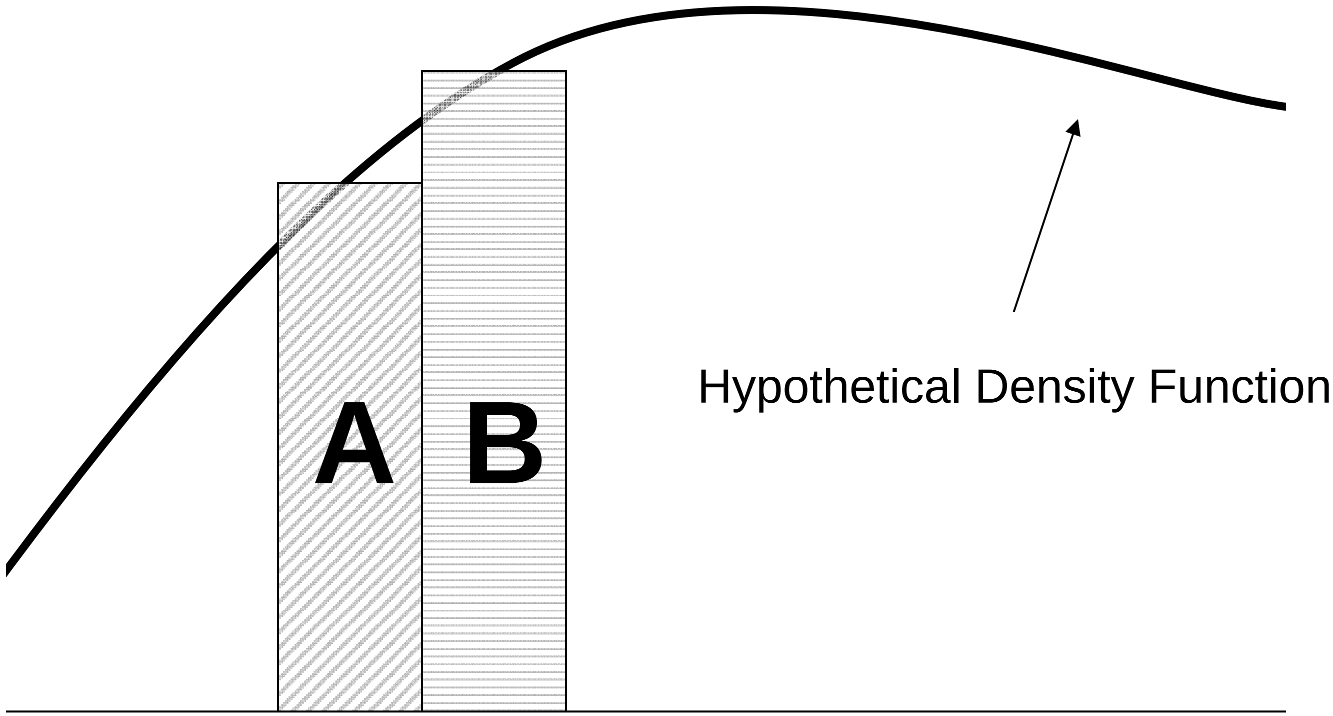


Figure I.B

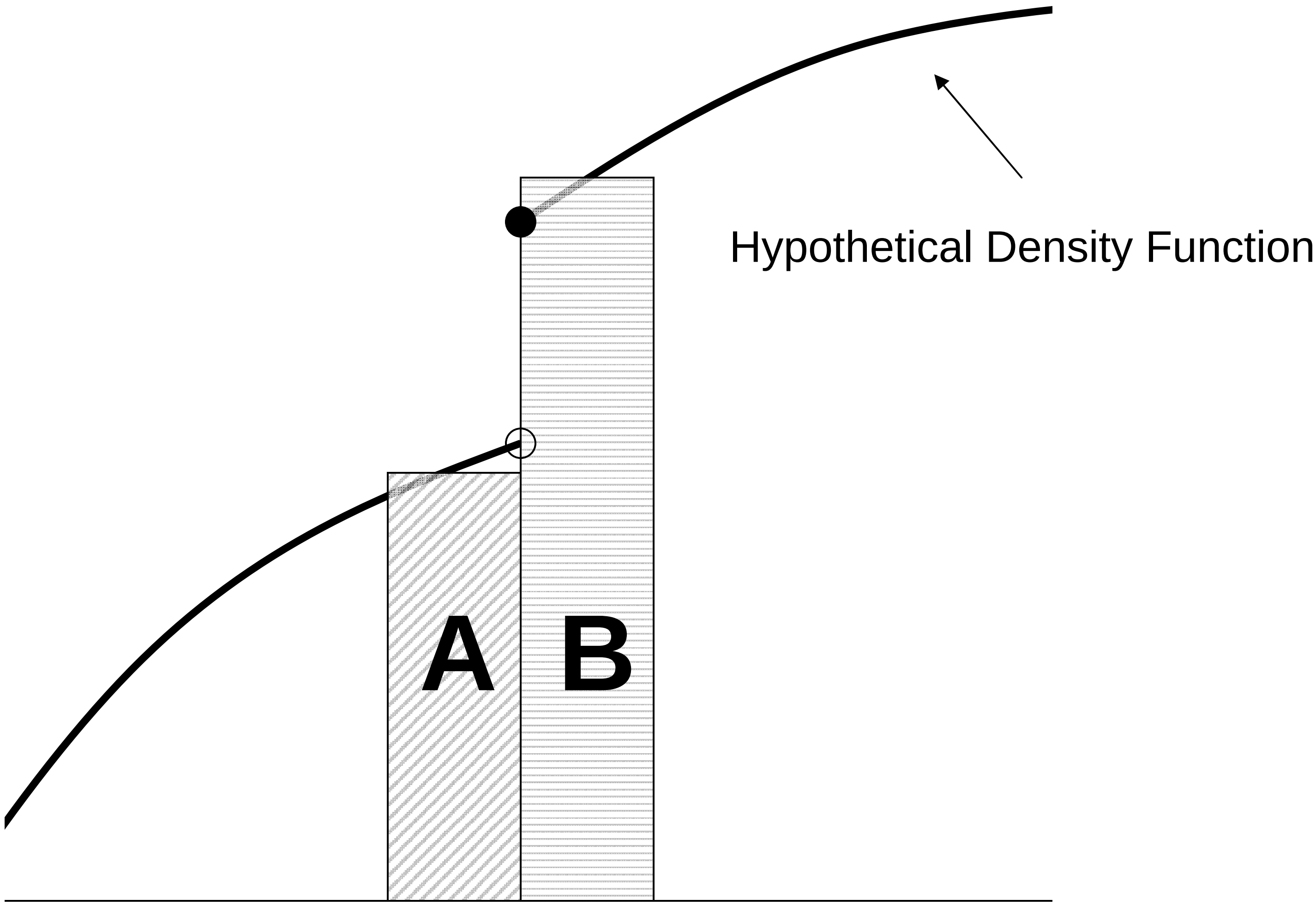


Figure I.C

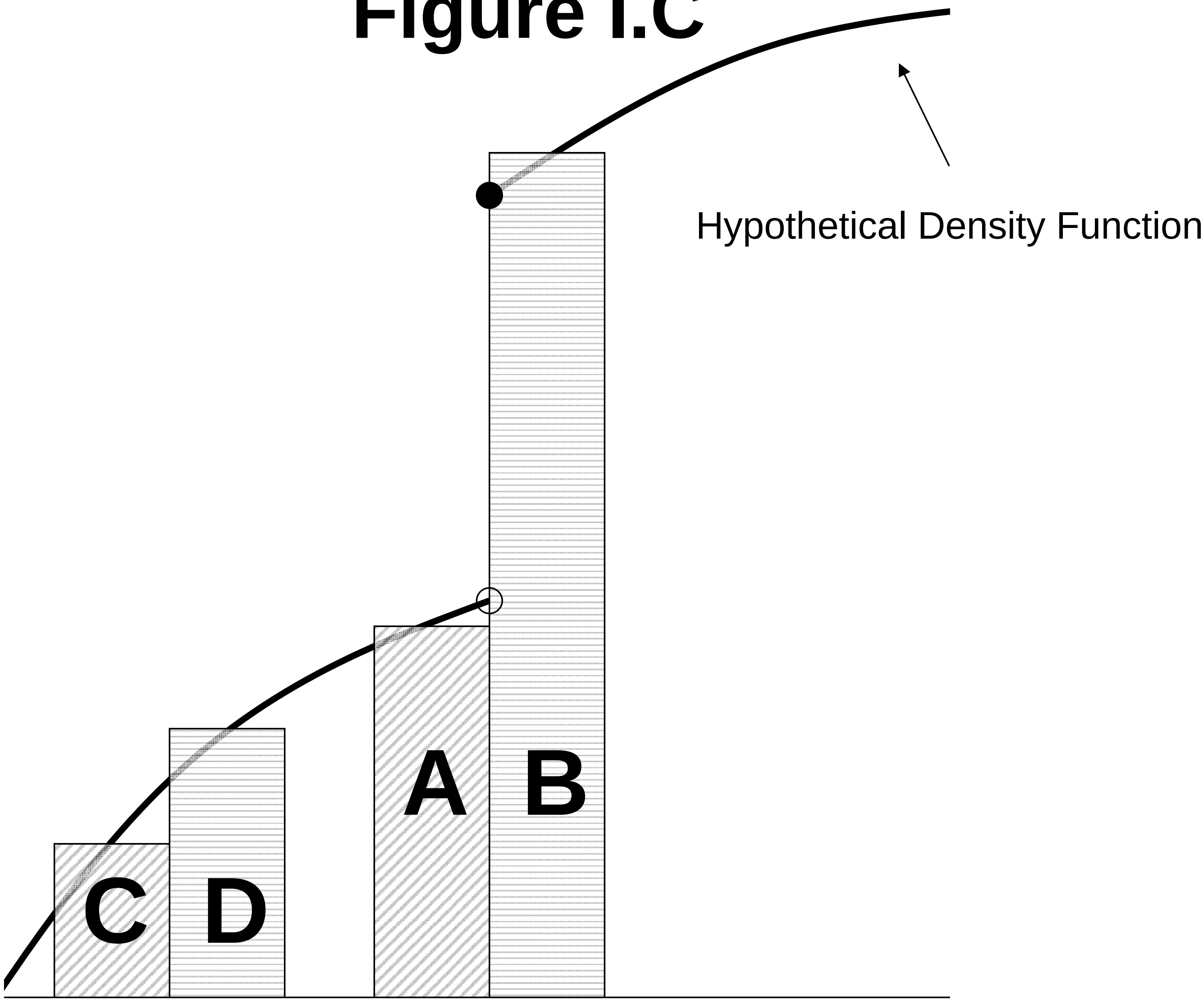
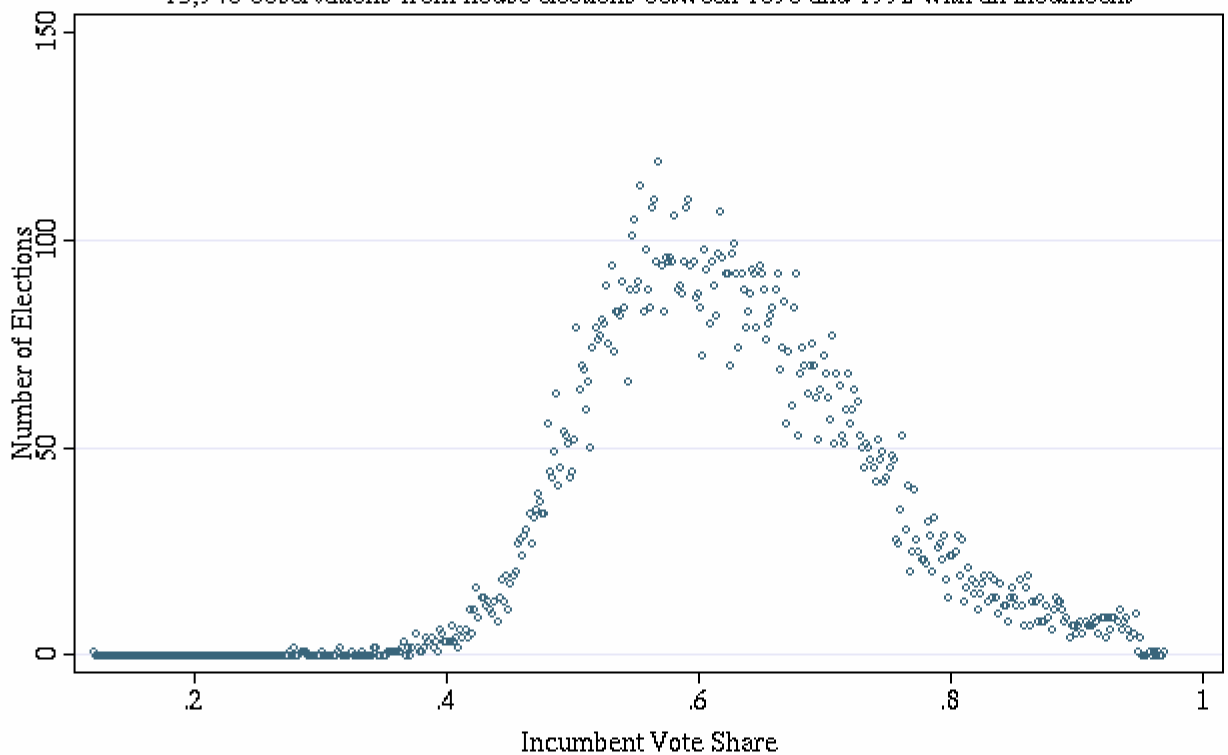


Figure II.A: Histogram of the Incumbent Vote Share Density Function

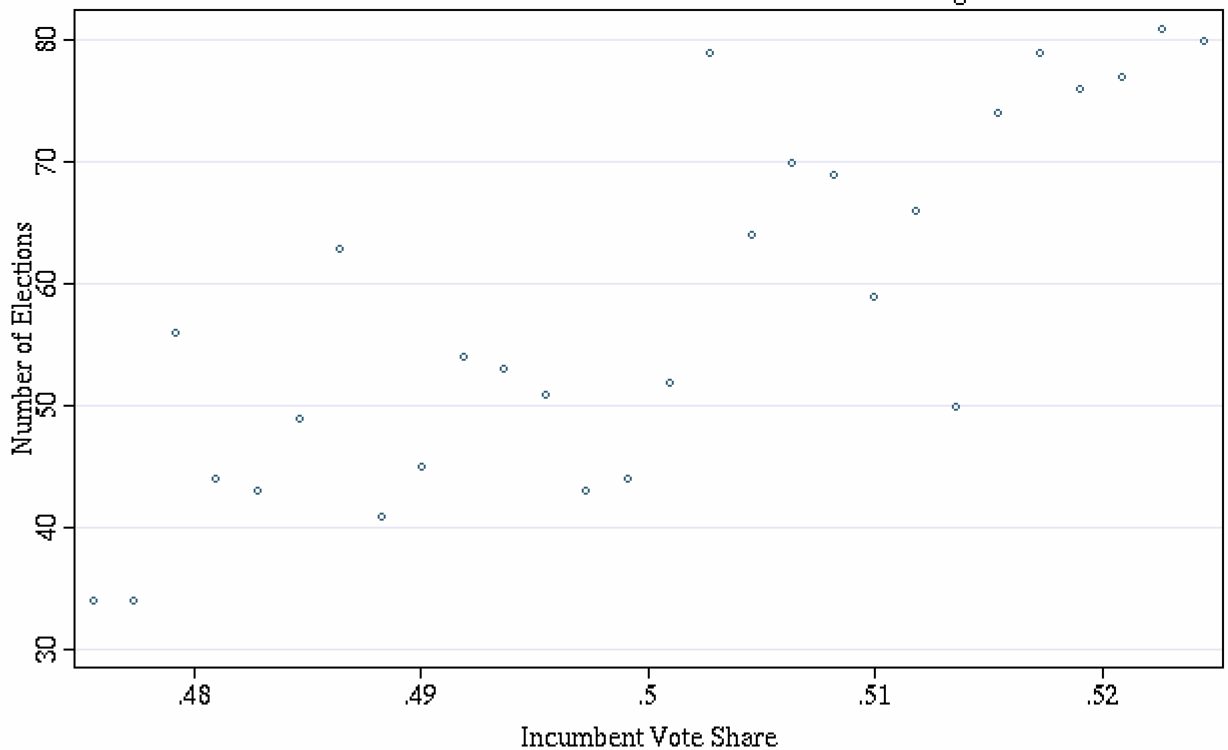
13,948 observations from house elections between 1898 and 1992 with an incumbent



Note: Bin Size = .0018

Figure II.B: Histogram of the Incumbent Vote Share Density Function

A Closer View Over the 47.5% to 52.5% vote share range



Note: Bin Size = .0018

Figure II.C: Histogram of the Incumbent Vote Share Density Function

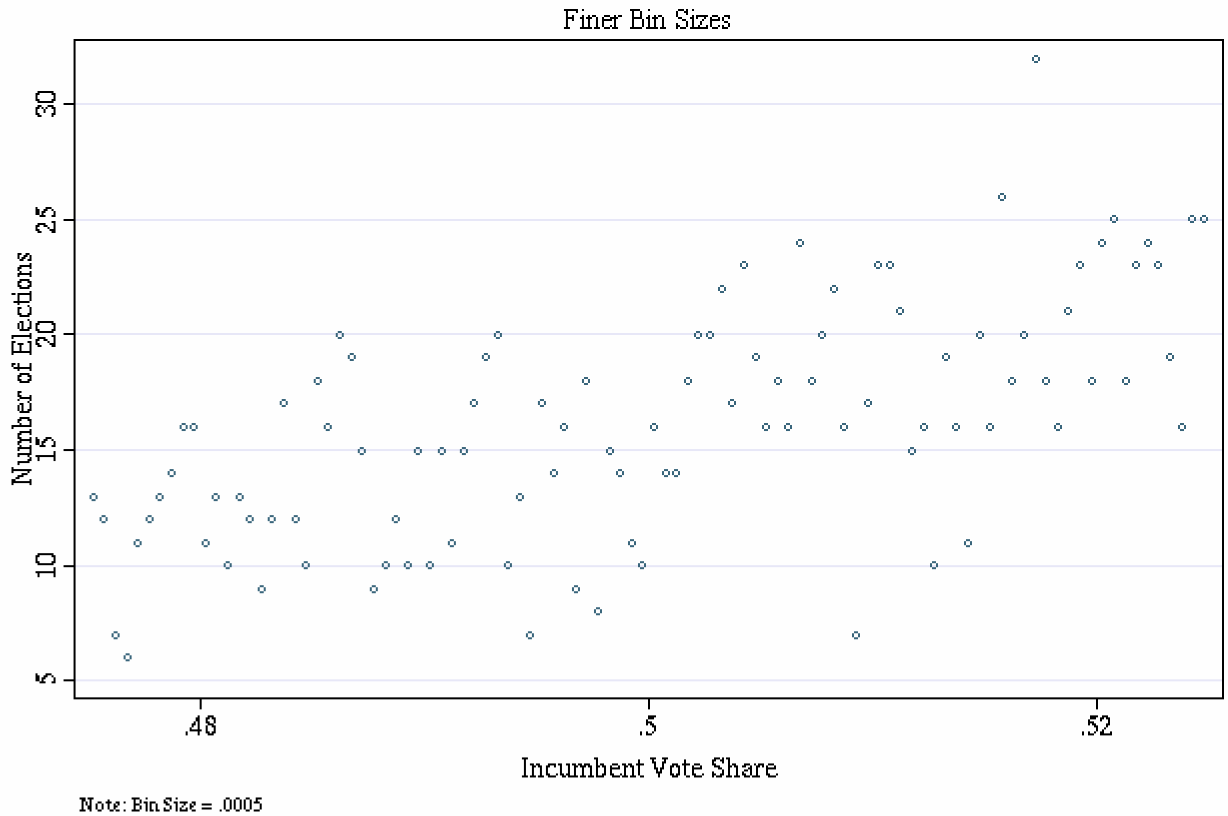
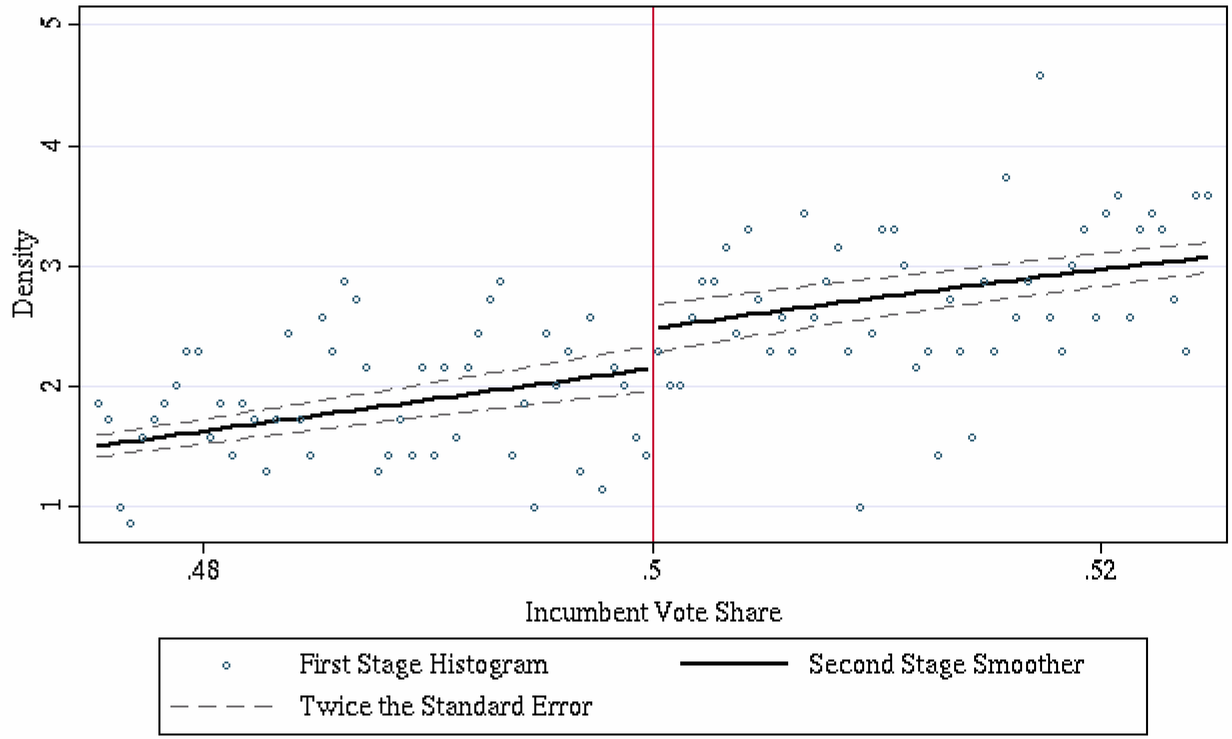


Figure III.A: Histogram of the Incumbent Vote Share Density Function

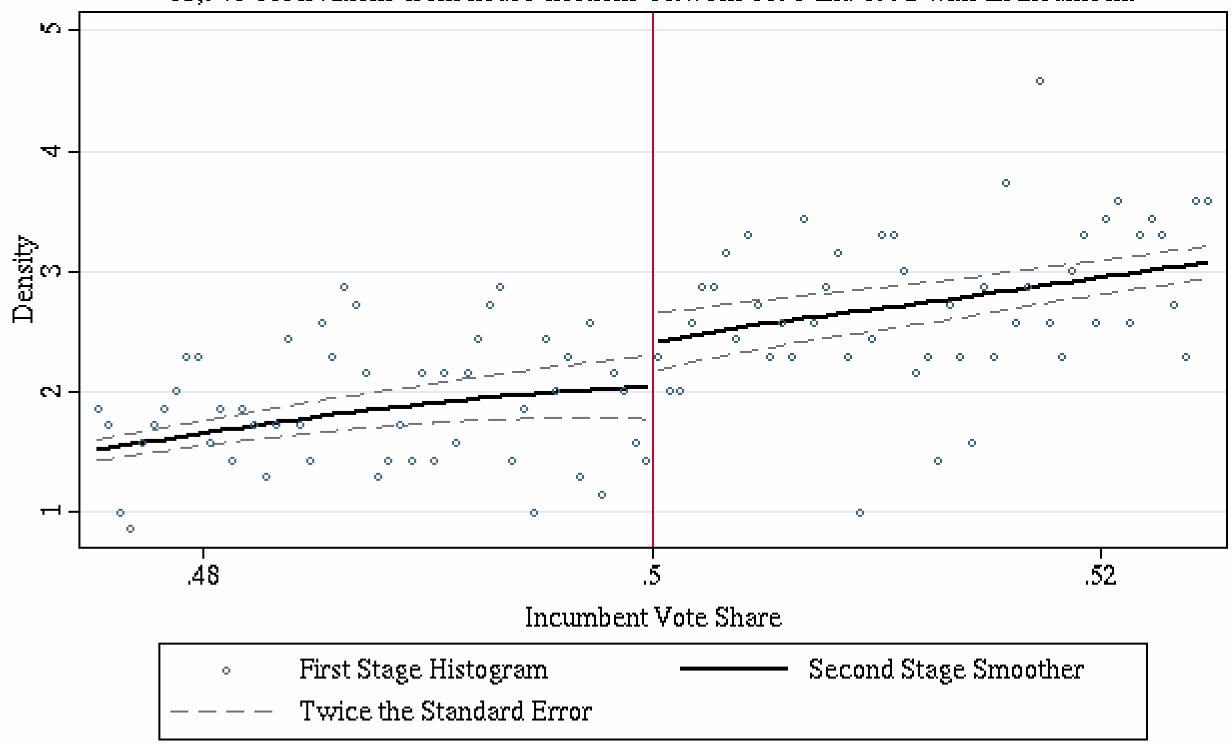
13,941 observations from house elections between 1898 and 1992 with an incumbent



Bin Size = .0005 Bandwidth = .08 T−Stat = 2.404

Figure III.B: Histogram of the Incumbent Vote Share Density Function

13,941 observations from house elections between 1898 and 1992 with an incumbent



Bin Size = .0005   Bandwidth = .04   T−Stat = 2.098

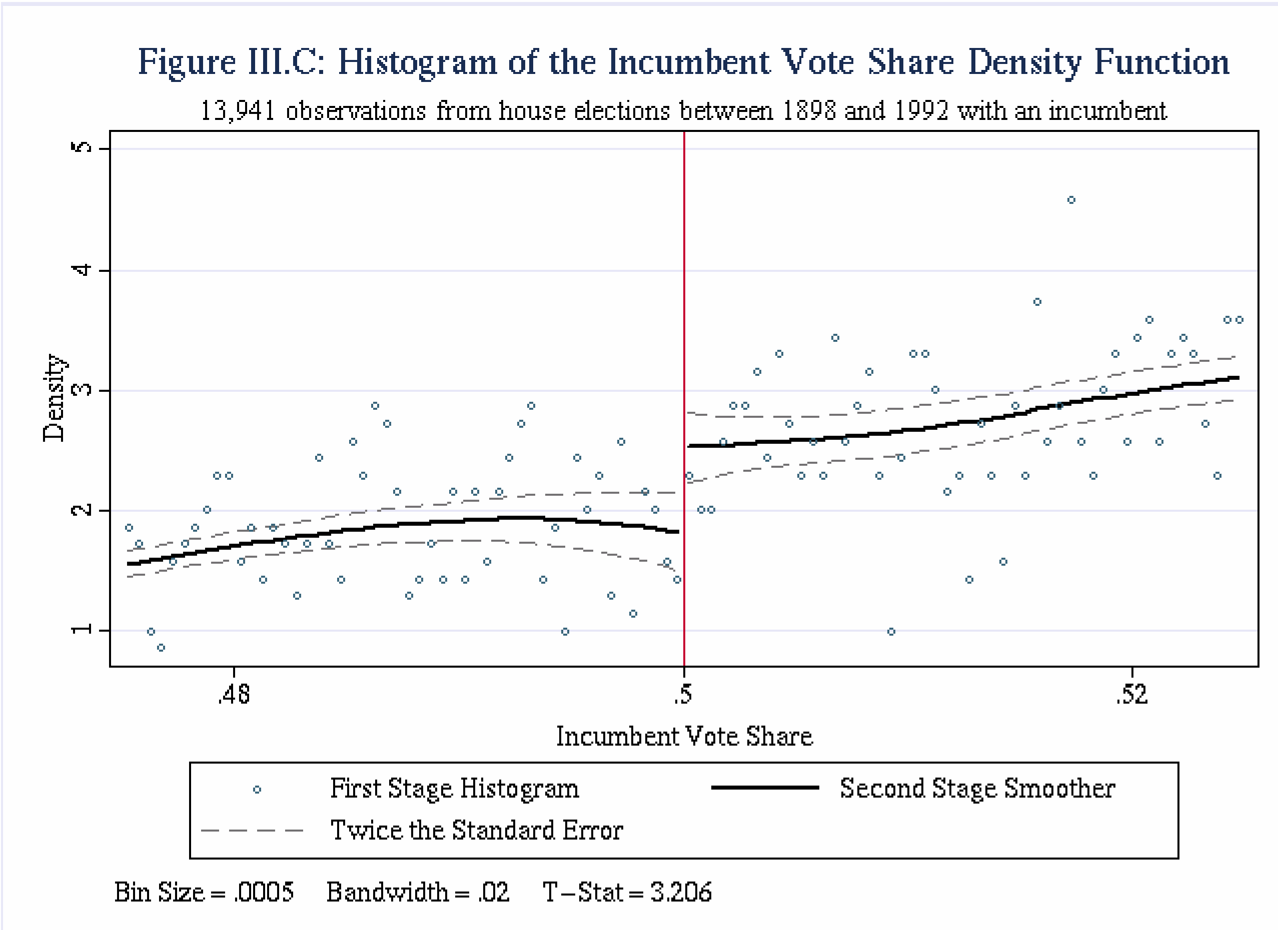


Figure III.C: Histogram of the Incumbent Vote Share Density Function
13,941 observations from house elections between 1898 and 1992 with an incumbent
Density
5
4
3
2
1
.48
.5
.52
Incumbent Vote Share
First Stage Histogram
Second Stage Smoother
Twice the Standard Error
Bin Size = .0005 Bandwidth = .02 T−Stat = 3.206

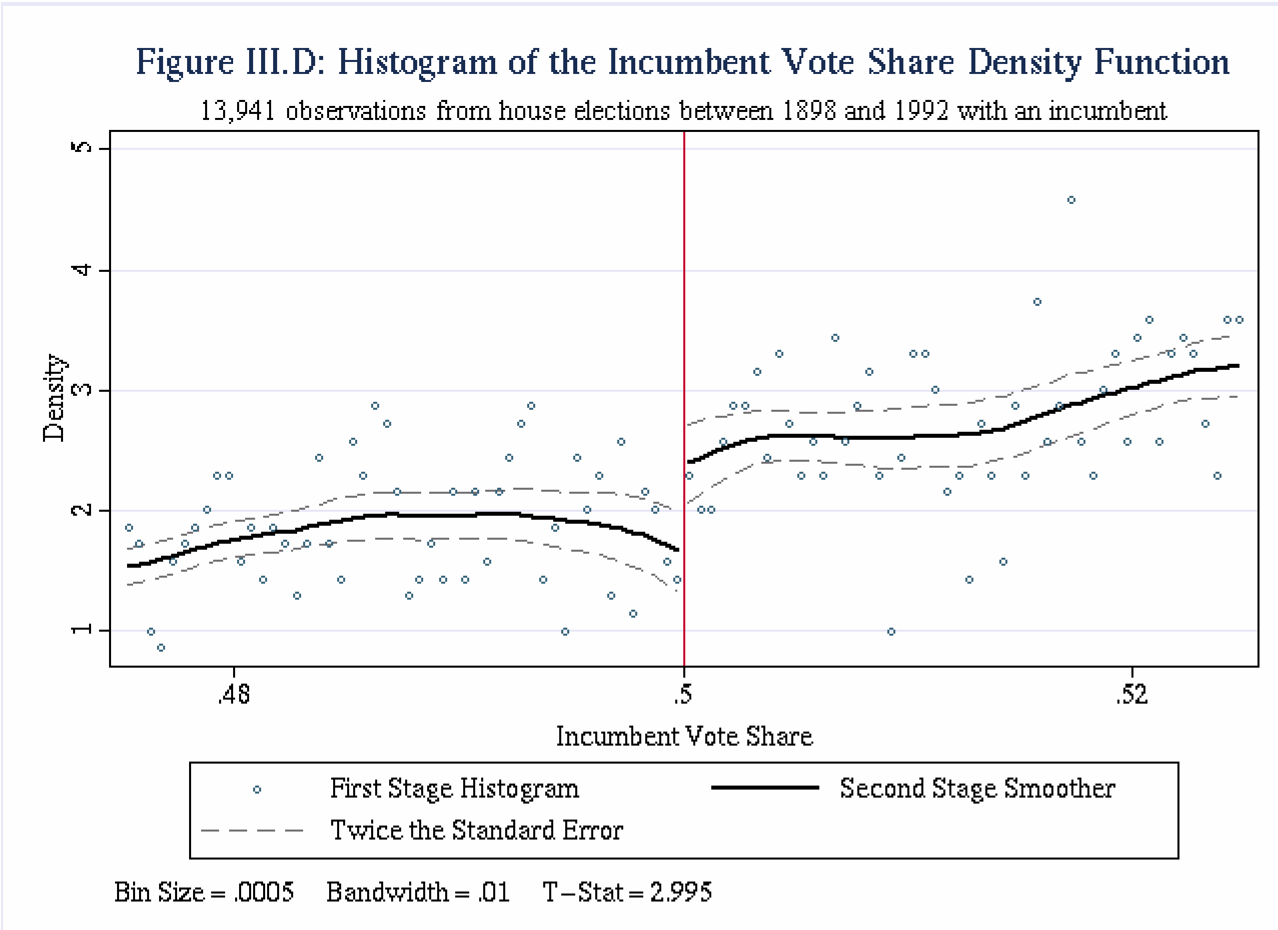


Figure III.D: Histogram of the Incumbent Vote Share Density Function
13,941 observations from house elections between 1898 and 1992 with an incumbent
Density
1
2
3
4
5
.48
.5
.52
Incumbent Vote Share
First Stage Histogram
Second Stage Smoother
Twice the Standard Error
Bin Size = .0005 Bandwidth = .01 T−Stat = 2.995

Figure III.E: Histogram of the Incumbent Vote Share Density Function

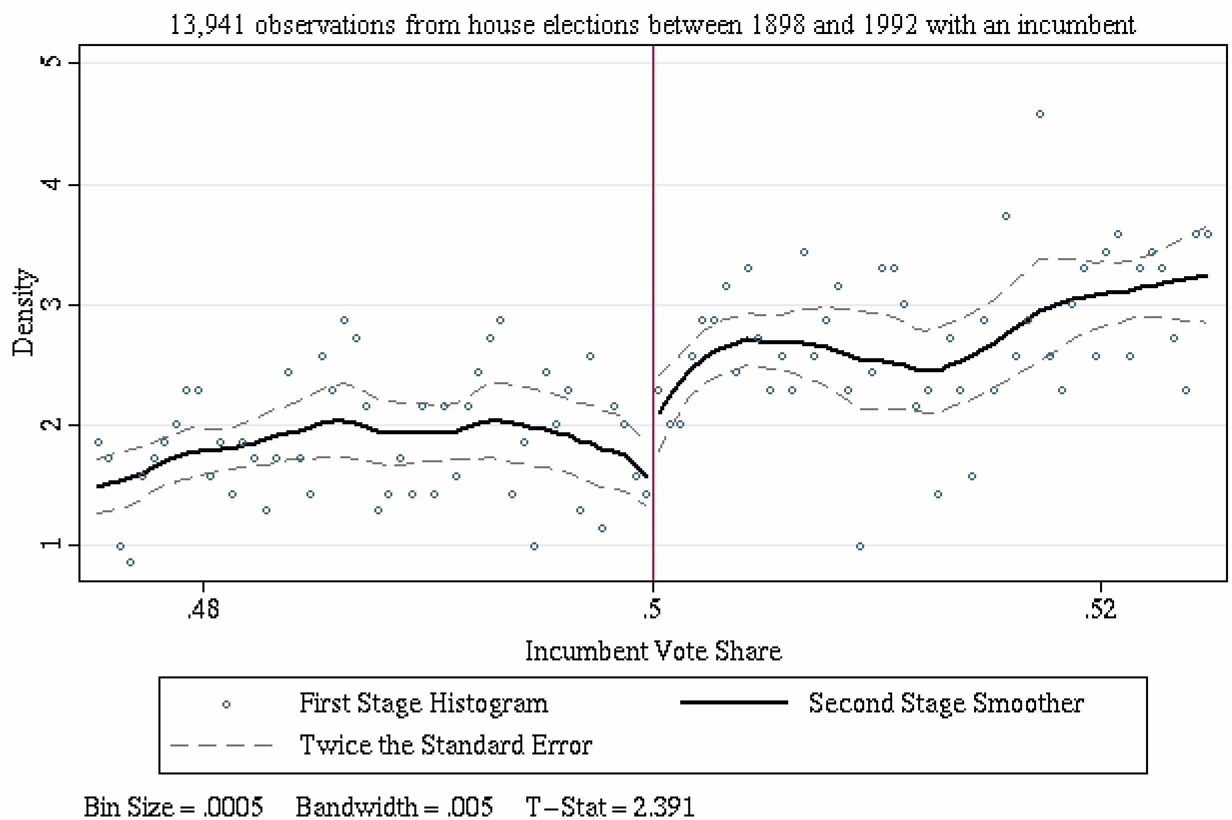
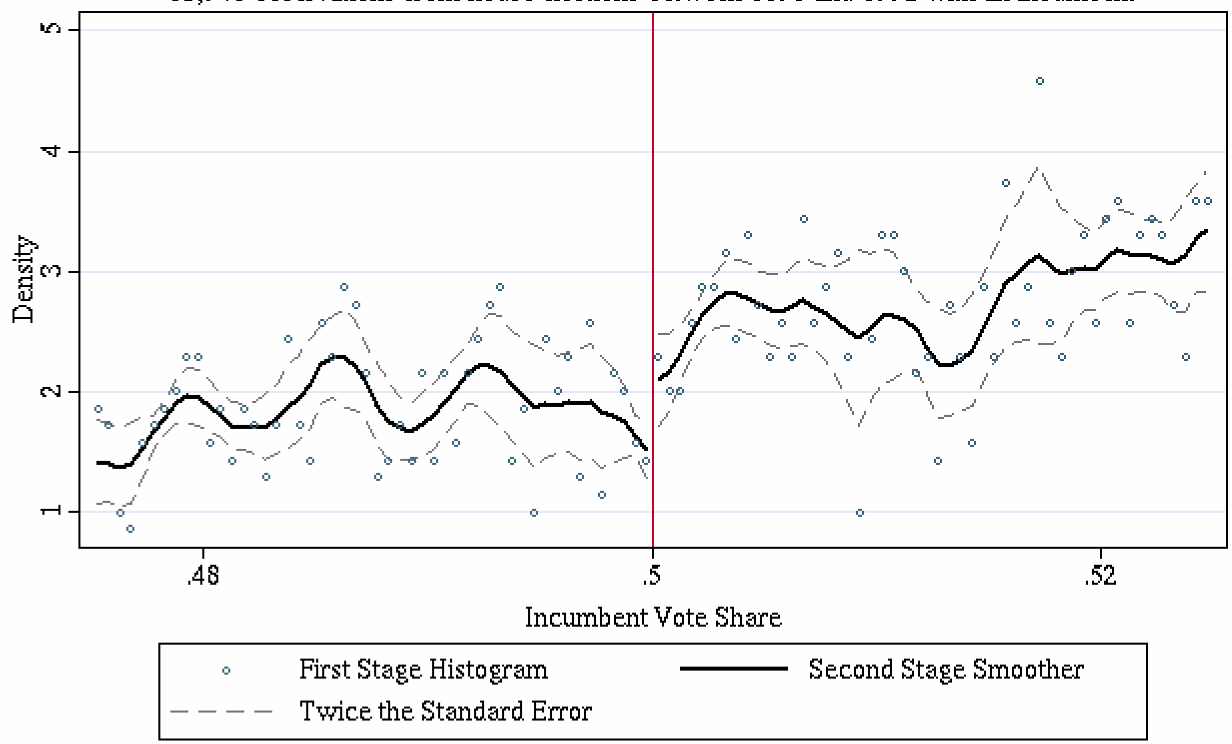


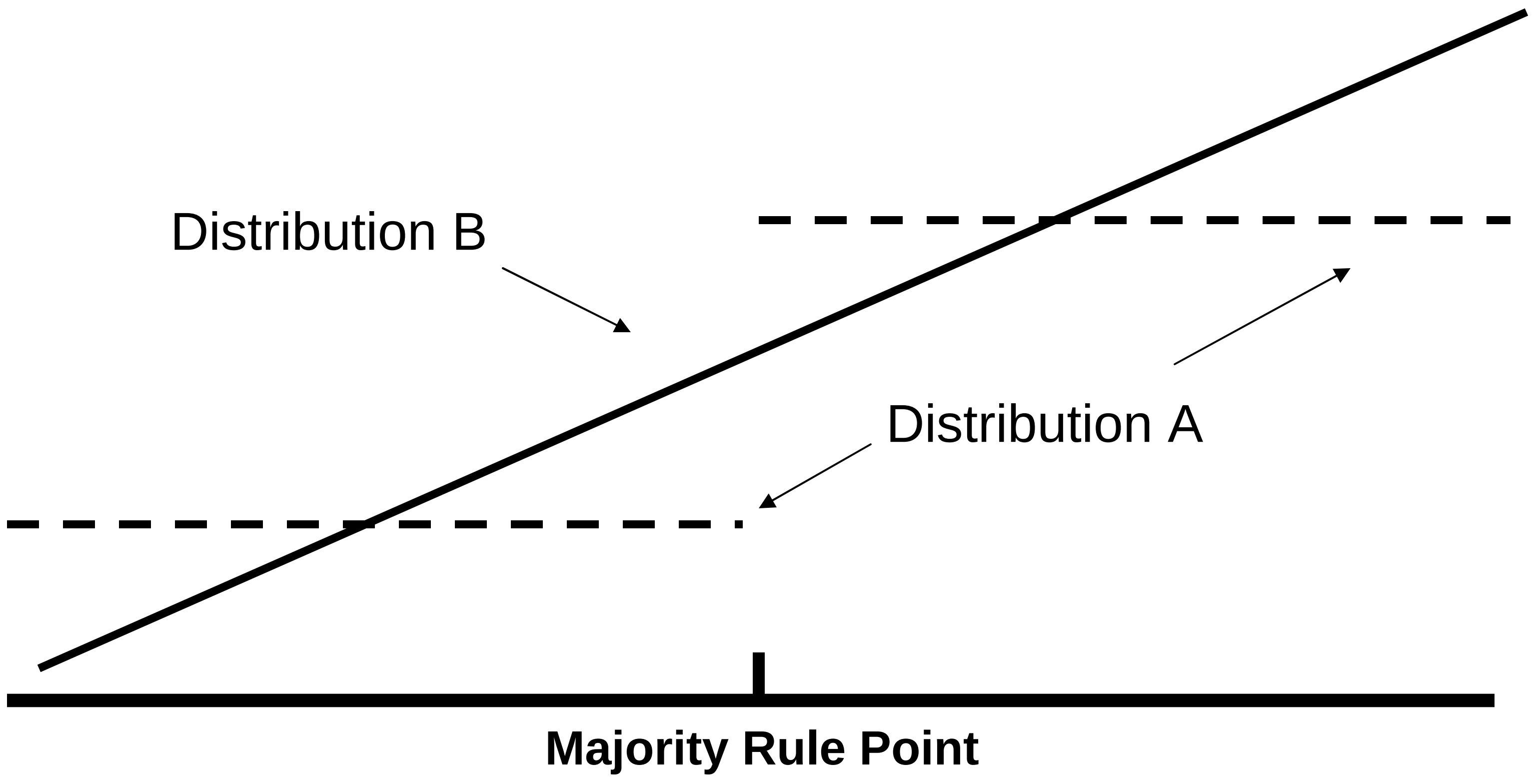
Figure III.F: Histogram of the Incumbent Vote Share Density Function

13,941 observations from house elections between 1898 and 1992 with an incumbent



Bin Size = .0005 Bandwidth = .0025 T−Stat = 2.479

Figure IV



Incumbent Vote Shares

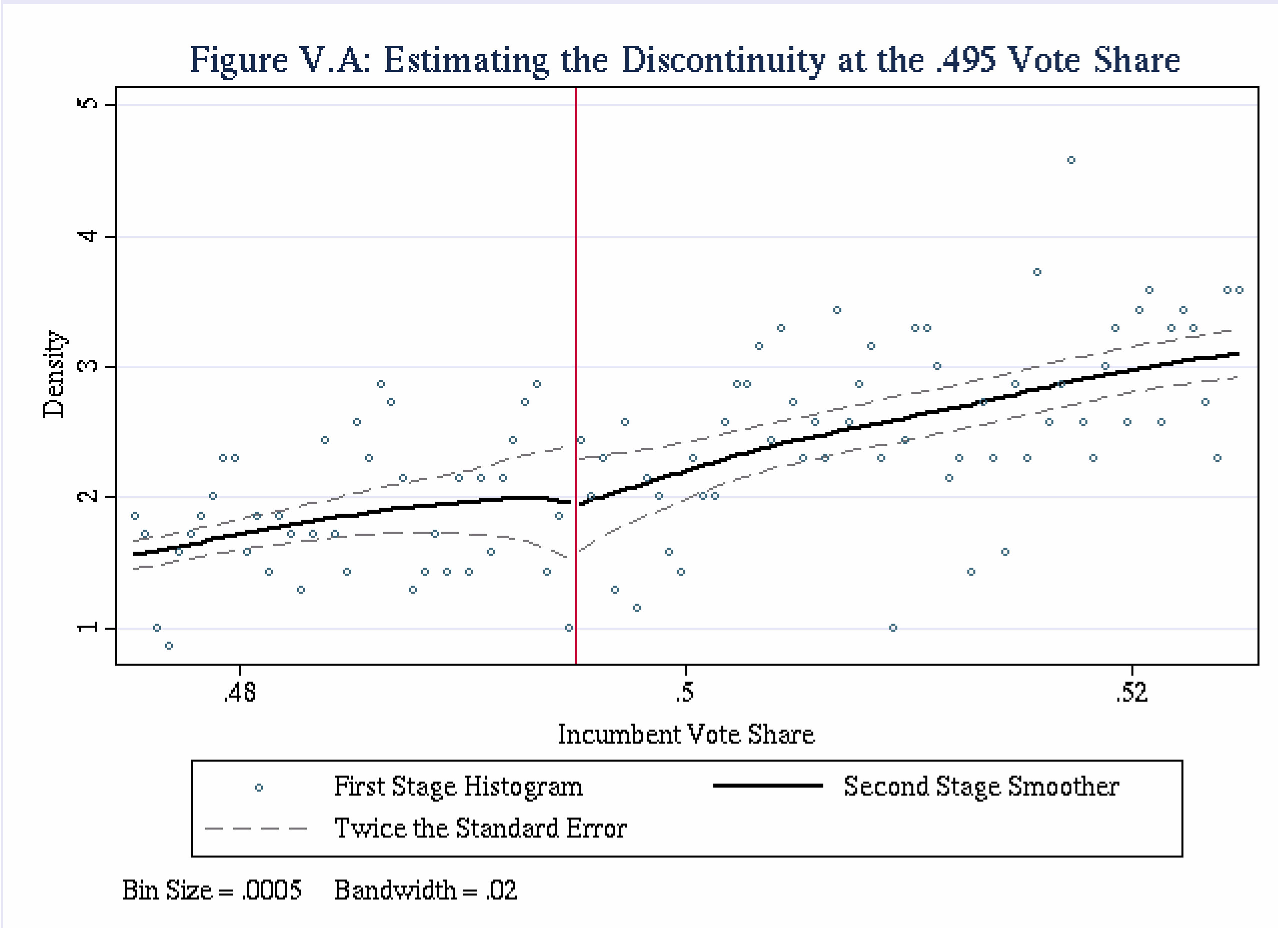
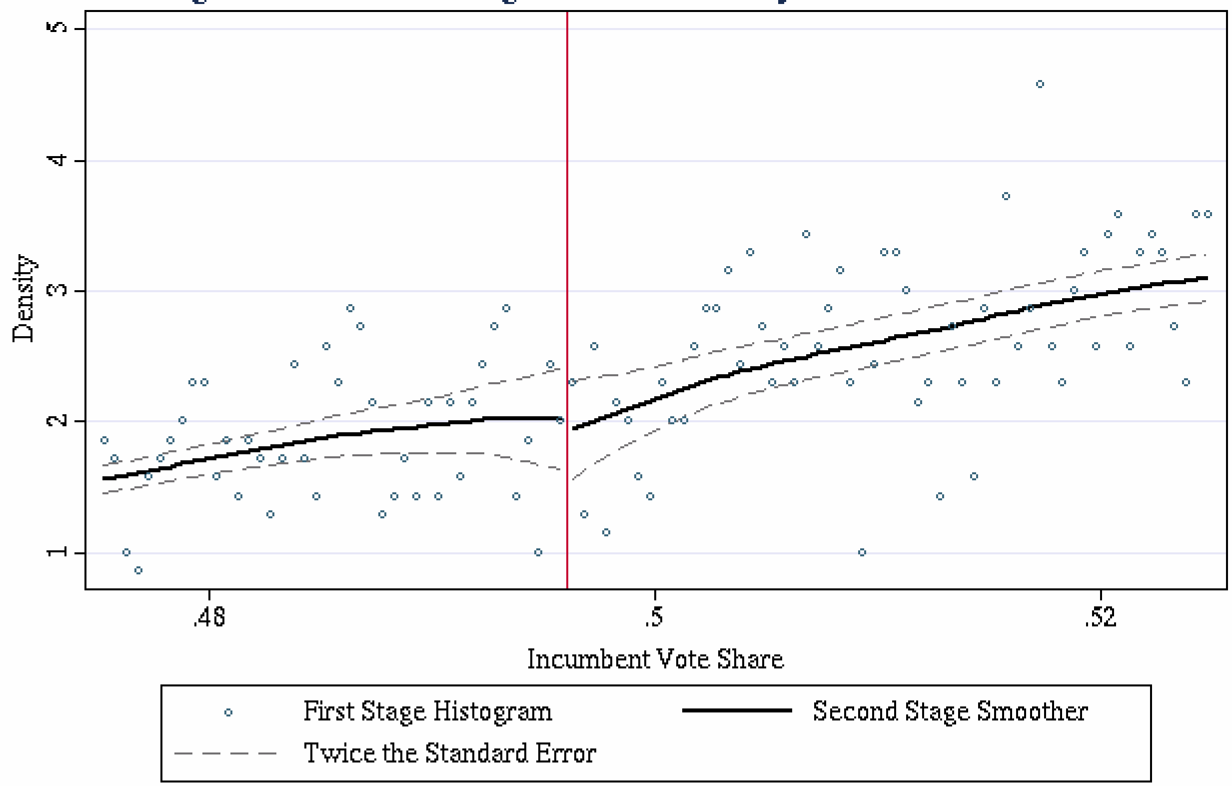


Figure V.A: Estimating the Discontinuity at the .495 Vote Share
Density
1
2
3
4
5
.48
.5
.52
Incumbent Vote Share
First Stage Histogram
Second Stage Smoother
Twice the Standard Error
Bin Size = .0005 Bandwidth = .02

Figure V.B: Estimating the Discontinuity at the .496 Vote Share



Bin Size = .0005 Bandwidth = .02

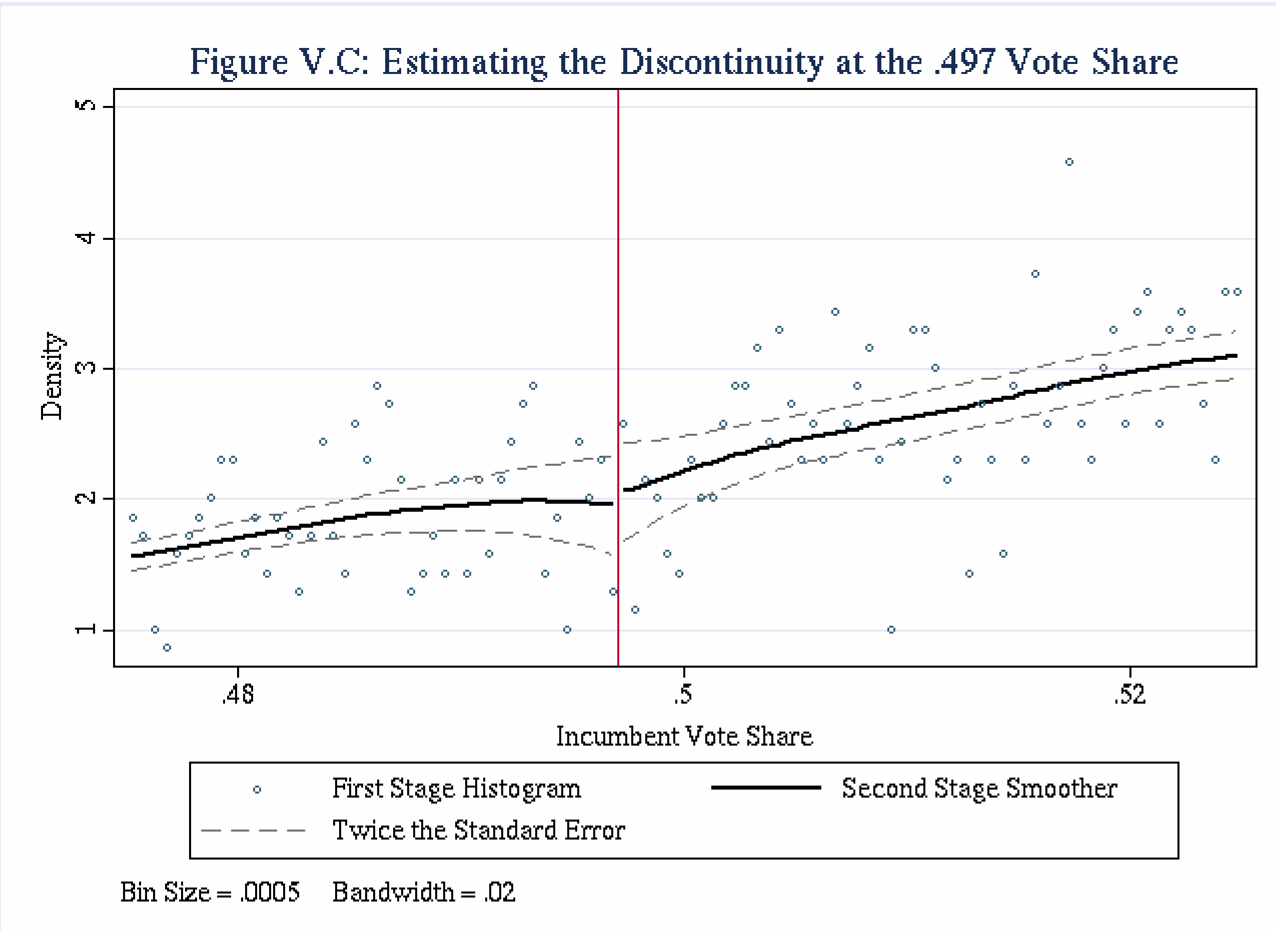


Figure V.C: Estimating the Discontinuity at the .497 Vote Share
Density
1
2
3
4
5
.48
.5
.52
Incumbent Vote Share
First Stage Histogram
Second Stage Smoother
Twice the Standard Error
Bin Size = .0005 Bandwidth = .02

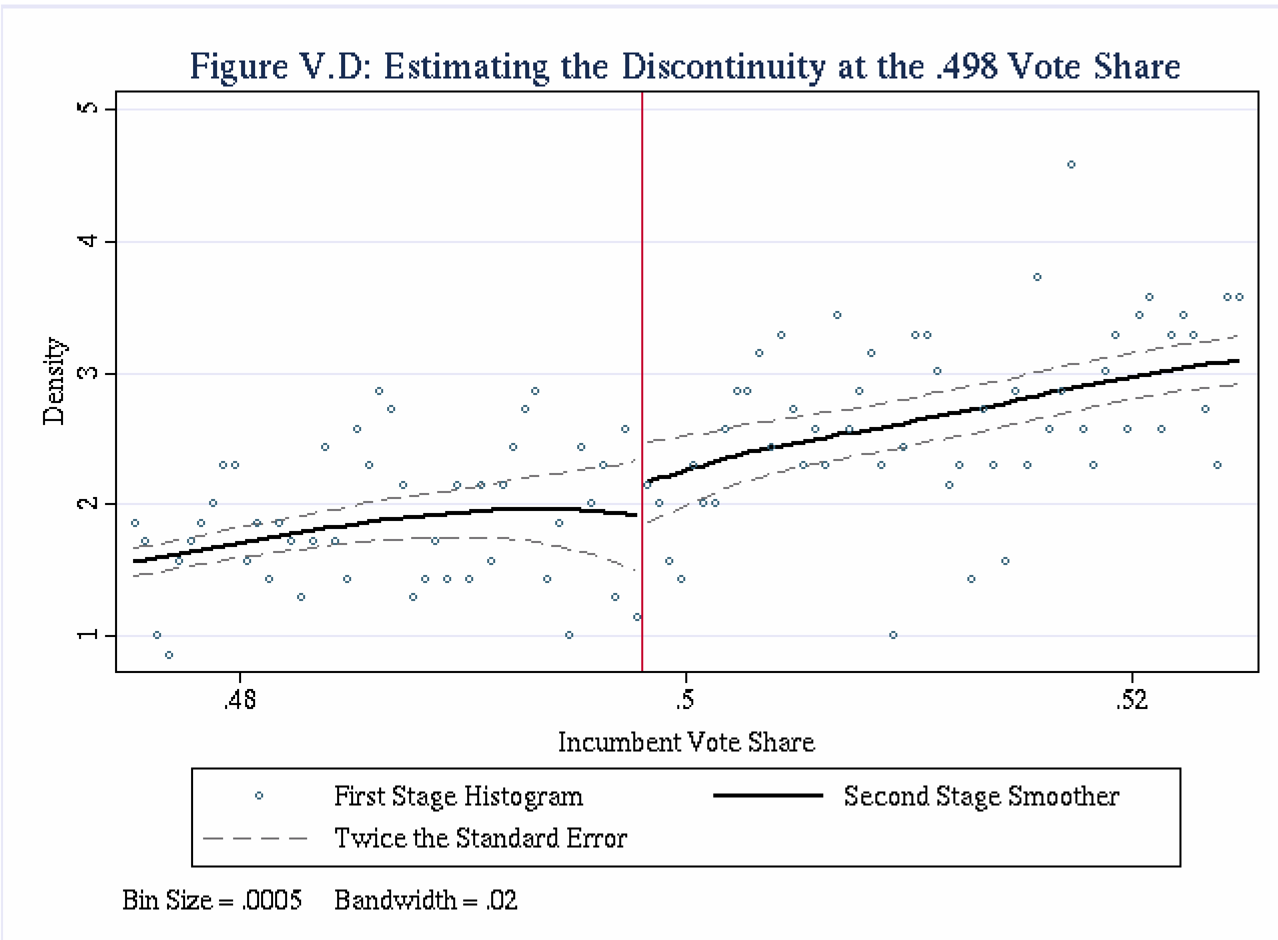


Figure V.D: Estimating the Discontinuity at the .498 Vote Share
Density
1
2
3
4
5
.48
.5
.52
Incumbent Vote Share
First Stage Histogram
Second Stage Smoother
Twice the Standard Error
Bin Size = .0005
Bandwidth = .02

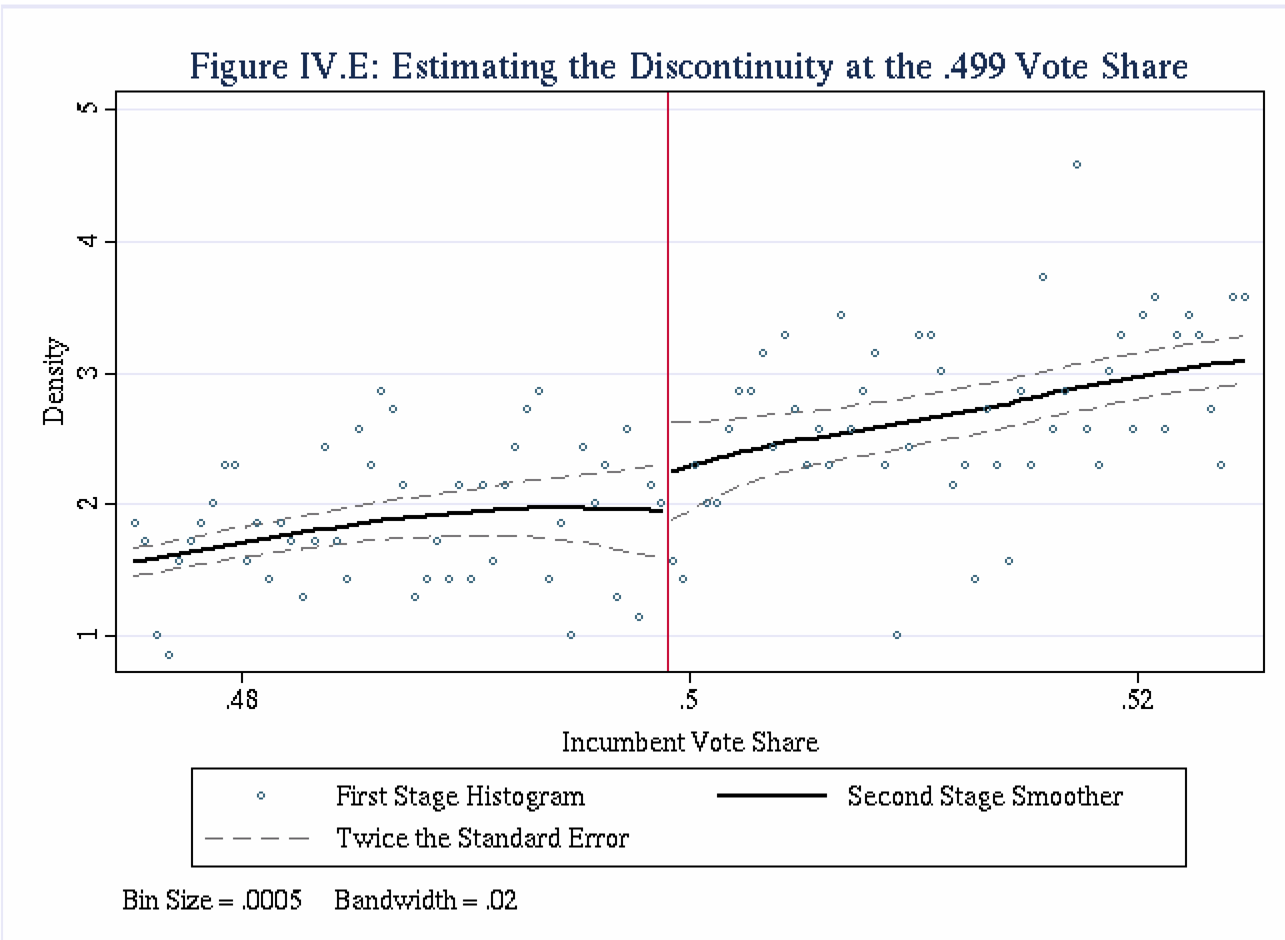


Figure IV.E: Estimating the Discontinuity at the .499 Vote Share
Density
1
2
3
4
5
.48
.5
.52
Incumbent Vote Share
First Stage Histogram
Second Stage Smoother
Twice the Standard Error
Bin Size = .0005
Bandwidth = .02

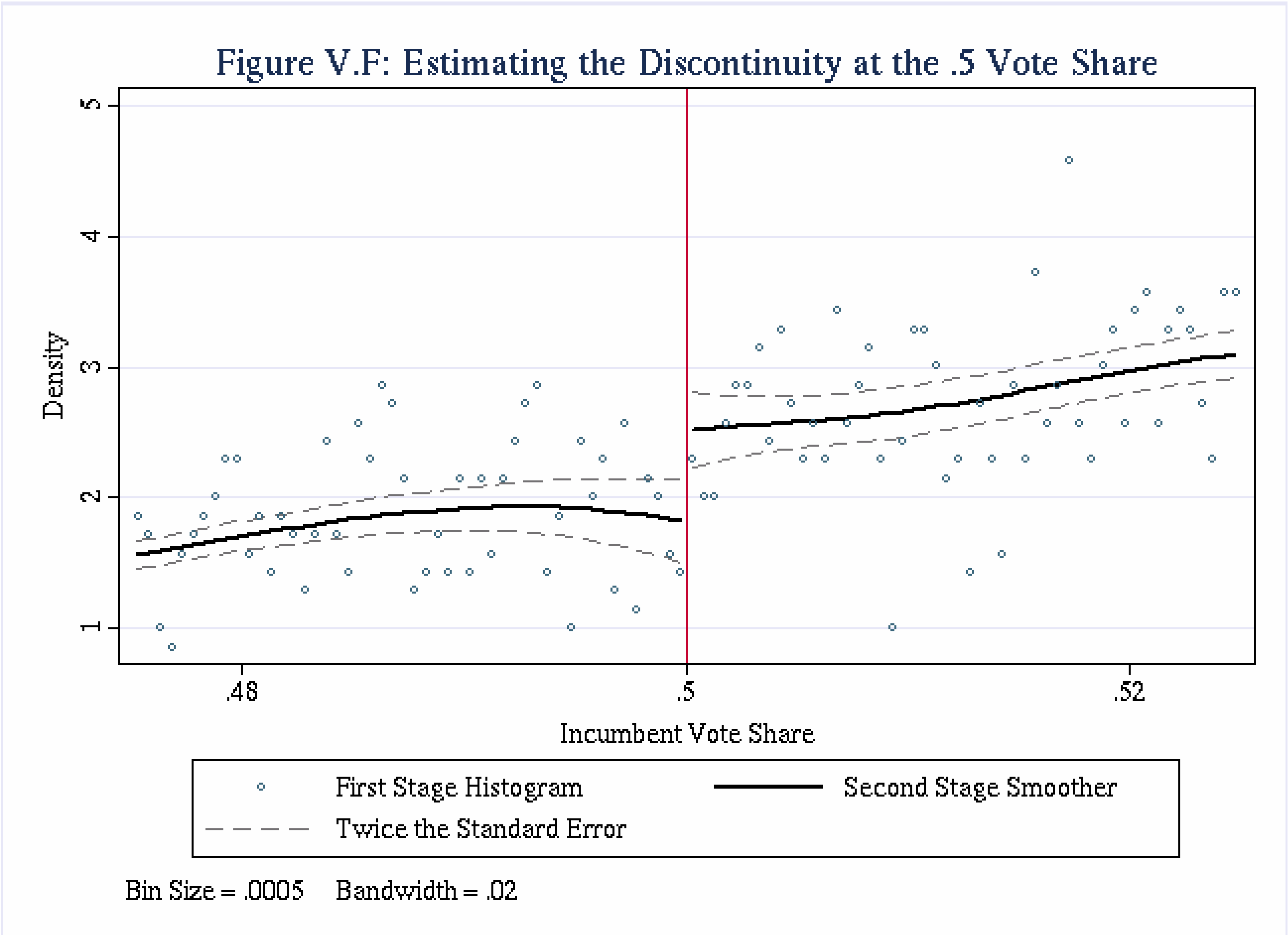


Figure V.F: Estimating the Discontinuity at the .5 Vote Share
Density
1
2
3
4
5
.48
.5
.52
Incumbent Vote Share
First Stage Histogram
Second Stage Smoother
Twice the Standard Error
Bin Size = .0005 Bandwidth = .02