



Working Papers

www.cesifo.org/wp

Precise Control over Legislative Vote Outcomes: A Forensic Approach to Political Economics

Ulrich Matter
Michaela Slotwinski

CESIFO WORKING PAPER NO. 6007
CATEGORY 2: PUBLIC CHOICE
JULY 2016

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

ISSN 2364-1428

Precise Control over Legislative Vote Outcomes: A Forensic Approach to Political Economics

Abstract

We propose a forensic approach to investigate the politico-economic forces that influence narrow vote outcomes in legislative assemblies. Applying nonparametric estimation techniques to a data set covering all roll call votes between 1990 and 2014, we can identify the existence of precise control over legislative vote outcomes in the U.S. House of Representatives. Several pieces of evidence indicate that this control seems to be, at least partly, driven by campaign finance donations. Moreover, control seems to be most prevalent in times of higher electoral competition, i.e. during election years. Our contribution sheds new light on the role of money in politics and, more generally, opens a novel perspective to empirical research on legislative voting.

JEL-Codes: D720.

Keywords: forensic economics, campaign finance, special interest groups, lobbying.

Ulrich Matter
University of Basel
Faculty of Business and Economics
Peter Merian-Weg 6
Switzerland – 4002 Basel
ulrich.matter@unibas.ch

Michaela Slotwinski
University of Basel
Faculty of Business and Economics
Peter Merian-Weg 6
Switzerland – 4002 Basel
michaela.slotwinski@unibas.ch

July 15, 2016

We are grateful to Joao Cunha, Patricia Funk, Armando Meier, Georg Nöldeke, Reto Odermatt, Marcus Roller, Mark Schelker, Kurt Schmidheiny, Thomas Stratmann, Alois Stutzer, Conny Wunsch, and conference participants at the CESifo Area Conference on Public Sector Economics 2016, the 2016 Meeting of the European Public Choice Society, the SSES Congress 2016, as well as seminar participants at the University of Basel, the University of Fribourg, and the University of Zurich for helpful remarks.

1 Introduction

A commonly held view of politics is that politicians are bound only to follow their voters' preferences as well as their own opinions and political ideology (Rice, 1924; Downs, 1957; Davis et al., 1970; Poole and Rosenthal, 1985). Political economics emphasizes the prevalence and relevance of additional forces and mechanisms that drive politicians' decisions when in office. For several decades, the economic analysis of politics has provided valuable insights into how politics works as a platform for exchange between individuals pursuing their own goals. As a central aspect of any representative democracy, legislative voting has been particularly in the focus of this research, taking into account the role of special interests, money, and lobbying. Contributions to this literature suggest that such external forces can influence legislators' decisions, i.e., buy votes, in order to procure a minimal winning coalition. However a broader assessment of how effectively and precisely vote outcomes can be affected or controlled is difficult to establish on the basis of the existing literature and methodology.

While many studies in the literature about legislators' voting behavior and the role of money in politics concentrate on specific votes in particular contexts when investigating how legislators' decisions are affected by external influences, we take a step back and start from a more global perspective. In this, we follow the forensic approach in economic research to identify hidden behavior in legislative decision making, taking a vast number of roll call votes into account. Forensic economics aims to detect hidden economic behavior by testing a null hypothesis which is in line with common economic theory in the absence of the investigated hidden action. This might, e.g., be a certain expected pattern in the data. Absence of the expected pattern violates the null hypothesis and indicates that an additional mechanism is at work which is not considered in the common theory.¹

In order to identify whether there is evidence for precise control over narrowly won votes in the U.S. House of Representatives, we first derive the link between the behavior of the vote outcome probability distribution and the presence/absence of precise control over legislators' votes: In the absence of precise control over legislators' votes the distribution of vote outcomes is continuous at any meaningful vote margin. In particular, we employ nonparametric estimation techniques to test our null hypothesis of 'no control', i.e. a continuous distribution of roll call outcomes, at the winning margin from a macro perspective. Following this line of thought, our finding of a discontinuous distribution of vote outcomes at the winning margin, just above 50%, indicates the existence of precise control over legislative voting behavior. It reveals that contested votes in the U.S. House are much more likely to be narrowly passed than narrowly declined.

In a next step we investigate whether this pattern can be explained by incentives set through campaign finance

¹ The term 'forensic economics', as we apply it here, was coined by Zitzewitz (2012) who surveys the related literature. There is another literature that also refers to forensic economics, applying the economic theory to a legal framework. Please note that we follow Zitzewitz (2012) in his definition of forensic economics.

contributions by special interest groups (SIGs). To assess whether money flows from SIGs to legislators are systematically related to the discontinuity in the vote distribution, we relate the timing of roll calls to the timed occurrence of campaign finance contributions. It is revealed that, on average, donations to congressmen around votes on bills that were passed with a very narrow margin are higher than around votes that were defeated by the same narrow margin. This finding suggests that precise control over vote outcomes is systematically related to donations. Moreover, the influence of money on the narrow passage of bills is particularly prevalent in election years and is mainly driven by corporate and party contributions. We not only find that donation flows react in general, but we document a specific time pattern of systematic donations around the date at which these particular votes take place. Specifically, systematic donations exhibit discontinuities in the days before the vote on a proposal, and thus before the outcome is established. Our empirical strategy draws its identification strength from the comparison of votes on bills that a priori (under H_0) should have a similar chance of passage. The methodological approach allows identifying specific patterns in the data which parametric methods would fail to uncover.

This paper is related to several strands of literature. First, it contributes to the politico-economic literature on money in politics that investigates how campaign contributions of special interest groups influence legislators' policy positions and voting behavior (see, for example, Kau et al. 1982 and Grossman and Helpman 1994 for theoretical considerations, Stratmann 1991 for empirical evidence, as well as Ansolabehere et al. 2003 and Stratmann 2005 for reviews of the literature). More specifically, our contribution adds to the literature aiming to identify the causal direction of donations on legislators' voting decisions. So far, it is generally unclear whether donors simply sympathize with and donate to politicians with political positions close to their own positions (i.e., donations are simply an expression of support for friends) or whether donations actually affect the politicians' positions (i.e., donations actually buy votes).² Second, it is related to contributions which advance a research question with the forensic approach (see references in Footnote 1) and more generally to the politico-economic literature on the analysis of congressional voting.

The contribution of this paper is threefold. First, we show that votes in the U.S. House are systematically more likely to be passed with a very narrow margin than they are to be defeated with that same narrow margin. Within the politico-economic and political science literature this finding poses a so far undocumented – but potentially important – stylized fact that deserves clarification. Second, this paper sheds new light on the role of money in politics, a strand of research in political economics that still involves many open questions. Finally, and more generally, this paper opens a novel perspective to empirical research on legislative voting.

² Bronars and Lott (1997) discuss this issue extensively. Contributions aiming to identify the causal direction of donations on legislators' voting decisions can only be found in the context of very specific policy domains with very few votes analyzed (see, e.g., Stratmann 2002; Dorsch 2011 for the case of financial services legislation).

The remainder of the paper is structured as follows. Section 2 develops our main theoretical considerations. Section 3 motivates and develops our empirical strategy. In Section 4 we disclose the data sources and describe the data used in our main analysis. Section 5 presents our main results. Finally, Section 6 offers concluding remarks.

2 Theoretical considerations on precise control

Our theoretical reasoning considers vote outcomes in legislative assemblies from a macro perspective. Thus, we hypothesise about how the distribution of these vote outcomes, i.e. the yes vote shares should behave at the winning margin, either with or without some form of precise control over legislators voting behavior. Therefore, we first derive the ‘natural’ behavior of vote outcome distributions in legislative assemblies at the winning margin, whereby the ‘natural’ state describes the situation in which legislators honestly vote according to their political positions and/or ideologies. The definition of this state leads to a null hypothesis that is tested against the data. The rejection of this Null would be evidence for some hidden action, in our case control over legislators’ votes, and the analysis proceeds by investigating what forces could lead to this deviation.

The derivation of our test for a precise control mechanism is based on our reasoning about the shape of the distribution of vote outcomes in legislative assemblies, meaning the probability density of yes vote shares. Taking the perspective of a spatial voting model (Downs, 1957; Davis et al., 1970; Poole and Rosenthal, 1985), in our intuitive and simple theoretical framework, the population of voters has a continuum of political preferences in various dimensions. Legislative assemblies consist of representatives elected by these voters. Legislators therefore represent a continuum of these political preferences. These legislators determine and draft the proposals upon which they vote. Consequently, the political positions of the proposals are also represented by a continuum, as they inherit the distribution of legislators’ positions. The distribution of the vote outcomes is a transformation of the legislators’ positions and decisions on which bill to agree to, which should translate to a smooth distribution of vote outcomes on these bills. Therefore, we would expect the vote distribution to be continuous at all meaningful vote thresholds in its natural state and thus in absence of any *deterministic* control over legislators’ votes.³ Importantly, even if there are forces influencing legislators’ positions as theorized in the current politico-

³ The pattern described is expected to hold in the U.S. House, as there is no strict (institutionalized) party discipline, but substantial heterogeneity in political positions within the two parties. There are notions in political science that in recent years party discipline in the U.S. Congress has strongly increased (particularly in the Republican party). However, there is no binding institutional rule forcing legislators to vote in line with the party or with other coalitions. If, for example, the German Bundestag were considered, it would not be surprising to observe a discontinuity in the vote outcome distribution. Members of the governing party coalition in the German Bundestag are bound by contract to vote in line with the coalition leaders’ decision. Thus it is more likely to observe vote outcomes with a yes-share that represents the governing coalitions’ share of seats than outcomes with a slightly lower yes-share. In that context, an important underlying ‘force’ driving the vote outcomes is the coalition contract posing strong incentives on coalition members to toe the line of the coalition leaders. However, in the U.S. Congress such a clear institutional incentive to follow party lines does not exist.

economic literature (such as informational lobbying, pork barrel politics, party leader influence, etc.), there is still a random chance that a legislator will deviate from the arrangement, which would also produce mass below the vote threshold. Consequently, as long as there is no force that has deterministic or at least very precise control over legislators' votes, we would expect the probability distribution of vote outcomes to be continuous at the winning margin (see Figure 1a). Please note that our understanding of the so called *external forces* is purposely very broad. It encompasses all politico-economic mechanisms that might influence legislators' voting behavior.

In order to describe what effect an external impact on legislators' votes has on the shape of the vote distribution, we consider forces that for example have a strong preference to pass bills, and thus try to influence legislators' votes to just provide the winning margin. If these forces have no precise control, they can increase the probability that a legislator will agree to a particular bill and therewith increase the probability that it will pass by increasing the yes share. But in this case, there would still be a random chance for a deviation of individual legislators. As a consequence, we would expect excess mass around the winning margin in the vote distribution, but no discontinuity. The lower the chance of a deviation, and thus the more precise the control over legislators' votes is, the higher would be the mass just above the winning margin compared to that below it. If control becomes almost deterministic, and given that there is a non ignorable number of such 'manipulated' votes, the chance to observe votes that narrowly pass would become disproportionately high compared to votes that were narrowly defeated, and if this mass is large enough the distribution could feature a discontinuity. In fact, it would lead to a positive discontinuity in the vote distribution, as indicated in Figure 1b. Some forces might, of course, also work in the opposite direction, towards a rejection of bills. If this is the case exclusively, the reasoning above is true vice versa. Under precise control we would expect a negative discontinuity at the winning margin. If both forces coexist and are equally strong there would be excess mass around the winning margin, but no discontinuity. If one force outweighs the other, and has precise control, the distribution of vote outcomes would again feature a discontinuity.

We underpin our theoretical considerations with a simulation of a stylized legislative process considering different levels of control precision by external forces and find that it supports our intuitive theory. In its natural state, meaning without any external influence, the vote outcome distribution is continuous at the winning margin. As long as the control over legislators' votes is not precise, and thus there is a random chance of deviation, the probability distribution stays continuous. However it would visually feature excess mass around the winning margin. The distribution becomes discontinuous as soon as the probability of deviation becomes marginal, i.e. the control over legislators' votes is almost deterministic. The simulation and its results are described in

Appendix A.I.⁴

Based on this reasoning we conclude that under the defined ‘natural’ state the probability distribution of vote outcomes is continuous at any meaningful winning margin, and a discontinuity would be evidence for some mechanism precisely controlling legislators’ votes. This reasoning translates to the basic null hypothesis to be tested with regard to the probability distribution.

‘No control’-hypothesis:

H_0 : The probability distribution of vote outcomes in legislative assemblies is continuous at the winning margin.

Thus, if the null hypothesis can be rejected, i.e., we find the probability density to be discontinuous at the winning margin, this is clear and rather strong evidence for precise control over legislators’ votes. If, however, the null hypothesis cannot be rejected, this does not automatically imply that there is no control over legislators’ votes. If there is no discontinuity, one might want to test for excess mass around the winning margin in order to get an idea about whether either control is not as precise or two forces are pushing in opposite directions. Note that we do not make any restrictions or claims concerning the overall shape of the vote distribution. It could be very skewed, symmetric, bimodal, or unimodal; this would not affect our theoretical proposition nor would it confound our empirical strategy.

There are, so far, few studies that consider precise control mechanisms and influence on legislative voting behavior from a macro perspective, and thus that theorize about its effects on the overall distribution of outcomes. Related considerations with respect to precise control over voting outcomes have been put forward by King and Zeckhause (2003). The authors’ theoretical work about precise vote manipulation by party leaders in narrow votes (by means of vote option; i.e. “if you need me” pledges or “hip-pocket votes”) predicts systematically fewer narrow losses and more narrow wins than under the natural vote distribution. They present circumstantial and descriptive evidence that is in line with the main prediction of their theory. Other theoretical insights are presented in Hummel (2009), who studies a setting where it is uncertain whether legislators will conform to an arrangement with a lobbyist when two opposing lobbyists compete for the vote. He predicts that lobbyists will bribe smaller majorities but pay higher bribes to the members of their coalition. This would also speak for the finding of narrow winning majorities.

The finding of a discontinuity in the vote distribution does not yet explain through which channel this mechanism might work. Our empirical approach does not depend on one specific theoretical consideration about channels

⁴ Please note that the assumption of a continuous distribution of political preferences within the population and the legislators is not a necessary condition for the resulting continuity under the natural state. As long as there is the chance of random deviations and variation in the acceptable distance on the bill level is not too low, such that no bill passes if political preferences are separated, the distribution without precise control is continuous at the pass margin. However, the continuity in preferences argument makes the concept intuitively accessible so we decided to retain it.

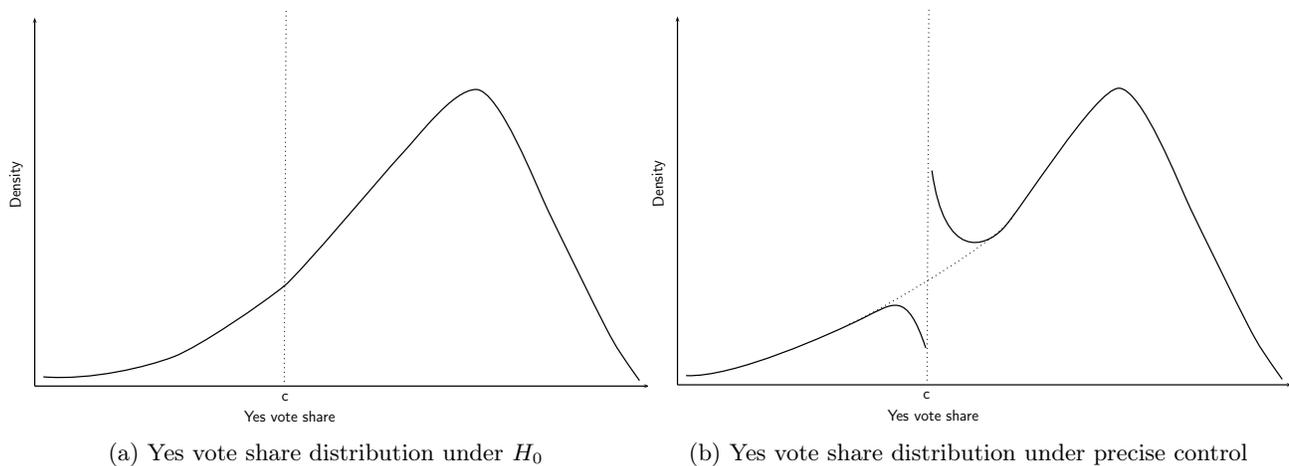


Figure 1: Vote distributions under H_0 and under precise control.

through which legislators' votes might be influenced, but allows us to test the outlined null hypothesis in a broad sense. There are many candidates from the politico-economic literature that could potentially be a driver and contributor to this precise control mechanism, and it is to be expected that in fact anything that can be observed from this macro perspective represents the result of a mixture of mechanisms. Yet, the a very prominent anecdotal and investigated explanation are financial incentives, i.e., the question of whether campaign finance donations in fact buy legislators' votes on specific bills. Donation flows could be part of a highly targeted incentivising and rewarding scheme by SIGs that have an interest in particular proposals. However, the evidence on this mechanism is mixed so far, the reason being that it is rarely evident whether a donation has been made because the donor likes the political position of a legislator and is close to his preferences, or whether the observed political position of legislators is close to that of the donor because he received donation flows (see, e.g., Bronars and Lott, 1997). We will thus use our empirical approach to investigate whether money can actually, at least in part, account for the observed precise control over narrow vote outcomes. There are, of course, possible alternative explanations for the observed pattern. Party leaders or party whips could apply pressure to pass bills they favour in a very precise manner, special interests could very precisely convince and inform legislators, or legislators could arrange a very reliable vote trading scheme. Furthermore, legislators might also bargain over the specific contents of bills such that the marginal legislator will just pass it because his interests are catered for.

3 Empirical strategy

We first describe the local likelihood density approach used to visualize the vote distribution and the testing for continuity at the winning margin. We further explain our approach to measure the characteristics of a potentially discontinuous vote distribution by applying bunching estimation in order to determine the excess mass above the threshold. Finally, the last part of this section illustrates how we investigate campaign finance donation flows as a potential determinant of a discontinuous vote distribution.

3.1 Discontinuity in density estimation

For nonparametric visual evidence on the distribution of yes vote shares we use local (linear) likelihood density estimation (LLD) (see, e.g., Loader, 1996; Otsu et al., 2013). Depending on the step of the analysis, we fit the LLD over the whole distribution of vote outcomes or from either side of the vote threshold, a yes-share just above 50%.

The probability density is estimated by solving the following maximization problems at each estimation point:

$$\max_{a,b} \left\{ \frac{1}{n} \sum K \left(\frac{x_i - x}{h} \right) (a + b(x_i - x)) - \int K \left(\frac{u - x}{h} \right) \exp(a + b(u - x)) du \right\}$$

where x is the estimation point in x_i , which in our application is the yes vote share. $K(\cdot)$ is a symmetric kernel weighting function and h the bandwidth used. The estimators of f are defined as $\hat{f} = \exp(\hat{a})$.⁵

To test H_0 , and thus whether the distribution of vote results is continuous at the vote threshold, we apply the discontinuity in density design proposed by Otsu et al. (2013).⁶

Following Otsu et al. (2013), we estimate the probability density separately from both sides of the threshold, using local (linear) likelihood density estimation. The point estimates of the density from the left (\hat{f}_l) and the right (\hat{f}_r) of the threshold result from solving the following two maximization problems:

$$\begin{aligned} \max_{a_l, b_l} \left\{ \frac{1}{n} \sum_{i: x_i < c} K \left(\frac{x_i - c}{h} \right) (a_l + b_l(x_i - c)) - \int_{u < c} K \left(\frac{u - c}{h} \right) \exp(a_l + b_l(u - c)) du \right\}; \\ \max_{a_r, b_r} \left\{ \frac{1}{n} \sum_{i: x_i \geq c} K \left(\frac{x_i - c}{h} \right) (a_r + b_r(x_i - c)) - \int_{u \geq c} K \left(\frac{u - c}{h} \right) \exp(a_r + b_r(u - c)) du \right\} \end{aligned} \quad (1)$$

⁵ More precisely for the graphical representation we use the local likelihood implementation (locfit) in the Chronux software package for Matlab (Bokil et al., 2010).

⁶ See the original paper of (Otsu et al., 2013) for a discussion of the advantages of their method over the approach of McCrary (2008), which is often used for the discontinuity in density estimation in a RDD context.

where c is the threshold value in x_i . The threshold is in our case the investigated vote threshold. $K(\cdot)$ is a symmetric kernel weighting function and h the bandwidth used. The triangular kernel is used throughout the nonparametric estimates, as we are interested in estimates at boundary points and this kernel has good properties at such points (for more details see Hahn et al. 2001 or Fan and Gijbels 1996). The estimators of f_l and f_r are calculated as $\hat{f}_l = \exp(\hat{a}_l)$ and $\hat{f}_r = \exp(\hat{a}_r)$, and the discontinuity parameter θ is estimated by their difference $\hat{\theta} = \hat{f}_r - \hat{f}_l$.

Inference is based on the empirical likelihood function for the parameter of interest θ , which is constructed as follows:

$$L(a_l, a_r, b_l, b_r) = \sup_{\{p_i\}_{i=1}^n} \prod_{i=1}^n p_i,$$

$$s.t. \quad 0 \leq p_i \leq 1, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i g_i(a_l, a_r, b_l, b_r) = 0,$$

with

$$g_i(a_l, a_r, b_l, b_r) = \begin{pmatrix} (1, x_{i,h})'(1 - I_i)K(x_{i,h}) - \int_{x < c} (1, \frac{x-c}{h})' K(\frac{x-c}{h}) \exp(a_l + b_l(x-c)) dx, \\ (1, x_{i,h})' I_i K(x_{i,h}) - \int_{x \geq c} (1, \frac{x-c}{h})' K(\frac{x-c}{h}) \exp(a_r + b_r(x-c)) dx \end{pmatrix}$$

Where g results from the estimating equations (or sample moment conditions), the FOC of (1), I_i is defined as $I_i = \mathbb{1}\{x_i \geq c\}$ and $x_{i,h} = \frac{x_i - c}{h}$. The dual form of the log empirical likelihood ratio function is

$$lr(a_l, a_r, b_l, b_r) = 2 \sup_{\lambda \in \mathbb{R}^4} \sum_{i=1}^n \log(1 + \lambda' g_i(a_l, a_r, b_l, b_r)),$$

and the concentrated log local likelihood ratio function, (profile) empirical likelihood ratio, of the parameter of interest ($\hat{\theta} = \hat{f}_r - \hat{f}_l$) is

$$lr(\theta) = \min_{a_l, a_r, b_l, b_r: \theta = \exp(a_r) - \exp(a_l)} lr(a_l, a_r, b_l, b_r).$$

Under the null hypothesis $H_0 : \theta = \theta_0$ the empirical likelihood function asymptotically follows a chi-square distribution $lr(\theta_0) \xrightarrow{d} \chi^2(1)$. The null hypothesis $H_0 : \theta_0 = \theta$ for some θ can be tested by $lr(\theta)$ using $\chi^2(1)$ critical values.⁷ For a more in-depth discussion and derivations see the original paper of Otsu et al. (2013).⁸

⁷ More precisely for matters of implementation, to test against a particular value of θ_0 g_i is set to $g_i(a_l, \log(\theta_0 + \exp(a_l)), b_l, b_r)$.

⁸ For a more general introduction to empirical likelihood methods and their computational issues see Owen (2001) or Kitamura (2008).

The null hypothesis H_0 predicts that the distribution of vote outcomes is continuous at the vote threshold. Thus we test $f_r - f_l = \theta = 0$ or $f_r/f_l = 1$. The empirical likelihood ratio is invariant to the formulation of the null hypothesis, and thus is identical for both the absolute and relative, formulation of H_0 .

3.2 Decomposition of vote bunching

If H_0 can be rejected, there is evidence that some force influences vote outcomes very precisely. To evaluate how targeted this control is and how many votes are concerned we decompose the pattern observed in the data using the idea of bunching measures inspired by, e.g., Saez (2010), Chetty et al. (2011), and Kleven and Waseem (2013). These authors investigate bunching behavior in order to identify structural earnings elasticities with respect to taxes. The basic idea can be described as follows: There is a threshold along a variable through which a treatment is assigned. Bunching describes the fact that individuals or units which would a priori lie at the disadvantageous side of the threshold, are able to and do systematically manipulate the assignment variable such that they are narrowly on the preferential side of the threshold. In the case where the treatment above the threshold is preferable this leads to the observation that there are disproportionately many observations narrowly overshooting the threshold and few narrowly undershooting it. This empirical approach has a particularly strong identification strategy, as it can compare a priori similar groups.

In some respects, voting behavior under precise control could be described as a form of bunching. The identification strategy naturally transfers to our application, as we compare votes that should be a priori very similar in their chance of being accepted, and thus form appealing groups for comparison. The basic idea behind bunching estimation is the construction of a measure of excess mass at the preferential side of a kink or a notch (the threshold), and locally comparing this mass to the mass in absence of the kink, or notch. Our vote threshold could be seen as a notch in the sense that an incremental change in the yes vote share produces a discontinuous change in the outcome and that there are systematic interests for placing a vote on the preferential side of the notch or threshold. This approach relies on the construction of a counterfactual distribution, which we describe below (for discussion of approaches applied in this context, see Spencer and Selin, 2014).

We propose to decompose the discontinuous vote outcome distribution under precise control over legislators' votes (like for example that presented in Figure 1b) as illustrated in Figure 2. Thereby, the solid line is the estimate of the empirical density, P_h is the point estimate of the density at the threshold coming from the right hand side, P_l is the corresponding point estimate of the density coming from the left hand side, P_c is the estimate of the counterfactual distribution (dashed line) at the threshold, B is a measure of the excess mass (which measures the number of votes which would not be there under the counterfactual scenario) and M

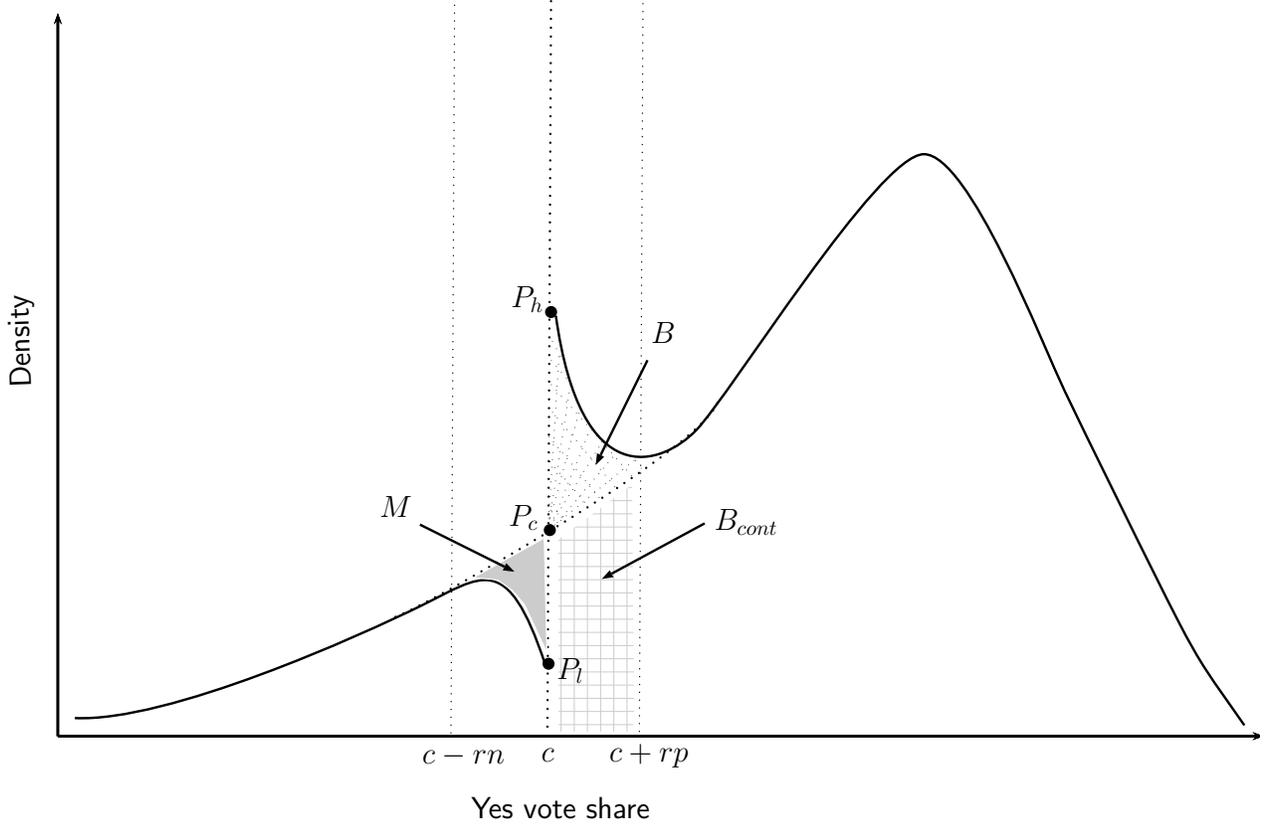


Figure 2: Yes vote share distribution under precise control. Decomposition of bunching in the vote distribution.

measures the missing mass (i.e., the number of votes which would have been on this side of the threshold under the counterfactual scenario).

For the estimation of these measures we mainly follow the strategy used in Chetty et al. (2011) and Kleven and Waseem (2013). They calculate the empirical income distribution by binned counts (f_c) and then estimate the counterfactual distribution by fitting a polynomial of order P to the empirical density (f_c), excluding observations in a “manipulation” range around an income threshold, defined as $(x - rn) < x < (x + rp)$. Instead of raw counts that have been used to estimate the empirical distribution in the past, we use a LLD estimate, separately from both sides of the threshold. Thus, in our case f_c is replaced by \hat{f} , the fitted values of a LLD estimation.

The estimated counterfactual distribution f_0 is then determined from the following regression,

$$\hat{f} = \sum_{p=0}^P \beta_p(x)^p + \sum_{j=rn}^{rp} \gamma_j \mathbf{1}[x = j] + \epsilon \quad (2)$$

where x is the respective vote share, P denotes the order of the fitted polynomial, and rn and rp stand for

the width of the excluded range to the left and right of the threshold. In our application we include three indicators on each side of the threshold whose coverage is defined by $rn/3$, or $rp/3$, respectively. The estimate of the counterfactual distribution is then constructed by $\hat{f}_0 = \sum_{p=0}^P \hat{\beta}_p(x)^p$, and thus the predicted values of this polynomial fit, ignoring the contribution of the indicators in the excluded range between rn and rp . The excess mass above the threshold (B) is estimated by the difference between the empirical and counterfactual distribution in the range between rp and the threshold c .

$$\hat{B} = \sum_{j=c}^{rp} (\hat{f} - \hat{f}_0)$$

The so called missing mass M is respectively calculated as the difference between the empirical and the counterfactual distribution in the range between rn and the threshold c .

$$\hat{M} = \sum_{j=rn}^c (\hat{f}_0 - \hat{f})$$

We would expect such missing mass to emerge if mainly votes that would normally be very narrowly defeated are manipulated precisely to pass the margin. If this missing mass is negative, there is also excess mass below the threshold. This would speak for imprecise control or two competing forces pushing in opposite directions. The remaining estimates are simply the point estimates of the empirical or the counterfactual distribution at the threshold c .

In addition to \hat{B} and \hat{M} , we will report the following statistics resulting from the econometric approach described: First, $\Delta(P_h P_c) = \hat{P}_h - \hat{P}_c$, which measures the discontinuity between the empirical and the counterfactual distribution at c and thus how many excess votes we observe precisely at the threshold. Second, $\Delta(P_h P_l) = \hat{P}_h - \hat{P}_l$, which again measures the discontinuity in the density at the threshold. The point estimates correspond to those in the discontinuity in density design. Third, $P_l = \hat{P}_l$, which measures whether the density just below the threshold is positive or whether we do not observe votes that are narrowly rejected. Additionally, we also report $\Delta(P_h P_c)/B$, which measures how many of the excess votes lie precisely at the threshold, and thus indicates how narrow the targeted majorities are. In other words, it measures which share of ‘manipulated’ votes is produced by minimum winning coalitions and which is produced by larger majorities. Finally, in a later step of the analysis, we need a measure of the mass below the counterfactual distribution within the manipulation range for the density, as a measure of the mass under the counterfactual scenario in the absence of control. We therefore additionally define B_{cont} by $\hat{B}_{cont} = \sum_{j=c}^{rp} (\hat{f}_0)$. Standard errors for the constructed measures are computed with a nonparametric bootstrap, using the bias corrected percentile method.

3.3 Determinants of a discontinuous vote distribution

For evidence on the determinants of the discontinuity in the vote distribution, we apply the estimation approach of a classical sharp regression discontinuity design, testing for a discontinuity in potential determining factors (y) at the threshold value. While there is no local randomization around the threshold, if there is bunching in the vote outcome distribution, a discontinuity in related determinants can tell us something about the drivers of the observed behavior. Therefore, we fit a local linear regression (LLR) separately from both sides of the threshold to see whether the behavior of the density could be explained by the movement of the determinant at hand.⁹ We apply the following estimation equation to every estimation point separately to both sides of the threshold (c).

$$\min_{\alpha\beta} \sum_{i=1}^N (y_i - \alpha - \beta(x_i - x))^2 K_h(x_i - x) \quad (3)$$

Where y is the dependent variable, and thus the determinant we are investigating, x_i represents the assignment or running variable - the yes vote share, h represents the bandwidth used $h > 0$, $K_h(x_i - x)$ represents the kernel weighting function at the estimation point x and is defined as $K_h(x_i - x) = \frac{1}{h} K\left(\frac{x_i - x}{h}\right)$. The discontinuity is then estimated by the difference of the limits of \hat{y} , coming from the right $\hat{\alpha}_r$ and coming from the left $\hat{\alpha}_l$, calculated as $\Delta(y) = \hat{\alpha}_r - \hat{\alpha}_l$. Where the subscripts indicate the point estimates at the threshold (c), either coming from the left l or coming from the right r , respectively. To be consistent, we again apply the same bootstrap procedure, as for the decomposition of the discontinuity in the density, to produce confidence intervals.

The primary determinant we investigate in this paper is campaign finance money donated to legislators around the vote dates. Consequently, the central dependent variable in our application y is *Money*, i.e., total daily campaign finance contributions around the vote date.

While the discontinuity in total daily donation flows is the main parameter of interest in the later empirical application, for one entity we aim to estimate, we need to estimate the counterfactual donation flows around the threshold, had there not been precise systematic influence. We approach this by using the same idea as in the estimation of the counterfactual density at the threshold value. In order to do so, we fit a local linear regression to total daily donation flows, as for the graphical representation and the discontinuity estimates. We then fit a polynomial, again including indicators to exclude the ‘manipulation’ range between $(x - rn) < x < (x + rp)$, to the local linear fit of donations \hat{m} , and calculate the equivalent measures as in the density case. As before, we estimate the discontinuity in donation flows at the winning margin $\Delta(PM_h PM_l) = \Delta(y)$, the discontinuity in

⁹ See, e.g., Porter (2003) or Hahn et al. (2001) for an introduction to RDD, and Fan and Gijbels (1996) for an introduction to LLR.

average total daily donation flows between the counterfactual and the empirical donation curve at the threshold $\Delta(PM_h PM_c)$, and the excess donation flows above the winning margin BM . Additionally, we need a measure of the counterfactual donation flows BM_{cont} as a measure of donations under the counterfactual scenario. Consistently with the distribution case we define BM_{cont} as $\hat{B}M_{cont} = \sum_{j=c}^{rp}(\hat{m}_0)$. All the measures of donation flows are estimated equivalently to those described in Section 3.2 just replacing the density estimate \hat{f} by a local linear estimate of total daily donation flows \hat{m} . Figure A4 in the Appendix visualizes these additional measures. Table 1 summarizes the most important generated measures that will be repeatedly used within the analysis.

Table 1: Summary of important measures

<i>Probability Density of Vote Outcomes</i>		<i>Donation Flows</i>	
$\Delta(P_h P_l)$	Total discontinuity in density	$\Delta(PM_h PM_l)$	Total discontinuity in donation flows
$\Delta(P_h P_c)$	Discontinuity between the empirical and the counterfactual density	$\Delta(PM_h PM_c)$	Discontinuity between the empirical and the counterfactual donation flow function
B	Excess mass above the counterfactual distribution $B = \sum_{j=c}^{rp}(\hat{f} - \hat{f}_0)$	BM	Total excess donations above the counterfactual donations $BM = \sum_{j=c}^{rp}(\hat{m} - \hat{m}_0)$
B_{cont}	Mass in manipulation range under the counterfactual scenario $\hat{B}_{cont} = \sum_{j=c}^{rp}(\hat{f}_0)$	BM_{cont}	Total donations under the counterfactual scenario $\hat{B}M_{cont} = \sum_{j=c}^{rp}(\hat{m}_0)$

The next section describes the data our empirical analysis is based on.

4 Data

We compile a data set covering all roll call records as well as additional information on all bills under consideration in the U.S. House of Representatives between 1990 and 2014. We link the roll call records to a vast data set covering all individual, party, and corporate/organization campaign finance donations to U.S. Representatives registered by the U.S. Federal Election Commission (FEC) during the same period of time. The compilation of our data takes advantage of newly available big public data sources providing highly detailed micro data and facilitating access for replication purposes. Our detailed micro data set on campaign donations allows us to assess the money flows to legislators on a daily basis. We can thus measure and statistically assess the timing

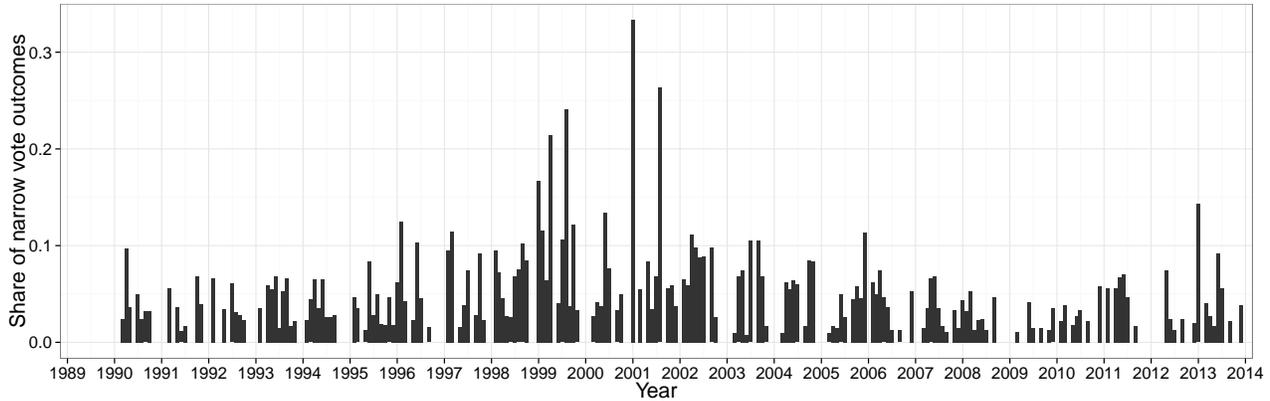


Figure 3: Share of narrow vote outcomes per month.

Data sources: GovTrack (primary sources: www.house.gov, Library of the Congress/THOMAS.gov, GOP).

of money flows with respect to the occurrence of voting outcomes in the U.S. House. This section discloses the data sources, briefly illustrates the structure of our data set, and presents some descriptive statistics of the key variables used in our main analyses.

4.1 Roll call records and bill information

Roll call data are based on the official roll call records published on the U.S. House’s website. All raw roll call data are provided by Civic Impulse LLC. (2015).¹⁰ Based on the official records, we compute the yes vote share for each vote in the House.

Of particular relevance for our empirical strategy are vote results with a very narrow outcome. Figure 3 illustrates how often narrow vote outcomes (results with a share of yes votes between 48% and 52%) occur over time in our main sample. The figure shows that there is no clear seasonal variation and that during most of the years in our observation period less than 10% of all votes were decided by a very narrow margin. However, during some congresses a substantial part of the votes was decided with a narrow margin.

¹⁰ Civic Impulse provides the information extracted from official U.S. government websites as Java Script Object Notation (JSON) data files via their website (www.govtrack.us). All data are freely accessible through GovTrack’s application programming interface. For large scale replication purposes, we recommend downloading the whole raw data set via GovTrack’s bulk download (see <https://www.govtrack.us/developers/data>).

4.2 Campaign finance contributions

The raw data on campaign contributions are provided by the Center for Responsive Politics (CRP) and are compiled from the Sunlight Foundation’s Influence Explorer database.¹¹ Each record refers to an individual donation and contains the exact date the donation was made as well as the type of donation and information on the contributor. Figure 4 shows the natural logarithm of the monthly aggregate flow of money to U.S. Representatives for the whole observation period in our study. The figure shows that overall campaign finance contributions have risen over the last decades. Contributions follow the seasonal pattern of the U.S. Congress, with peaks shortly before elections (i.e., November of even calendar years; indicated with red dashed vertical lines) and relatively few donations during the first months of a new congress.

For our main analyses concerning the role of donation flows, we compute the total daily campaign contributions donated to U.S. Representatives over an interval of 4 weeks (28 days) before and after each day a roll call vote took place. As a consequence we assign 57 data points¹², each representing the donation total for one day, to each roll call outcome we observe, whereby the voting date serves as the assignment indicator. We do this separately for three contributor types: corporate/organization, individual and party campaign donations.¹³ This allows us to investigate the time structure of donations and their origins.

5 Results

The results of our empirical analysis are presented in three steps. We start with testing our basic null hypothesis H_0 , i.e., that the distribution of vote outcomes is continuous at a certain vote threshold, to determine whether there is precise control over legislators’ votes. To this end, we apply the approach described above to the most common vote threshold or winning margin employed in the U.S. House (with 435 members), i.e., the simple majority at above 50%. Thus, the vote threshold in the following analysis is the first possible yes vote share above 50%, i.e., a yes vote share of 50.1149%. Votes with a share of 50% are rejected and votes exceeding this margin are passed. In a second step, we estimate the decomposition parameters of the observed bunching behavior. Finally, we investigate whether campaign donation flows are systematically related to the observed vote bunching, try to access the anatomy of this relation and offer one possible way to measure its scope.

¹¹ For large scale replication purposes we recommend using the bulk download of the full raw data set (around 10 GB) provided by the Sunlight Foundation under <http://data.influenceexplorer.com/bulk/>.

¹² These are composed of 28 days before the vote, 28 days after the vote, and the day of the vote itself.

¹³ See the data appendix (A.III) for a complete list of which donation types/recipient types (in accordance with CRP’s categories) are considered in each of our three main categories (individual, corporate/organization, and party).

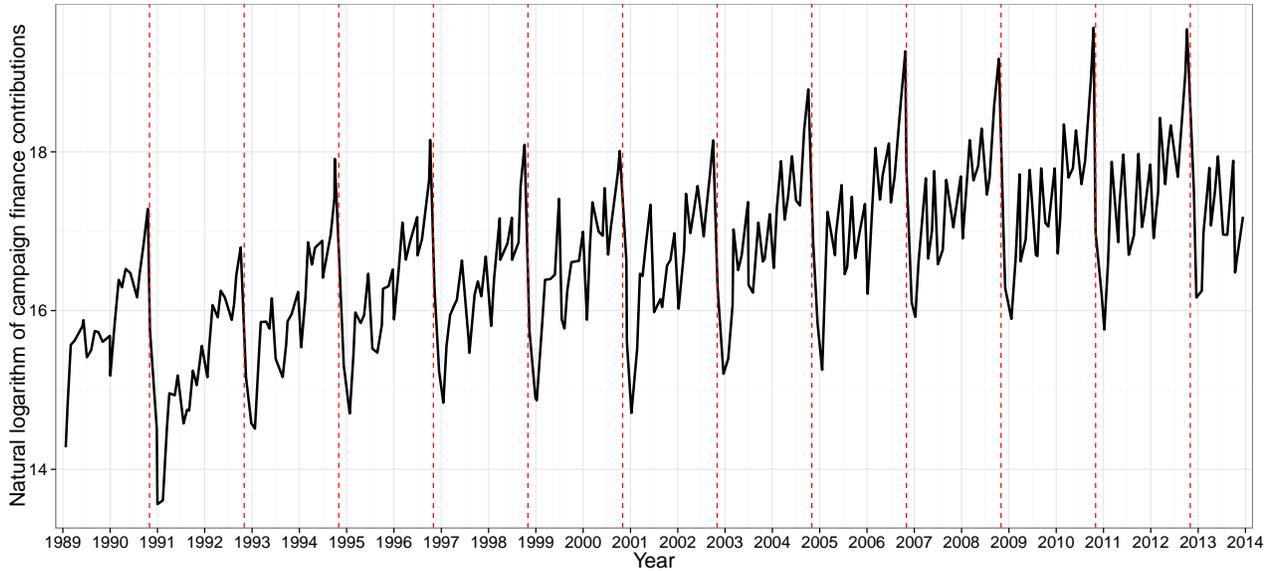


Figure 4: Monthly campaign finance contributions to U.S. Representatives.

Notes: The time series depict the natural logarithm of campaign finance contributions in monthly intervals. Campaign finance contributions are aggregated as the monthly sum of all donations (individual, party, and corporate/organization) to all respective U.S. Representatives (and respective candidates). The vertical dashed red lines indicate election months. *Data sources:* CRP and Sunlight Foundation.

5.1 Testing the continuity of the vote distribution

We first illustrate the vote outcome distribution with a simple histogram, presented in Figure 5. The vertical red line indicates the respective vote threshold of 50.1149%. Evidently, votes with a yes vote share of narrowly above the threshold seem to be disproportionately frequent. To see whether the distribution features a discontinuity, we estimate a LLD separately from both sides of the threshold, as introduced in Section 3.1, using a bandwidth of $h = 1$. The resulting graph can be found in Figure 6a. Visually the distribution seems to feature a clear positive discontinuity at the vote threshold. Votes seem to bunch above the threshold.¹⁴ This first piece of evidence would speak for a rejection of our null hypothesis. In order to test whether the observed discontinuity is systematically different from zero, we estimate the discontinuity in the density, based on the empirical likelihood approach described above. Results are reported in Table 2, where the corresponding discontinuity estimate for Figure 6a is presented in row I. We find that overall the discontinuity in the probability density of vote outcomes at the vote threshold is positive and significantly different from zero. Thus, we clearly reject the null hypothesis of continuity. Consequently, there seem indeed to be forces that can exert precise control over the outcome of legislative votes. The finding of a positive discontinuity suggests that forces that lead to passing bills outweigh

¹⁴ We obtain a qualitatively similar picture if applying the method used in McCrary (2008) to estimate discontinuities in densities by smoothing binned counts. The equivalent graph for Figure 6b is presented in Figure A10 in the Appendix.

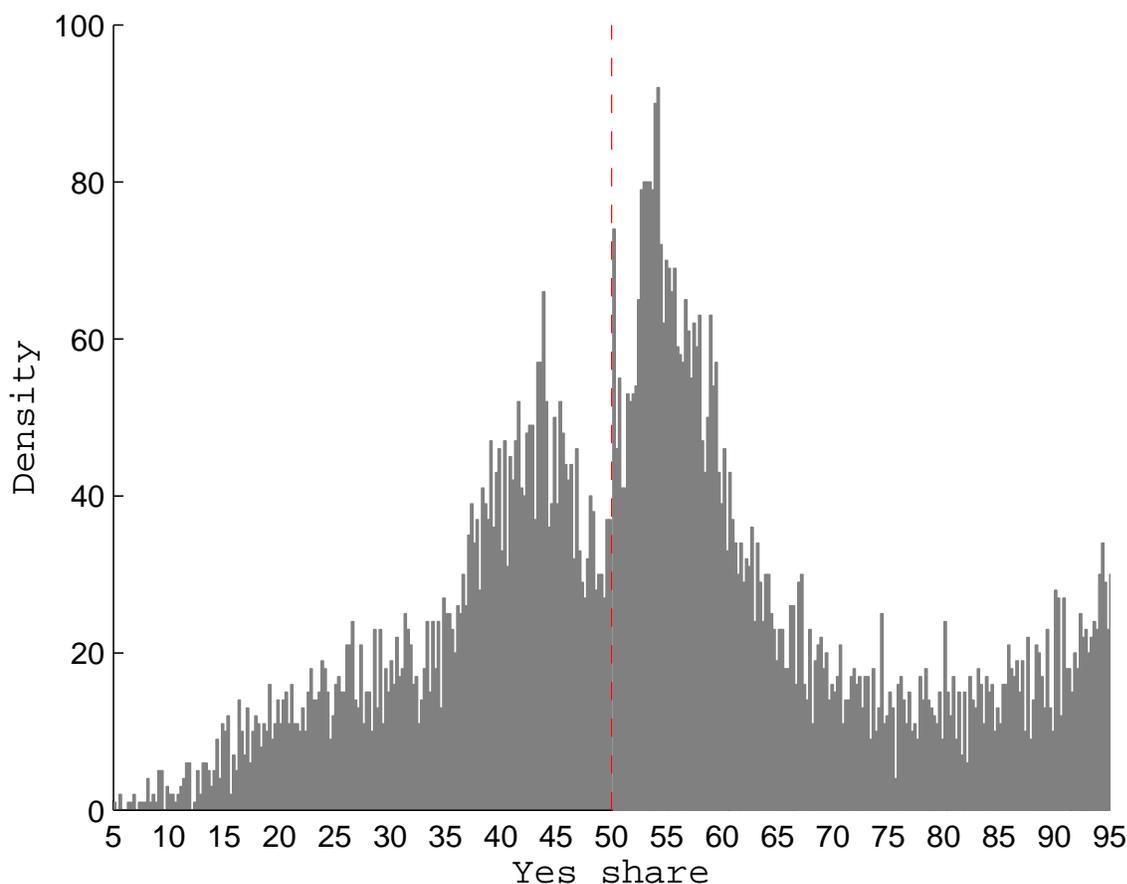


Figure 5: Histogram of overall vote distribution using a binwidth of 0.2, separately on both sides of the threshold.

forces that would lead to rejection of the bills. Thus, in the presence of forces that prefer a rejection, and exert control, our estimates can be understood as a lower bound of the real impact. Looking at the relative measure of the discontinuity \hat{f}_r/\hat{f}_l , the point estimate of the density just above the threshold is about two times higher than the point estimate of the distribution just below the threshold.

In a next step we distinguish between votes in election- and non-election years, as the scope of action for legislators might well differ depending on the intensity of electoral competition. On the one hand, interest groups have easier access to legislators during election years, as they meet more often (i.e., during fund raising events, etc.) and might thus more easily press their interests (in exchange for, e.g., financial campaign support). In particular, during election years legislators are in need of donations to finance their campaigns. This could make legislators more prone to the influence of specific interest groups. On the other hand, the discretionary leeway for legislators might be higher in years without elections, as there is less public attention and media coverage. Pressure to maintain their reputation might thus force legislators to make a trade-off between direct electoral support and

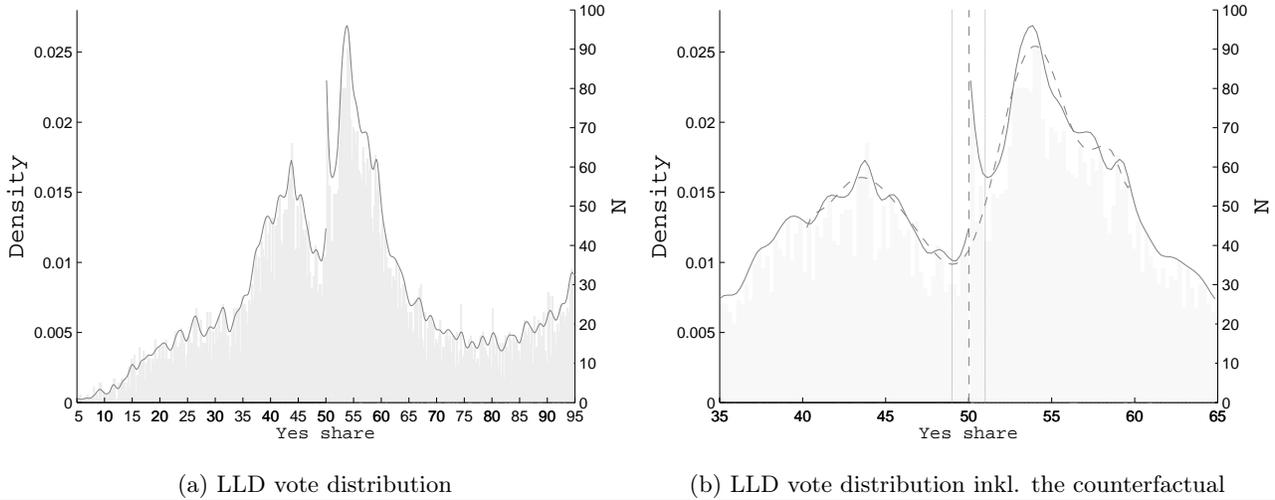


Figure 6: Overall vote distribution. The solid line represents the LLD fit to both sides of the threshold, the shaded area represents the histogram of the underlying data using a binwidth of 0.2 separately on both sides of the threshold, and the long dashed line indicates the fit of the counterfactual distribution (i.e., under the assumption that the distribution is continuous). The corresponding density discontinuity estimates can be found in row I of Table 2.

indirect electoral support through SIGs (in the form of, e.g., campaign finance contributions).¹⁵ This would speak for external forces having more influence on legislators’ voting behavior in years without elections. To see whether there is a difference, we repeat the estimates additionally for the subsample of votes in election and non-election years.¹⁶ The results can be found in Table 6, rows II and III. The distribution remains systematically discontinuous in the election year case, while the discontinuity is not significantly different from zero in the non-election year sample. The relative measure \hat{f}_r/\hat{f}_l in election years is about 3, while it is only about 1.2 in years without elections.

Summing up, we find that the vote outcome distribution in the U.S. House features a clear and systematic discontinuity at the winning margin. This means that the number of narrowly passed votes in the U.S. House is systematically higher than the number of votes rejected with the same narrow margin. This finding leads to the rejection of the null hypothesis of a continuous distribution, and according to our theoretical reasoning is evidence for the existence of precise control over legislators’ voting behavior.

¹⁵ See Matter and Stutzer (2015) for the role of public attention on how campaign contributions by special interests affect legislator’s policy positions.

¹⁶ House elections take place in even years. Thus, in our sample period these are 1990, 1992, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, and 2014. Thereby, all seats in the U.S. House and one third of the seats in the U.S. Senate are up for election.

Table 2: Density discontinuity estimates

	c	h	\hat{f}_l	\hat{f}_r	$\hat{\theta}$	\hat{l}_r	\hat{f}_r/\hat{f}_l	p-value	N
Overall									
I.	50.1149	1	0.0124	0.0230	0.0106	5.9466	1.8496	0.0147	12,404
Election year									
II.	50.1149	1	0.0089	0.0277	0.0187	8.5755	3.1053	0.0034	5,651
Non-election year									
III.	50.1149	1	0.0155	0.0193	0.0038	0.4232	1.2446	0.5153	6,753

Notes: Local likelihood ratio results for the discontinuity in the vote distribution. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.

5.2 Decomposition of discontinuity in vote distribution

Having found that the vote outcome distribution is discontinuous at the threshold investigated, we proceed by the proposed decomposition of vote bunching behavior. Therefore, we estimate the counterfactual density by fitting an 11th order polynomial between $x \geq 40$ and $x \leq 60$ to a LLD fit, separately from both sides of the threshold, using a bandwidth of $h = 1$, and excluding a range of 1 around the threshold ($rn = 1$ and $rp = 1$). The excluded range is determined visually at the point where the density estimate seems to start deviating (see, e.g., Figure 6b). The resulting counterfactual density is included in Figure 6b as a long dashed line, the excluded range is indicated by the two vertical dashed lines. The results of the decomposition estimates are reported in Table 3, where we provide 95% bootstrap confidence intervals. In the overall estimate the measure for the excess mass is significantly different from zero, while that for the missing mass is not. The estimate of the discontinuity at the threshold remains significantly different from zero using this estimation approach. There are about 357 ($B \times N$) potentially controlled votes, of which about 41% lie precisely at the threshold value and the rest lie somewhat above, providing a cushion for winning the vote. The distribution just below the threshold is still positive $\hat{P}_l \neq 0$. Comparing the point estimates for the subsamples of election and non-election years it is again clear that, as in the discontinuity estimation, control seems to be more pronounced in election than in non-election years.

While finding a force that is apparently able to exert rather precise control over legislators' votes is interesting in itself, it does not yet tell us anything about what this force is and through what channel the control is exerted. In principle there are many potential factors discussed in the politico-economic literature that could at least partly account for the observed regularity, e.g., party leaders' pressure, well informed and precise logrolling, or SIGs that somehow precisely inform and convince legislators. But a very prominent hypothesis would be that it is money from campaign donations that plays a role. This is the channel we want to concentrate on in the next

Table 3: Discontinuity in Density and Decomposition

	B	M	$\Delta(P_h P_c)$	$\Delta(P_h P_l)$	P_l	$\Delta(P_h P_c)/B$
Overall						
<i>parm</i>	0.0288	-0.0034	0.0119	0.0106	0.0124	0.4118
<i>lower95%</i>	0.0146	-0.0148	0.0055	0.0026	0.0079	0.2999
<i>upper95%</i>	0.0436	0.0086	0.0192	0.0187	0.0174	0.548
<i>h</i>	1					
<i>N</i>	12,404					
<i>Nboot</i>	1,000					
Election Years						
<i>parm</i>	0.0430	0.0051	0.0169	0.0187	0.0089	0.3933
<i>lower95%</i>	0.0205	-0.0107	0.0066	0.0057	0.0043	0.2436
<i>upper95%</i>	0.0690	0.0198	0.0306	0.0322	0.0170	0.5266
<i>h</i>	1					
<i>N</i>	5,651					
<i>Nboot</i>	1,000					
Non-election years						
<i>parm</i>	0.0173	-0.0106	0.0079	0.0038	0.0155	0.4582
<i>lower95%</i>	-0.0013	-0.0277	0.0009	-0.0082	0.0090	0.1484
<i>upper95%</i>	0.0347	0.0050	0.0170	0.0144	0.0259	1.8308
<i>h</i>	1					
<i>N</i>	6,753					
<i>Nboot</i>	1,000					

Notes: Estimates are calculated as described in Section 3.2. 95% bootstrap confidence intervals are reported.

step by investigating whether money flows from SIGs could, at least partly, account for this pattern in the data.

5.3 Precise control and donation flows

Based on the finding that the vote distribution is discontinuous at the winning margin, we investigate whether the campaign finance donations to legislators around the time of votes are discontinuous when comparing narrow votes just above and below the vote threshold at narrowly above 50%. Thus, we test whether donation flows resemble the behavior around the threshold we observe for the vote distribution.

Narrow vote outcomes around a 50% yes-share indicate that the bill was potentially contested, which in turn might indicate the particular attention of opposing SIGs actively attempting to influence the outcome. We would thus expect donation flows to peak around these contested votes, as there is ex ante no certainty whether the vote will be won or not and SIGs donate more to legislators around the votes on these contested bills to influence their voting behavior in one direction or the other. Thus, we might observe a peak in donations at the threshold, but not a discontinuity, if money was not a determinant of the precise control mechanism. If money flows to legislators feature a positive discontinuity at the particular threshold, this would be evidence for systematically

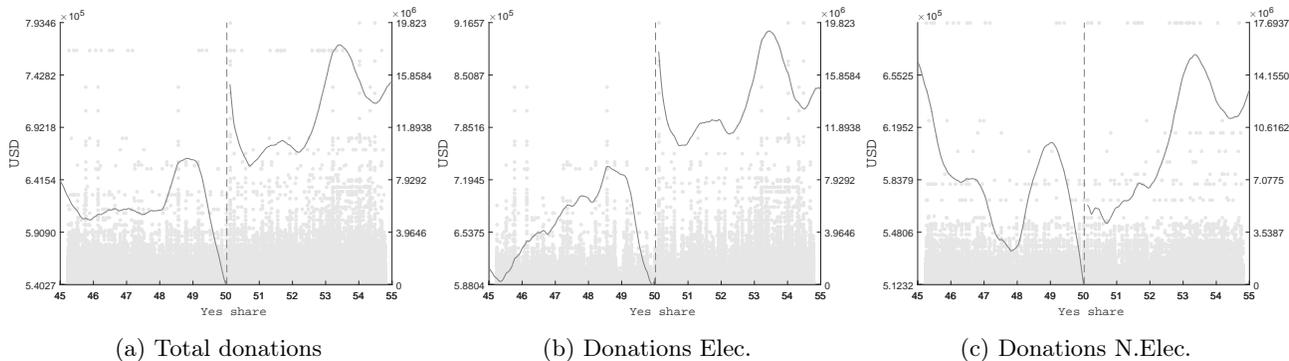


Figure 7: The solid line visualizes the local linear smooth of total daily donation flows within a time span of ± 4 weeks around the vote, the gray markers represent the raw data points using all votes in the sample, or votes in election and non-election years respectively. The left y-axis corresponds to the smooth and the right to the data points.

higher money flows around votes that are narrowly won, compared to those which are narrowly defeated. This pattern could be explained by the fact that donors knew in advance that the votes would be narrowly won and donated more around these votes. However, a precise prediction of a vote outcome should not be feasible; and even if it could be predicted, a certain pass of a vote would not provide any incentives to donate more around these votes if there were no incentives driven by a previous promise. Thus we would only expect a discontinuity in donation flows if donors are firstly part of the decision making process and thus have an expectation as to whether the vote will be narrowly won, and secondly donate more around these votes as part of a bargain with the legislators, and thus the incentive to donate more around these bills comes from a previous commitment. Therefore a discontinuity in total daily donations would lead to the conclusion that the excess narrowly passed votes were won *because* there were systematically higher money flows. Thus, a positive discontinuity in money flows at the threshold indicates that the discontinuity in the density is, at least in part, driven by money flows to legislators.

5.3.1 Behavior of donation flows around the winning margin

In order to investigate whether donation flows feature a discontinuity, we fit a local linear regression to the money flows over the vote share (x) separately to both sides of the threshold, using a bandwidth of $h = 1$. We define our dependent variable in this first step as the total daily donations to legislators between 4 weeks before and 4 weeks after the vote. Donation flows are assigned to specific votes through the donation date. As mentioned before, we link votes to donations at each day in the defined time range, thus around the date on which the respective vote took place, and then calculate the daily total of these donations for each vote.¹⁷

¹⁷ The advantage of this procedure is that each vote gets the same weight in the estimation, irrespective of whether there were many or few single donation flows.

Table 4: Discontinuity in total daily donations

	I	II	III	IV	V	VI	VII	VIII	IX
Sample									
Donations Years	Total Overall	Total <i>Elec.</i>	Total Non <i>Elec.</i>	Indiv <i>Elec.</i>	Indiv Non <i>Elec.</i>	Corp <i>Elec.</i>	Corp Non <i>Elec.</i>	Party <i>Elec.</i>	Party Non <i>Elec.</i>
$\Delta(\text{Money})$	193,638.84	284,856.84	54,092.21	56,232.15	34,926.58	148,152.61	-2,026.85	80,472.08	21,192.49
<i>lower</i> 99%	123,038.81	160,052.71	-29,210.03	4,776.61	-12,050.31	93,510.51	-34,502.07	38,716.89	-8,043.55
<i>upper</i> 99%	267,299.69	417,371.14	158,709.01	100,646.98	86,505.53	207,056.33	24,533.44	125,537.28	61,722.97
<i>h</i>	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
<i>N</i>	707,028.00	322,107.00	384,921.00	322,107.00	384,921.00	322,107.00	384,921.00	322,107.00	384,921.00
<i>Nboot</i>	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00

Notes: Local linear discontinuity estimates including 99% bootstrap confidence intervals. *Money* stands for total daily donations within a range of 4 weeks around each vote.

Figure 7a shows the local linear fit of the daily total donations within a time span of 4 weeks around the vote date for the whole sample of votes. Visually, donation flows feature a wing-like pattern around the threshold, and a positive discontinuity in total daily donation flows at the vote margin. Strikingly, we observe that on average total daily donation flows drop sharply for narrowly rejected bills and rise sharply for narrowly passed bills. It suggests that, on average, donation flows around bills that were narrowly passed are higher than around bills that were narrowly defeated, although they should a priori be equally contested. Thus, money flows strongly resemble the behavior of the vote distribution at the threshold. Fig 7b and 7c show the visual representation of the pattern in total daily donations around the threshold for election and non-election years separately. The wing like behavior around the threshold is striking in election and less distinct in non-election years.

In a next step we test whether these visual discontinuities are systematic by estimating the discontinuity in total daily money flows at the threshold value and report the results in Table 4 in columns I to III. Overall, the discontinuity in total daily donation flows within 4 weeks before and after the vote, is systematic and positive at the threshold. On average, total daily donation flows jump by about USD 190,000. Decomposing this pattern further and re-estimating the discontinuity separately for election and non-election years, we find that, as the graphs suggested, the discontinuity is primarily driven by election years. The discontinuity estimate in election years is positive and significantly different from zero. Donation flows in the 4 weeks before and after votes that narrowly passed are on average higher by about USD 280,000 than around votes that narrowly lost in election years. The point estimates are not significantly different from zero in non-election years.

So far we have analyzed total donations, adding donations from all contributions sources to legislators around votes. But we can additionally distinguish between three broad sources of donations: *individual*, *corporate/organization*, and *party* donations. While we cannot tell to what extent which source drives the observed behavior, it is interesting to see whether the wing-like discontinuous behavior is confined to one of them or is shared by

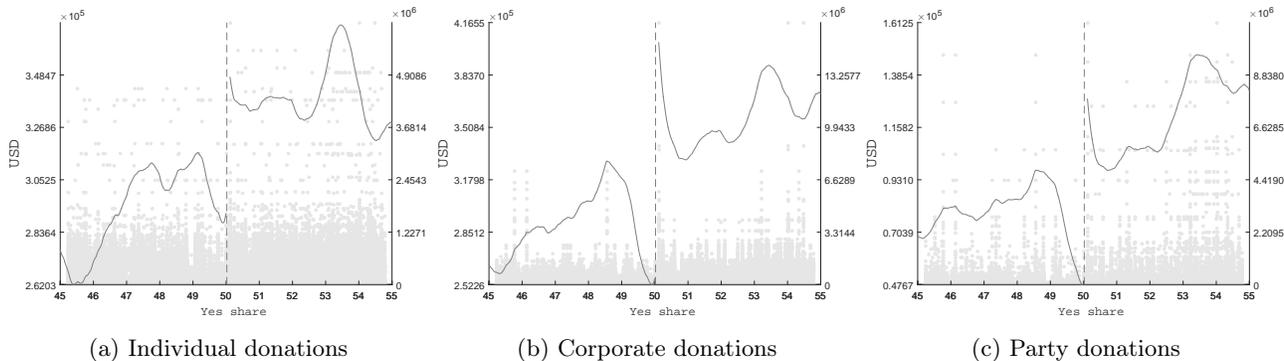


Figure 8: The solid line visualizes the local linear smooth of total daily donation flows within a time span of ± 4 weeks around the vote, the gray markers represent the raw data points using all votes in election years, and the total of individual, corporate/organization, or party donations respectively. The left y-axis corresponds to the smooth and the right to the data points.

all contributor types. Figure 8 shows the local linear fit of the daily total donations within a time span of 4 weeks around the vote date for the sample of votes in election years and all three contribution types. Figure 8a presents the pattern for individual, Figure 8b for corporate/organization, and Figure 8c for party donations. Visually, all three donation types feature a wing-like pattern around the threshold, and a positive discontinuity in total daily donation flows at the vote margin, while the wings are least distinct for individual donations. The respective figures for non-election years are presented in Figure A6 in the Appendix. Again the wing-like pattern seems to be specific to election years.

Columns IV to IX in Table 4 present the estimates of the discontinuity in total daily money flows at the threshold value for all three donation types. All feature systematic positive discontinuities in election and not in non-election years. Depending on the donation source, total daily donation flows on average jump by between USD 50,000 to USD 150,000 in election years. Corporate and party donations feature the largest discontinuities, although on average total daily donation flows around the margin are comparable between individual and corporate donations and individual donations are considerably higher than party donations on average. This would suggest that, if the size of donations is proportional to the influence, corporate and party donations drive most of the observed behavior.

If the variation in money flows is really due to strategic donating in order to influence legislators' voting behavior, we would not expect very low individual donation flows to drive our effect.¹⁸ These low donations are normally sympathy statements from private persons that arguably should not be meant to impact legislators' voting. This is exactly what we find when repeating our estimates for subsamples of 'low' individual donation flows, below

¹⁸ High campaign finance donations by individuals might still capture some special interests as shareholders, CEOs, and board members of specific companies might intend to specifically influence legislators through donations in the context of legislation that is particularly relevant for their industry. In contrast, very low donations by a broad range of individual donors are less likely the result of coordinated SIG efforts.

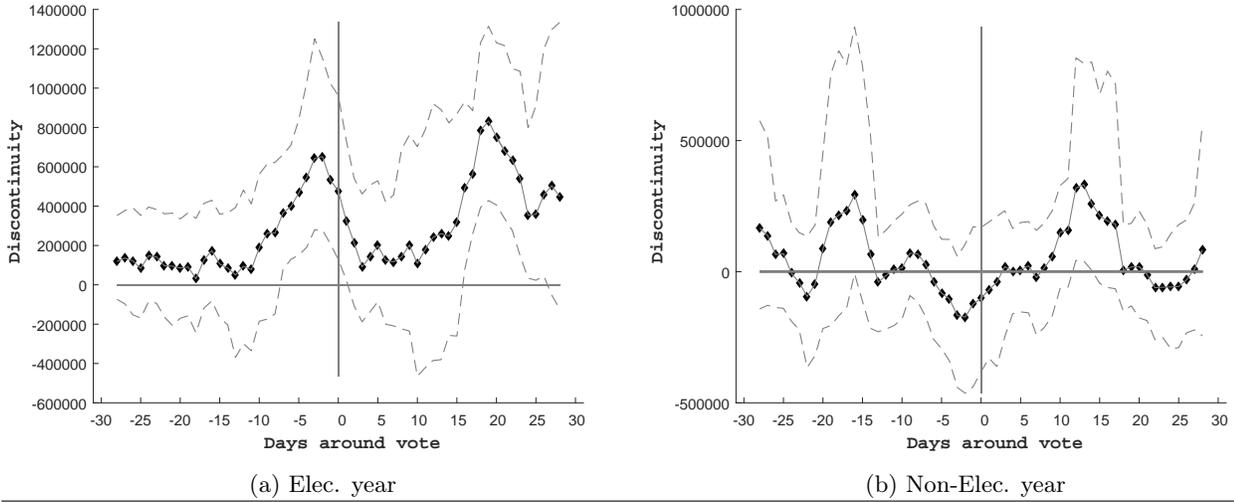


Figure 9: Rolling discontinuity estimates in total donations for election and non-election years within a 6-days band. The dots represent the discontinuity estimates, and the dashed lines represent the 99% confidence bands.

USD 30, 40, 50, and 100. The detailed results are presented in Table A1 in the Appendix.

Summing up, we find that there are on average systematically higher total daily donation flows to congressmen around votes that were narrowly passed compared to those which were narrowly defeated. This is especially true for election years. This result is very much in line with our observations in the vote distribution, where we found that the vote outcome distribution is discontinuous in election but not in non-election years, suggesting that the precise control mechanism is in force during election years. The results for the total daily donation flows resemble this behavior and suggest that donation flows are related to the passage of contested bills.

5.3.2 Time structure of discontinuities in donation flows

In a next step, we investigate whether there is a specific time pattern in the discontinuities within the 4 weeks before and after the votes. Therefore, we repeat the discontinuity estimates in a rolling manner, always including daily donations within a 6-day time band. Results for the sample of votes in election years are presented in Figure 9a and for non-election years in Figure 9b. Each dot represents a discontinuity estimate (using a bandwidth of 1%). The x-axis represents the ± 28 days around the vote, centered on the vote date. The estimates include the daily donations on the date they are assigned to on the x-axis and the 5 following days, such that, e.g., the estimate at $x = 0$ includes donations on the day of the vote itself and on the 5 following days.

We find that the donation flows show systematic discontinuities in the period about 10 days before the vote

Table 5: Discontinuity in total donations to legislators who vote yes or no

	I	II	III	IV	V	VI
Sample						
Years Voted (Yes/No)	Overall <i>Yes</i>	<i>Elec.</i> <i>Yes</i>	<i>NonElec.</i> <i>Yes</i>	Overall <i>No</i>	<i>Elec.</i> <i>No</i>	<i>NonElec.</i> <i>No</i>
$\Delta(Money)$	1,569,347.53	3,414,589.71	-597,052.38	29,604.97	4,131.03	-487,445.43
<i>lower</i> 99%	826,844.42	2,219,621.63	-1,369,124.79	-571,390.64	-954,338.32	-1,075,303.53
<i>upper</i> 99%	2,391,758.63	4,724,345.18	-10,272.40	516,403.02	872,973.61	33,516.41
<i>h</i>	1.00	1.00	1.00	1.00	1.00	1.00
<i>N</i>	707,028.00	322,107.00	384,921.00	707,028.00	322,107.00	384,921.00
<i>Nboot</i>	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00

Notes: Local linear discontinuity estimates including 99% bootstrap confidence intervals. *Money* stands for total donations that were made to legislators who voted yes/no on the respective day of the vote within a range of 4 weeks before or after the vote date.

until the day after the vote, and also in about the 15 to 24 days after the vote in election years. This could be interpreted to mean that money is a motivating device just before the vote takes place, and that legislators who stick to the agreement are rewarded after the vote passed. The corresponding time pattern for non-election years features few significant point estimates and does not reveal any clear pattern. The finding that those donations temporarily close to the voting days drive the effect are consistent with the findings of Stratmann (1995) and Stratmann (1998) who investigates the timing of donations and votes on bills to distinguish whether SIGs donate to support their preferred legislator or to impact legislators' voting behavior. By investigating several farm bills in the 1980s and 1990s he finds that the timing of bills and donation flows are strongly interrelated, which would speak for the attempt to influence voting behavior. The reasoning is, that if contributions were a pure appreciation statement, they should in no way be systematically related to the timely occurrence of votes on specific bills.

5.3.3 Money flows and support for bills

An obvious question is whether the positive discontinuity in donation flows that we observe is mainly driven by donations to legislators who approve the bill, i.e., if the control mechanism indeed works through pre-vote incentivising and post-vote rewarding donations, the discontinuity of donation flows to legislators who vote 'yes' should be at least stronger than for those who voted 'no'. Even if special interests trying to exert control donate more to all legislators around the contested votes, they would probably only reward those who kept their part of the bargain. Consequently, total donations to legislators voting in favor of the bill should feature a stronger discontinuity.

To see whether legislators who voted 'yes' really received higher donation flows and thus drive the discontinuity, we estimate the discontinuity in total daily donation flows within ± 4 weeks of the vote, over all donations

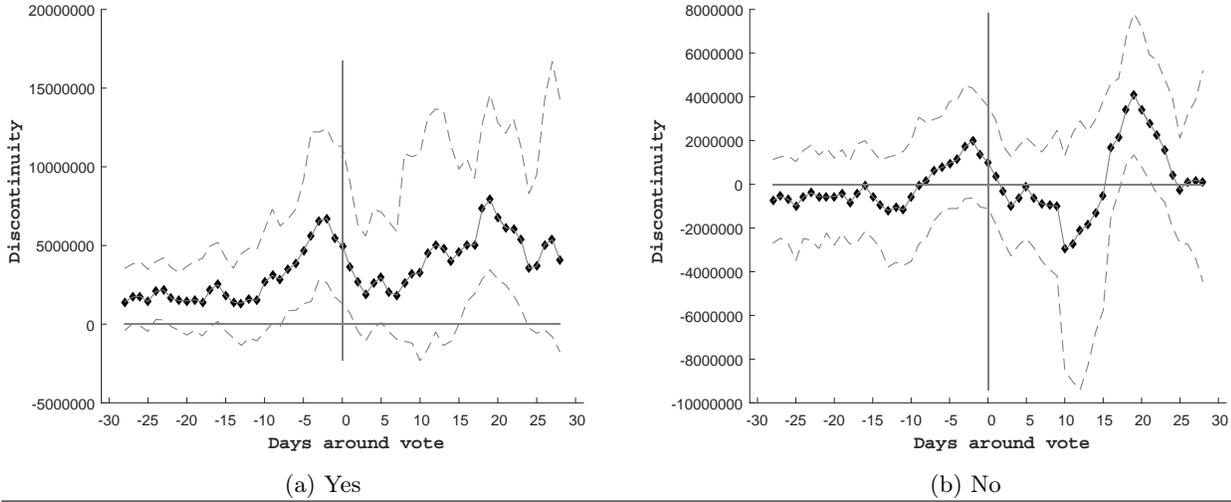


Figure 10: Rolling discontinuity estimates in total donations to ‘yes’ and ‘no’ voters within a 6-days band in election years. The dots represent the discontinuity estimates, and the dashed lines represent the 99% confidence bands.

made to legislators who voted ‘yes’ or ‘no’. Therefore we calculate the total of daily donations to legislators that approved one bill on a specific day and those who did not approve a bill on that day. This is of course an approximation, but as we can only assign donation flows to days and not directly to bills it is the closest thing we can do. In doing this, the systematic element should remain and the rest should only introduce some noise. Results for estimating the discontinuities in daily total donation flows to ‘yes’ and ‘no’ voters are summarized in Table 5. We find that only total donation flows to legislators who voted ‘yes’ in election years feature a positive significant discontinuity, and that there is no systematic discontinuity for legislators who voted ‘no’. This is a first piece of evidence that donors systematically target those legislators who are willing to pass the bill.

The graphic representation of the behavior of donation flows to ‘yes’ and ‘no’ voters around the threshold can be found in the Appendix (Figure A7 presents results for election years, and Figure A8 for non-election years). In election years donations to ‘yes’ voters feature a distinct positive jump and the expected wing-like behavior around the threshold, while donations to those who voted against peak around the threshold but feature no distinct discontinuity. We see that the observed pattern of donation flows in non-election years, although overall featuring a significantly negative drop at the threshold (see Table 5), does not look systematic at all. There is no distinct pattern around the margin. This is also confirmed by the rolling estimates in Figure A9 in the Appendix. If we take a small subset of donations, the significance in non-election years for donations to ‘yes’ and ‘no’ voters disappears, while donations to ‘yes’ voters in election years show the same time pattern as the overall donations investigated previously.

In principle, one could argue that such a positive discontinuity in total daily donations could be produced if

SIGs observe the voting behavior of legislators on bills they are interested in. If they observe that a legislator has approved such a bill they might be willing to support the legislator more. To produce the wing-like behavior around the winning margin they should be more willing to donate higher amounts if it narrowly passed than if it was passed with a higher margin. Abstracting from the idea that already passed bills should not be an incentive to donate more, there is another fact that cannot be explained by this line of thought. We find that donations feature a positive jump, even before the vote on the narrowly passed bills takes place. This, however, cannot be explained by the observed future voting behavior if the voting decision of legislators is not already known in advance with certainty.

Summing up, our results so far show that donation flows are, at least in election years, systematically related to the discontinuous behavior of the vote distribution. This suggests that money flows are one of the causal drivers of the precise control over legislators' votes and that the discontinuity in the vote distribution is, at least in part, driven by SIGs' financial campaign support to congressmen. Moreover, we find that donations that drive the discontinuous behavior mainly occur just the week before the vote and about two weeks after the vote has taken place, suggesting that legislators are first motivated and later rewarded for catering to the interests of SIGs. This is an important finding with respect to a pending question in the literature on money in politics. Our result can be seen as first macro evidence in line with the hypothesis that donations actually buy votes. The timing of donations would be irrelevant with respect to the timing of votes if it were the case that donors only support their friends in politics, and donations did not work as incentives to affect voting decisions in particular roll call votes. However, if it were the case that donations are, at least in part, aimed at convincing legislators to pass specific bills, we would expect the timing of donations to be systematically close to votes that pass with a very narrow margin (for a similar argument see Stratmann, 1998). The latter is exactly what we find.

5.4 Quantifying the impact: Elasticity of passing a bill with respect to donations

Given the documented effect of campaign donations on the narrow passage of votes, we want to investigate the question of how strong this effect is and how expensive it is to 'buy' a vote. We approach this question through a back-of-the-envelope calculation of the elasticity of passing a bill with respect to donation flows, and thus by a measure of how many percent just passed bills are increased as donation flows at the margin are increased by 1 percent. Therefore, we relate the measures of excess votes from the vote distributions to those of the excess donations from the money estimates.

We use a standard elasticity definition of the form:

$$e = \Delta_{votes}/\Delta_{money} \times Money/Votes$$

This approach implicitly assumes that all excess votes we observe have been produced by donation flows, thus ascribing excess votes that might have been produced by other forces to the money channel. As the decomposition into those bills driven by money and those by other forces is not possible, any measure we produce can only be an upper bound for the influence of money. Nonetheless, this analysis helps to get an idea of how large the influence might be under this assumption.

The measures that we combine for the elasticity calculation below were motivated in Section 3 and are summarized in Table 1.¹⁹

We use $rn = rp = 1$ and the bandwidth of 1% as before. In principle, there are several ways to define the components in e , as we can produce various measures of excess votes. One would be to measure e very locally at the winning margin. This would imply that the whole discontinuity in the vote distribution is produced by the whole discontinuity in donation flows, therefore, calculating $\hat{e}_1 = \Delta(P_h P_l)/\Delta(PM_h PM_l) \times PM_l/P_l$. Another way would be to take the discontinuity between the empirical and estimated counterfactual distribution and donation function as the change in votes and using the value of the counterfactual distribution at the threshold as the counterfactual level. Then the elasticity would be estimated by $\hat{e}_2 = \Delta(P_h P_c)/\Delta(PM_h PM_c) \times PM_c/P_c$. The underlying assumption would be that the excess votes at the threshold are produced by the excess money at the threshold value. However, these approaches ignore the fact that there are also excess votes somewhat above the winning margin. Therefore, we also propose a third measure that includes votes farther above the threshold by defining $\hat{e}_3 = B/BM \times BM_{cont}/B_{cont}$. This measure assigns all excess votes above the threshold to all excess donation flows within the manipulation range, taking the mass below the counterfactual outcome and distribution as the baseline.

We estimate all three measures of the elasticity on the basis of all votes in the U.S. House in election years. We use the systematic part of donation flows which resulted from the time structure estimates. These are total daily donation flows between the 10 days before and the day after the vote and the 15 to 24 days after the vote.

In a first step we calculate the elasticities using the total daily donation flow measure composed of all donations from all three sources to all legislators. Results for the three measures are presented in Table 6 in columns I to III, again reporting bootstrap confidence bands. In all three alternatives we find the elasticity to be around

¹⁹ Please note that we fit a 7th order polynomial between $x > 45$ and $x < 55$, to estimate the counterfactual donation flows around the winning margin. We use $rn = rp = 1$ and the bandwidth of 1% as before.

Table 6: Elasticity of passing a bill with respect to donations

	I	II	III	IV	V	VI
Elasticity type	\hat{e}_1	\hat{e}_2	\hat{e}_3	\hat{e}_1	\hat{e}_2	\hat{e}_3
Donations	Total			'Yes' voters		
e	2.79	2.33	2.46	1.41	1.50	1.58
<i>lower99%</i>	1.41	1.20	1.20	0.75	0.83	0.95
<i>upper99%</i>	7.73	8.15	12.82	3.34	3.33	3.13
<i>Nboot</i>	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00

Notes: Elasticity estimates during election years including 99% bootstrap confidence intervals. Donation flows between 10 days before and the day after the vote and the 15 to 24 days after the vote are used.

2.5. Thus, under the restrictive assumption that all excess votes are driven by the measured donation flows, a 1 percent increase in donations would at the margin increase the number of just passed bills by about 2.5 percent. In a second step we measure the elasticity with respect to donation flows directed to legislators that approved a particular bill. As documented above, these donation flows seem to drive the effect, and the discontinuity estimates are considerably larger. Using the more relevant money flows for an elasticity estimate can thus be considered a more conservative measure. Results for this donation measure are presented in Table 6 in columns IV to VI. We find the elasticity, as expected, to be smaller. Again all three measures are similar in size. The elasticity is about 1.4, which is still substantial if one considers that under a one percent increase in donations at the margin, excess passed votes would increase by 1.4 percent. The fact that all three approaches to measuring the elasticity render similar estimates suggests that donations and excess votes seem to evolve in a proportional manner. This additionally supports our conclusion that narrowly passed bills are systematically related to donation flows.

According to these results, donating at the margin can be seen as a potentially profitable and effective investment given that the stakes in passing certain bills are high and often lead to considerable redistribution for some groups. As the underlying assumption is rather strong, these results have however to be interpreted with caution. They are likely an upper bound estimate.

5.5 Robustness

In order to validate the robustness of our methodological approach and our results we perform several validation checks on the sample of votes in election years, which drive our main findings. First, we repeat our analysis of the density at several placebo thresholds without incentives for strategic influence. We do not observe any systematic discontinuities in the density at these placebo thresholds (see Figure A11 in the Appendix). Not only are these estimates insignificant, but are also much smaller than those at the investigated vote threshold. This

is also true for estimates of the discontinuity in total daily donation flows at the alternative thresholds presented in Figure A15 in the Appendix. The fact that we do not find any systematic response at thresholds without systematic incentives supports the causal interpretation of our findings and defuses concerns that our findings are driven by the estimation techniques used.

Second, we check whether our results could be driven by single influential observations or years. In Figure A12 we present the discontinuity in density estimates in which we drop one year at a time. The estimates are rather stable and stay significantly different from zero at the 5% level. We also repeat this strategy for the discontinuity in donation estimates, using the time range between two days before up to three days after the vote in election years. Again, the results do not seem to be driven by single observations or years, as the estimates stay rather stable (as shown in Figure A13 in the Appendix). Our results are moreover robust to the exclusion of the years 2001 and 2002, when seats in the U.S. House of Representatives were divided equally between Democrats and Republicans. Thus our results are not driven by a strict party alignment. Further, our results are also robust to the exclusion of votes in November of election years, where donations peak. This is also true excluding, e.g., votes between October and December.²⁰

As a last robustness check, we repeat our estimates for a range of bandwidth, and do not find a particular sensitivity. Results of this analysis for the discontinuity in density estimates can be found in Table A2 in the Appendix. The check for bandwidth choice sensitivity for the discontinuity in donation flow estimates, using the time range between two days before up to three days after the vote in election years, is presented in Figure A14 in the Appendix.

6 Concluding remarks

Contested bills in the U.S. House of Representatives are systematically more likely to pass with a very narrow margin than they are to fail with such a margin. This statistical regularity indicates that the passage of contested bills is influenced by a mechanism that exerts precise control over the actual vote outcomes. The finding poses a stylized fact, which, to our best knowledge, has not been studied in the politico-economic literature so far and which demands clarification.

We investigate the role of campaign finance contributions as one potential channel for the exertion of such precise control over vote outcomes. Our results indicate that money flows to congressmen can, at least in part, account for the observed discontinuity in the distribution of vote outcomes. We find that campaign contributions are systematically higher around the votes on bills that were passed with a very narrow margin than around votes

²⁰ The results on these additional analyses are available upon request.

that were narrowly rejected with the same very narrow margin. This positive discontinuity in donation flows seems to be mainly driven by donations to legislators who were willing to pass the bill.

The empirical analysis is driven by our suggestion of a forensic approach to political economics, which is particularly promising due to novel econometric methods and the increasing availability of large and highly detailed data on the political process. The econometric analysis applies nonparametric estimation techniques to a large data set including all roll call records in the U.S. House and all campaign contributions to U.S. Representatives between 1990 and 2014.

Our work suggests the need for further research in several directions. First, the finding of the existence of a politico-economic force that exerts precise control over vote outcomes deserves careful theoretical consideration as well as further empirical investigation. A potentially fruitful approach is to investigate the passage of bills in legislative assemblies in a cross-sectional setting in order to exploit the role of institutional differences. Second, with respect to the additional findings on the role of campaign finance contributions, the results presented can be seen as first macro evidence for a reward and motivation scheme based on legislators' voting behavior vis-à-vis a scheme in which SIGs simply channel financial support to legislators with policy positions close to their own. This is an important result, as it is generally difficult to empirically determine whether the causal direction runs from legislators' policy positions to campaign support from SIGs or vice versa. Consequently, empirical work in this context usually only considers very specific bills/policy issues which allow for a clear prediction of the direction of legislators' deviation if money indeed plays a role. It is, however, still unclear to what extent such financial awards affect policy making from a broader perspective. Future studies could apply our approach in order to investigate, for example, under what circumstances actual vote buying is more or less likely to occur. Third, financial incentives set by SIGs potentially only account for part of the documented exertion of control over vote outcomes. Future research, both theoretical and empirical, could investigate additional explanations. The theoretical contribution on the influence of party leadership by King and Zeckhauser (2003) might be a good starting point. Finally and more generally, we believe that the suggested forensic approach to the economic study of legislative voting has a lot of potential for other areas of political economics. The detection of hidden behavior is symptomatic of politico-economic research questions and the application of nonparametric econometric methods in combination with big public data might well shed new light on many of them.

References

- Ansolabehere, S., de Figueiredo, J., and Snyder, J. M. (2003). Why Is There So Little Money in U.S. Politics? *Journal of Economic Perspectives*, 17(1):105–130.
- Bokil, H., Andrews, P., Kulkarni, J., Mehta, S., and Mitra, P. P. (2010). Chronux: A Platform for Analyzing Neural Signals. *Journal of Neuroscience Methods*, 192(1):146–151.
- Bronars, S. G. and Lott, J. R. J. (1997). Do Campaign Donations Alter How a Politician Votes? Or, Do Donors Support Candidates Who Value the Same Things That They Do? *Journal of Law and Economics*, 40(2):317–350.
- Chetty, R., Friedman, J. N., Olsen, T. V., and Pistaferri, L. (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, 126(2):749–804.
- Civic Impulse LLC. (2015). GovTrack.us - About [online, last checked 2015-10-07]. Available from: <https://www.govtrack.us/about>.
- Davis, O., Hinich, M. J., and Ordeshook, P. C. (1970). An Expository Development of a Mathematical Model of the Electoral Process. *American Political Science Review*, 64(2):426–448.
- Dorsch, M. (2011). Bailout for Sale? The Vote to Save Wall Street. *Public Choice*, 155(3):211–228.
- Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 65(2):135–150.
- Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*, volume 66 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, London.
- Grossman, G. M. and Helpman, E. (1994). Protection for Sale. *American Economic Review*, 84(4):833–850.
- Hahn, J., Todd, P., and van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, 69(1):201–209.
- Hummel, P. (2009). Buying Supermajorities in a Stochastic Environment. *Public Choice*, 141(3-4):351–369.
- Kau, J. B., Keenan, D., and Rubin, P. H. (1982). A General Equilibrium Model of Congressional Voting. *Quarterly Journal of Economics*, 97(2):271–293.
- King, D. and Zeckhause, R. (2003). Congressional Vote Options. *Legislative Studies Quarterly*, 28(3):387–411.

- Kitamura, Y. (2008). Empirical Likelihood Methods in Econometrics: Theory and Practice. *Cowles Foundation Discussion Paper No.*, 1254.
- Kleven, H. and Waseem, M. (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.
- Loader, C. R. (1996). Local Likelihood Density Estimation. *The Annals of Statistics*, 24(4):1602–1618.
- Matter, U. and Stutzer, A. (2015). Public Attention and the Influence of Special Interest Groups: Policy Positions on SOPA/PIPA Before and After the Internet Blackout. *Mimeo*.
- McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, 142(2):698–714.
- Otsu, T., Xu, K.-l., and Matsushita, Y. (2013). Estimation and Inference of Discontinuity in Density. *Journal of Business & Economic Statistics*, 31(4):507–524.
- Owen, A. B. (2001). *Empirical Likelihood*, volume 92 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, Boca Raton and Fla.
- Poole, K. T. and Rosenthal, H. (1985). A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science*, 29(2):357–384.
- Porter, J. (2003). Estimation in the Regression Discontinuity Model. *Mimeo, Department of Economics, Harvard University*.
- Rice, S. A. (1924). The Political Votes as a Frequency Distribution of Opinion. *Journal of the American Statistical Association*, 19(145):70–75.
- Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 2(3):180–212.
- Spencer, B. and Selin, H. (2014). Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule. *Journal of Public Economics*, 109:36–49.
- Stratmann, T. (1991). What Do Campaign Contributions Buy? Deciphering Causal Effects of Money and Votes. *Southern Economic Journal*, 57(3):606–620.
- Stratmann, T. (1995). Campaign Contributions and congressional Voting: Does the Timing of Contributions Matter? *Review of Economics and Statistics*, 77(1):127–136.

Stratmann, T. (1998). The Market for Congressional Votes: Is Timing of Contributions Everything? *Journal of Law & Economics*, 41(1):85–114.

Stratmann, T. (2002). Can Special Interests Buy Congressional Votes? Evidence from Financial Services Legislation. *Journal of Law and Economics*, 45(2):345–373.

Stratmann, T. (2005). Some Talk: Money in Politics. A (Partial) Review of the Literature. *Public Choice*, 124(1/2):135–156.

Zitzewitz, E. (2012). Forensic Economics. *Journal of Economic Literature*, 50(3):731–769.

Appendix

A.I Simulation of vote outcome distribution

In order to underpin our intuitive theory about the shape of the distribution of vote outcomes, i.e., the probability distribution of yes vote shares in legislative assemblies, we perform a simulation study of a stylized legislative process. Therefore, we simulate vote outcome distributions in their natural state (under the assumption that legislators vote according to their political position and ideologies and there is no control over their voting behavior) and investigate how precise, or imprecise, control manifests in its shape. In this we are especially interested in the distributions' shape and behavior at the winning margin.

We simulate the legislative process as follows. We draw a population of 70,000 individuals represented on a one dimensional policy space from a normal distribution with a random mean between 0.4 and 0.8 and with a standard deviation of 0.5. As in the investigated legislative assembly, we assume that there are 435 legislators and that a vote is accepted if it exceeds the 50% winning margin, such that the first possible vote share that would pass a bill is 50.1149%. Legislators are represented by a random draw (with replacement) from the population and stay in office for two periods. In each period legislators vote on 350 bills, which are random draws (with replacement) of the political positions of all legislators. Whether a legislator for votes for or against a bill depends on the distance between his own political position and that of the bill. Thus legislators vote for a bill if it is not too far from their own political position. Distance is measured in x times the standard deviation (std) of the political positions of legislators, and x is a normal random number with mean 1.3 and standard deviation 0.5, such that the distance which they are willing to accept varies on the bill level. Legislators whose political position lies within the range of $\pm x \times std$ accept the bill. Legislators are absent on votes randomly with an expected value between 0 and 30 percent, this probability is randomly assigned to single votes. The whole procedure is repeated over 20 legislative periods, and thus leaves us with 14,000 vote shares, which form the vote distribution in its natural state.

In a next step, we introduce one possible form of 'manipulation' of vote outcomes. We assume that the manipulating agent ('force') has a one-directional preference to pass bills he is interested in. Thus, if the 'natural' a priori vote share on a bill he or she is interested in is below or equal to 50% the 'force' tries to push it above 50%. The bills that are targeted are chosen randomly from all bills with an a priori yes vote share below or equal to 50%, where the probability that it is targeted rises linearly the closer the bill is to 50% with a slope of 0.2%.

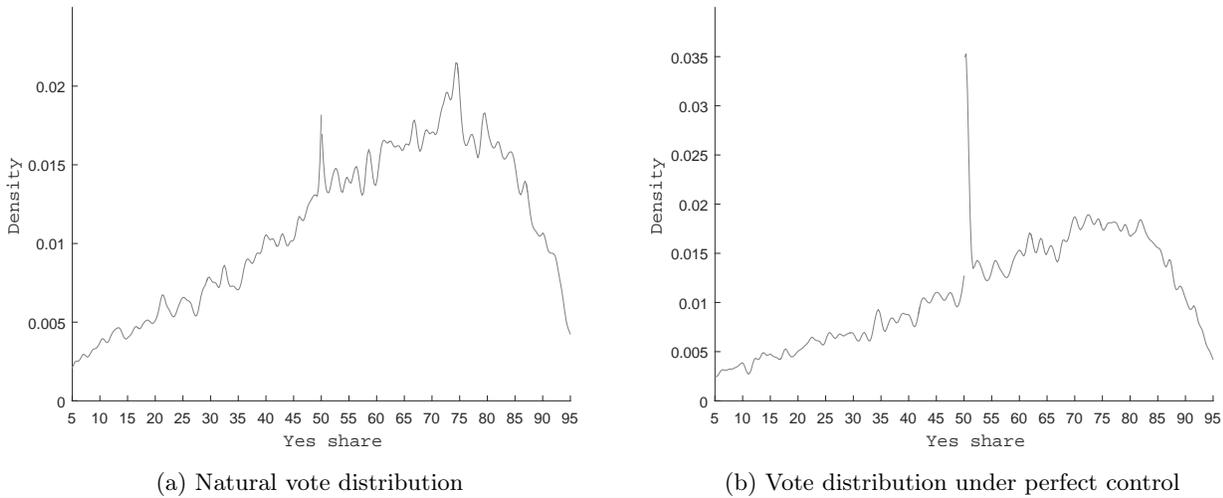


Figure A1: Simulated vote distribution in its natural state and under precise control. Local likelihood density estimates separately from both sides of the threshold.

The yes share that the manipulating agent targets lies somewhere close above 50%²¹ and is determined per vote by 50.1149% plus a random number between 1 and 2. Then, the difference in yes-votes between the a priori vote share and the targeted vote share is the number of legislators who have to be additionally convinced to pass the bill.

Under the most extreme scenario the ‘force’ has perfect control over legislators’ votes, and the probability that the manipulating ‘force’ manages to convince legislators is one. Thus targeted votes are placed somewhere just above 50% with certainty. Under imprecise control each legislator deviates from the agreement with a certain probability. Such that some legislators conform and some do not. Consequently, the probability of passing the bill increases, and the mass in the vote distribution is shifted into the direction of the threshold, but with no certainty whether it will exceed it.

Figures A1 and A2 show exemplary distributions for four cases. We simulate each scenario 200 times and calculate the proportion of runs in which a discontinuity that is significantly different from zero on the 99% confidence level is observed. Thereby, the continuity at the winning margin is tested using the discontinuity in density approach described in Section 3.1. Figure A1a shows the vote distribution in its natural state for one exemplary simulation run. It is estimated using a LLD separately from both sides of the threshold. Visually, there is no distinct pattern around the vote threshold, and a discontinuity is found in only 2.5% of the runs. Figure A1b presents an exemplary run for the distribution under precise control over legislators votes. It features a clear and distinct discontinuity at the winning margin. This observation is clearly confirmed by the simulation

²¹ The argument why a vote share of closely above 50% is targeted is simply that it is assumed that manipulation is a costly effort.

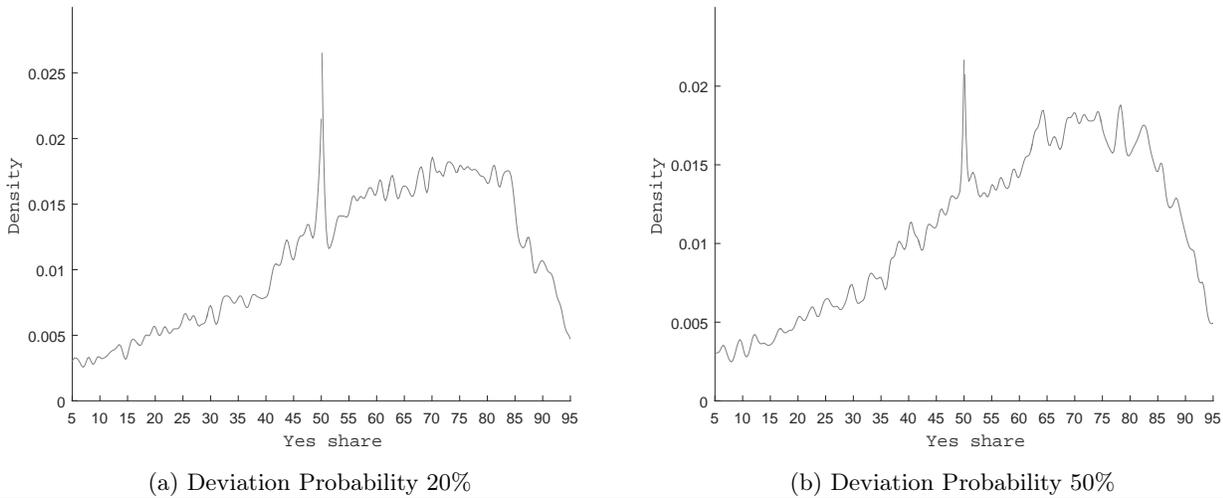


Figure A2: Simulated vote distribution for two level of deviation probabilities. Local likelihood density estimates separately from both sides of the threshold.

in which we find a discontinuity in 97.5% of the runs. Figure A2a shows the distribution in the case where legislators deviate with a 20% probability, thus the probability of conforming is still 80%. There appears to be excess mass around the threshold but no visible discontinuity. A statistically significant discontinuity is only found in 2.5% of 200 simulation runs. Finally, Figure A2b assumes a deviation probability of 50%, and thus whether a legislator conforms or not is a coin toss. Again there seems to be some excess mass around the threshold but no discontinuity. We do not find a discontinuity in any run.

In a next step we perform a fine-grained simulation, increasing the deviation probability by 1% in each step to get an idea of how precise the control has to be to produce a discontinuity in the distribution. The result is summarized in Figure A3. The probability of finding a discontinuity is rather high as long as control over legislators' votes is almost deterministic, and drops sharply as soon as the deviation probability exceeds about 2%.

This simulation backs our intuitive reasoning that control over legislators' votes needs to be very precise to produce a discontinuity in the vote distribution. Further, it illustrates that the continuity of the distribution does not automatically imply that there is no impact on legislators' votes. In this case, one might want to have a look at whether there is a striking excess mass around the threshold, which would imply imprecise control or two strong forces pushing in opposite directions.

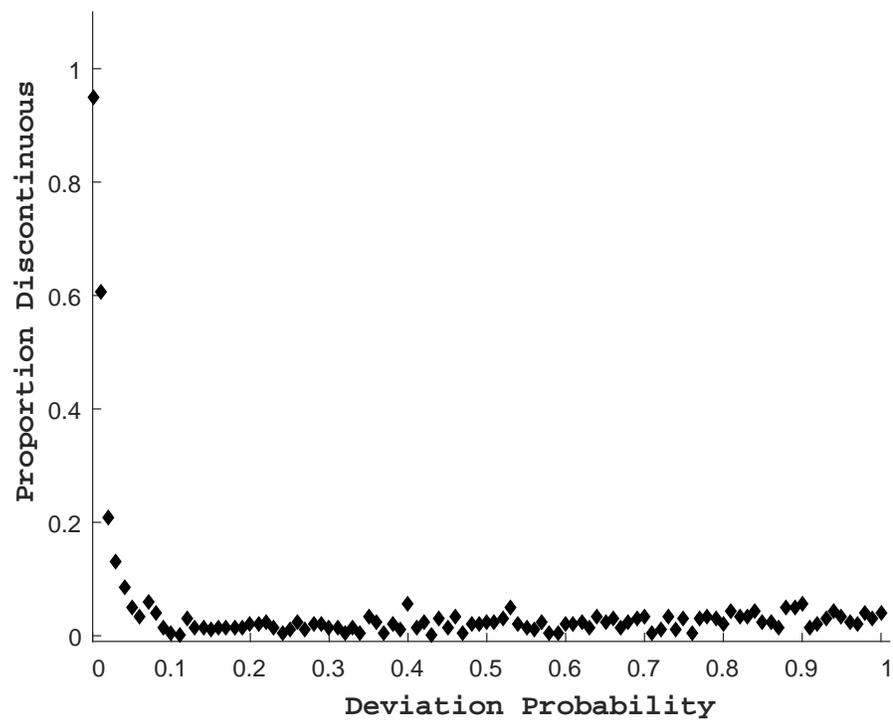


Figure A3: Probability of finding significant discontinuities in 200 simulation runs, increasing the deviation probability by 1%.

A.II Additional Materials

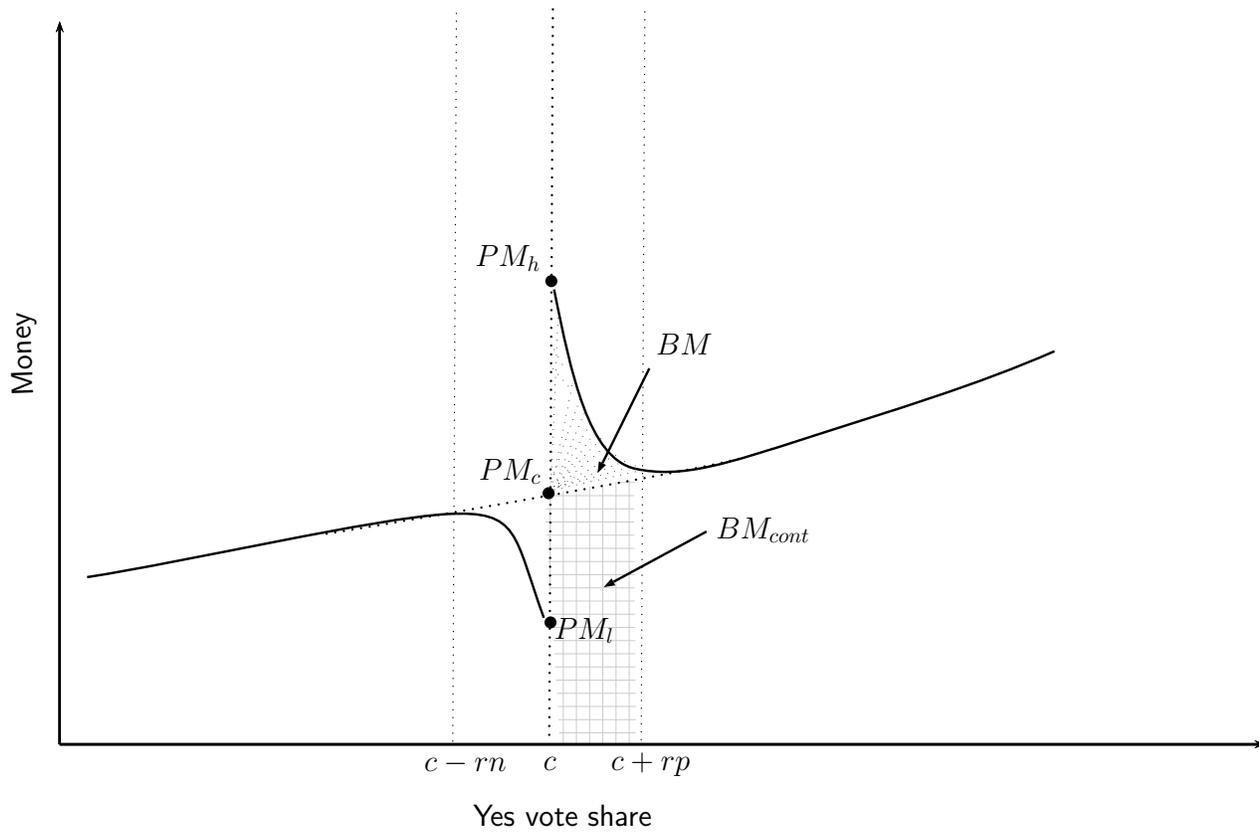


Figure A4: Function of donation flows (m) around the threshold, assuming that donations are related to the precise control mechanism. Decomposition of discontinuity in donation flows.

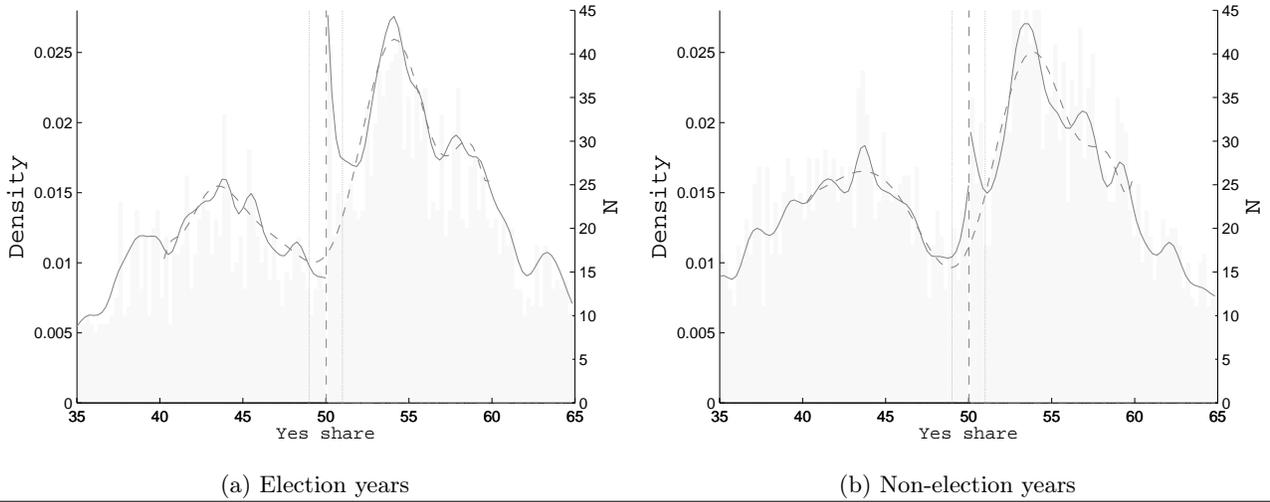


Figure A5: Vote distributions in election and non-election years. The solid line represents the LLD fit, the shaded area represents the histogram of the underlying data in 0.2 bins, and the long dashed line indicates the fit of the counterfactual distribution (i.e., if the distribution was continuous).

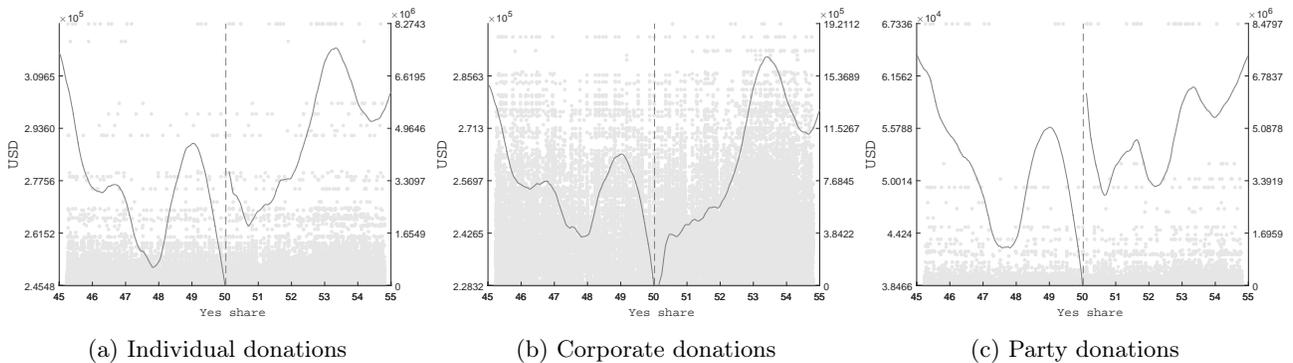


Figure A6: The solid line visualizes the local linear smooth of total daily donation flows within a time span of ± 4 weeks around the vote; the gray markers represent the raw data points using all votes in non-election years, and the total of individual, corporate/organization, or party donations, respectively. The left y-axis corresponds to the smooth and the right to the data points.

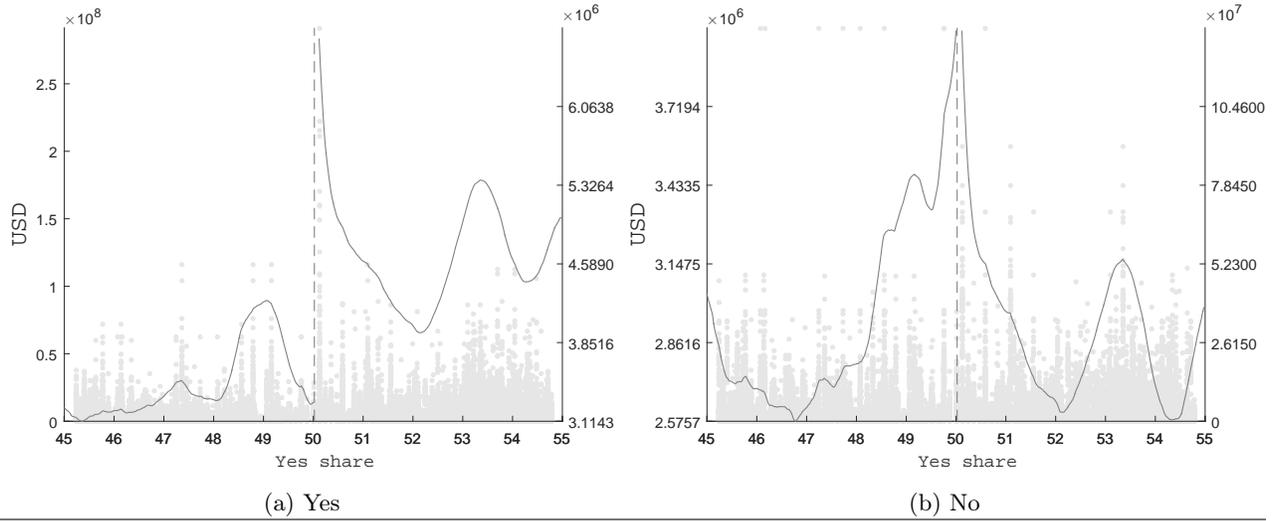


Figure A7: The solid line visualizes the local linear smooth of total daily donation flows within a time span of ± 4 weeks around the vote; the gray markers represent the raw data points using all votes in election years, and total daily donations to ‘yes’ or ‘no’ voters, respectively. The left y-axis corresponds to the smooth and the right to the data points.

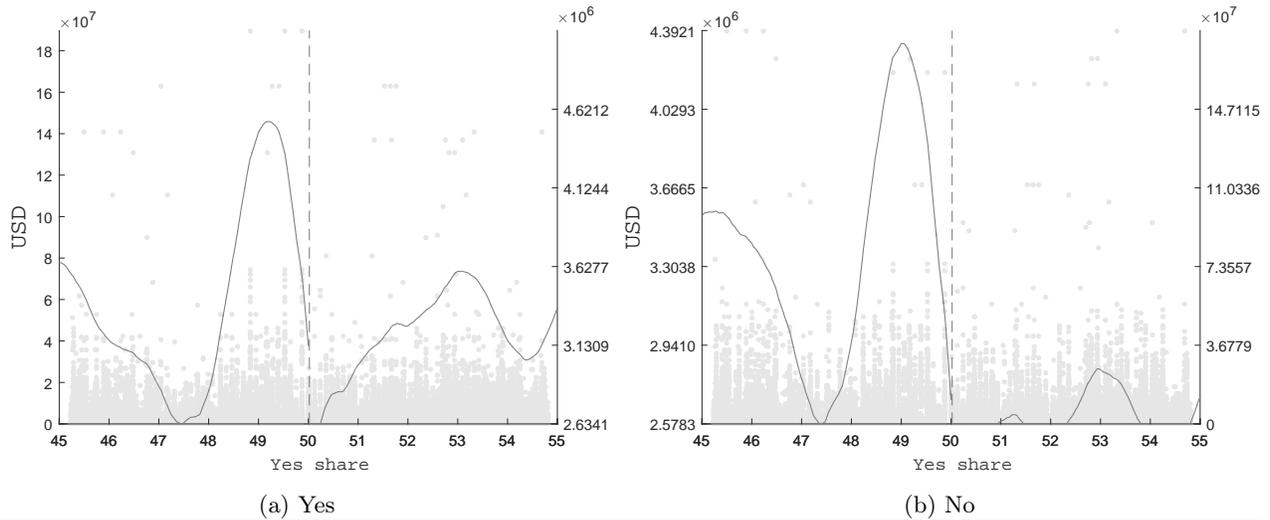


Figure A8: The solid line visualizes the local linear smooth of total daily donation flows within a time span of ± 4 weeks around the vote; the gray markers represent the raw data points using all votes in non-election years, and total daily donations to ‘yes’ or ‘no’ voters, respectively. The left y-axis corresponds to the smooth and the right to the data points.

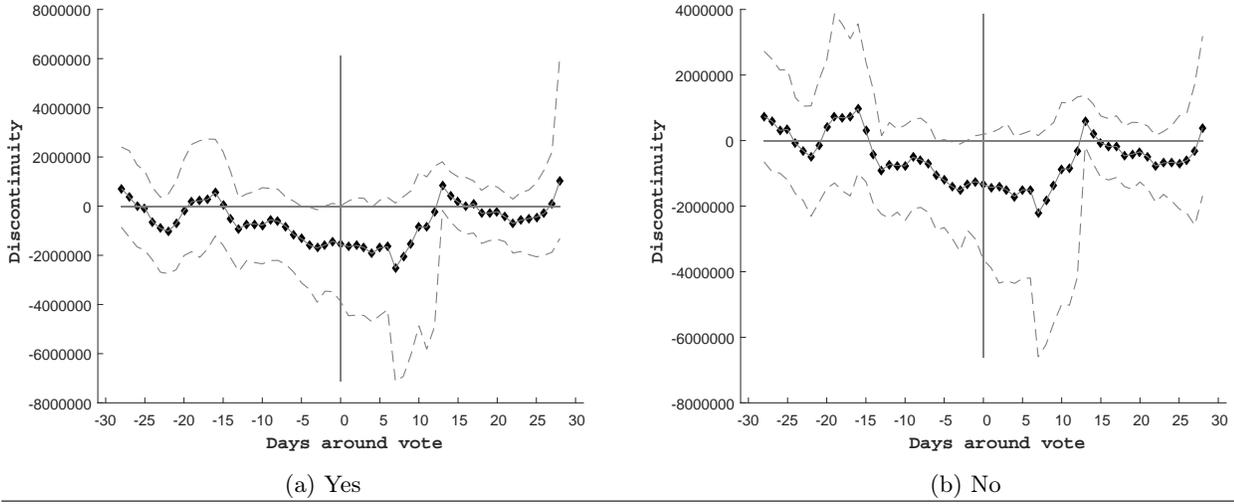


Figure A9: Rolling discontinuity estimates in total donations to ‘yes’ and ‘no’ voters within a 6-day band in non-election years. The dots represent the discontinuity estimates, and the dashed lines represent the 99% confidence bands.

Table A1: Discontinuity in total ‘low’ individual donations

	I	II	III	IV
Sample				
Donations	<i>Don<30</i>	<i>Don<40</i>	<i>Don<50</i>	<i>Don<100</i>
Years	<i>Elec.</i>	<i>Elec.</i>	<i>Elec.</i>	<i>Elec.</i>
$\Delta(Money)$	-0.06	0.36	0.31	1.63
<i>lower99%</i>	-0.82	-0.58	-0.54	-13.66
<i>upper99%</i>	0.44	1.13	1.14	12.70
<i>h</i>	1.00	1.00	1.00	1.00
<i>N</i>	322,107.00	322,107.00	322,107.00	322,107.00
<i>Nboot</i>	1,000.00	1,000.00	1,000.00	1,000.00

Notes: Local linear discontinuity estimates, including 99% bootstrap confidence intervals. *Money* stands for total daily individual donations within a range of 4 weeks around each vote, including only ‘low’ donation flows for below USD 30, 40, 50, and 100, respectively, in election years.

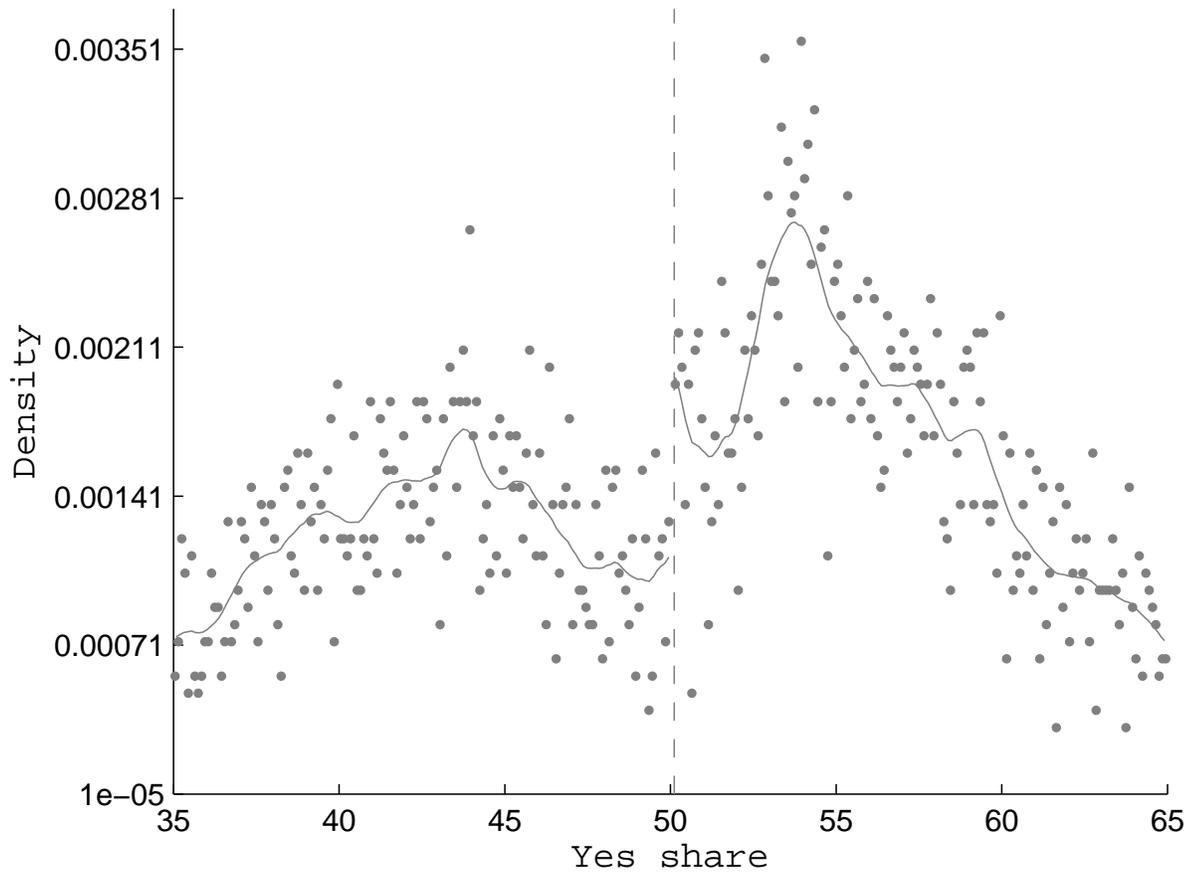


Figure A10: Vote distributions estimate using smoothed binned counts as in McCrary (2008). The smooth is a local linear smooth using a triangular kernel and the bandwidth of 1%. The bin size is 0.1%.

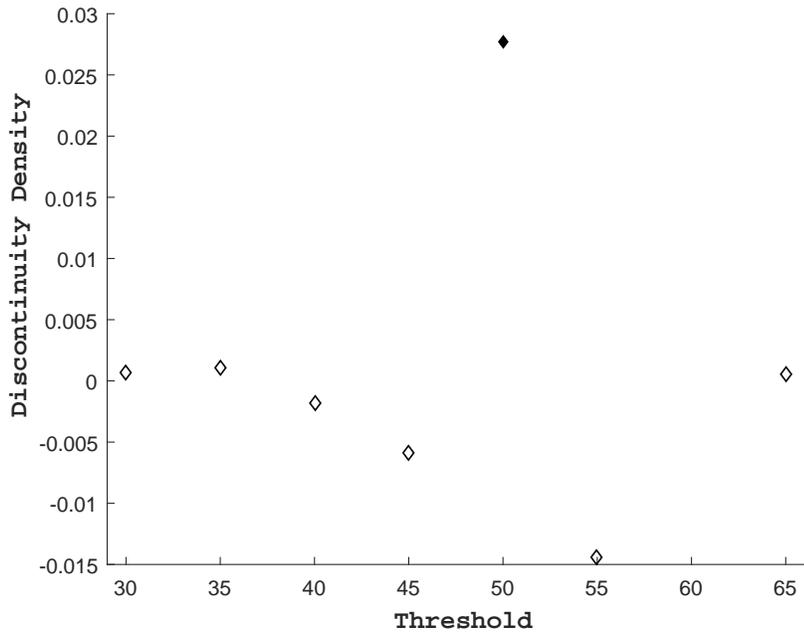


Figure A11: Discontinuity in density estimates for votes in election years at several thresholds. Estimates that are significantly different from zero at the 5 percent level are marked with filled points.

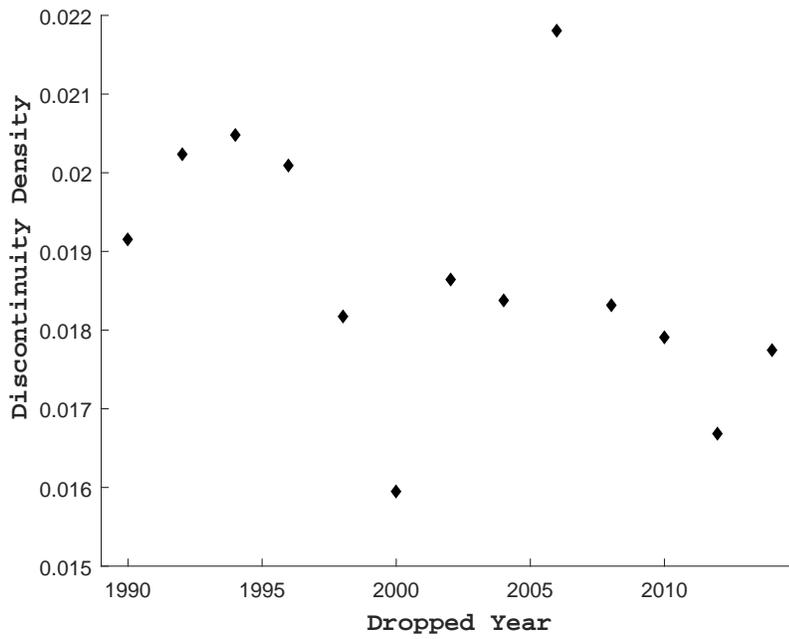


Figure A12: Discontinuity in density estimates for votes in election years dropping one year at a time. Estimates that are significantly different from zero at the 5 percent level are marked with filled points.

Table A2: Density discontinuity - Sensitivity Bandwidth Choice

	c	h	\hat{f}_l	\hat{f}_r	$\hat{\theta}$	\hat{l}_r	\hat{f}_r/\hat{f}_l	p-value	N
Election years									
I.	50.1149	0.5000	0.0110	0.0387	0.0277	4.7561	3.5085	0.0292	5,651
II.	50.1149	0.7500	0.0094	0.0328	0.0234	7.6494	3.4975	0.0057	5,651
III.	50.1149	1.0000	0.0089	0.0277	0.0187	8.5755	3.1053	0.0034	5,651
IV.	50.1149	1.2500	0.0094	0.0253	0.0159	8.5918	2.7007	0.0034	5,651
V.	50.1149	1.5000	0.0090	0.0235	0.0145	9.6469	2.6176	0.0019	5,651
VI.	50.1149	1.7500	0.0086	0.0225	0.0138	11.0875	2.6034	0.0009	5,651
VII.	50.1149	2.0000	0.0083	0.0219	0.0136	12.9333	2.6264	0.0003	5,651
VIII.	50.1149	2.2500	0.0083	0.0216	0.0133	14.4692	2.6021	0.0001	5,651
IX.	50.1149	2.5000	0.0083	0.0211	0.0127	15.2205	2.5232	0.0001	5,651
X.	50.1149	2.7500	0.0085	0.0205	0.0120	15.2951	2.4091	0.0001	5,651
XI.	50.1149	3.0000	0.0087	0.0199	0.0113	15.0817	2.2962	0.0001	5,651
XII.	50.1149	3.2500	0.0088	0.0193	0.0105	14.6498	2.1931	0.0001	5,651
XIII.	50.1149	3.5000	0.0088	0.0187	0.0098	14.3839	2.1143	0.0001	5,651
XIV.	50.1149	3.7500	0.0089	0.0182	0.0093	13.9875	2.0371	0.0002	5,651

Notes: Local likelihood ratio results for the discontinuity in the vote distribution for votes in election years and for diverse bandwidths. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.

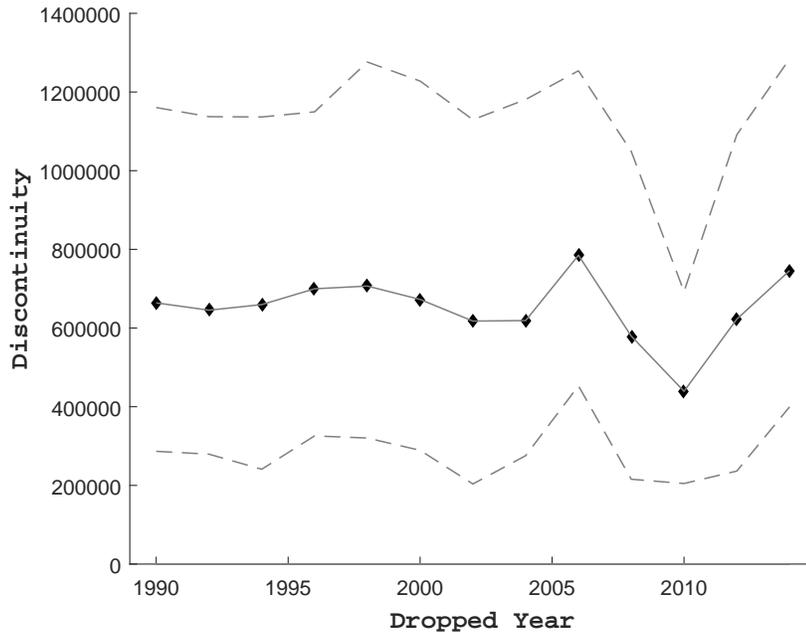


Figure A13: Discontinuity in total daily donations around votes in election years between the two days before the vote and up to three days after the vote, dropping one year at a time. The dots represent the discontinuity estimates, and the dashed lines the 99% confidence bands.

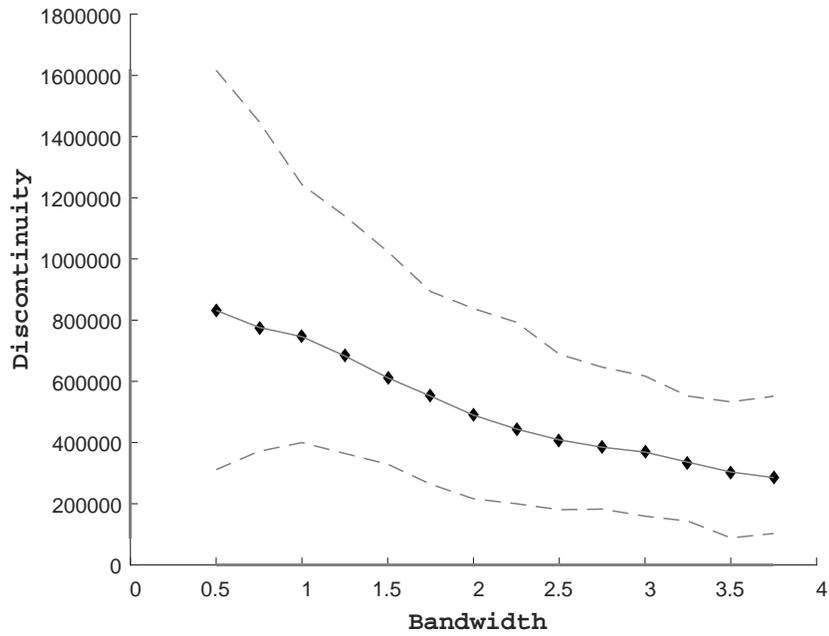


Figure A14: Discontinuity in total daily donations around votes in election years between the two days before the vote and up to three days after the vote, for diverse bandwidths. The dots represent the discontinuity estimates, and the dashed lines the 99% confidence bands.

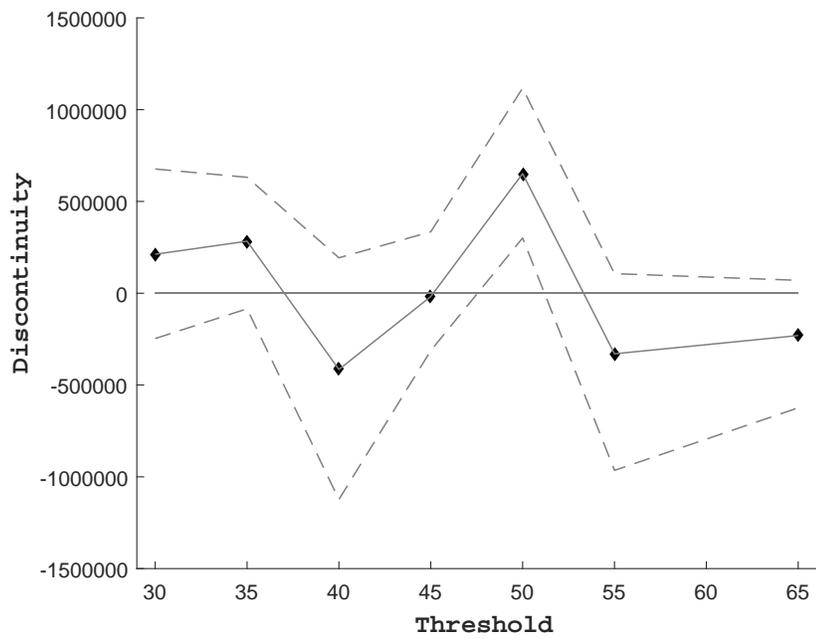


Figure A15: Discontinuity in total daily donations around votes in election years between the two days before the vote and up to three days after the vote, for several thresholds. The dots represent the discontinuity estimates, and the dashed lines the 99% confidence bands.

A.III Data appendix

Our data set on campaign finance contributions consists of all campaign finance donations to U.S. Representatives registered by the Federal Election Commission (FEC) between 1 January 1989 and 31 January 2015. The raw data are prepared by the Center for Responsive Politics (CRP) and provided by the Sunlight Foundation’s Influence Explorer database (<http://data.influenceexplorer.com/bulk>). Each row in the raw data set consists of one registered donation, mentioning the amount (in USD), the exact donation date, as well as additional information describing both the recipient and the contributor. Two additional variables are particularly relevant for our empirical analyses:

- *Contributor-type*: This categorical variable indicates whether a donation was made by an individual (a natural person) or an organization, i.e., a committee (political action committee, party committee, or candidate committee). Importantly, corporations, unions, trade organizations, and other legal entities are not allowed to (or have restricted possibilities to) directly contribute to a legislator’s campaign as per the Federal Election Campaign Act (FECA). They can, however, contribute to political action committees (PAC) which in turn can contribute to a candidate’s/legislator’s campaign. The contributor type variable allows us to filter the data for contributions that were made by individuals only or by organizations (PACs) that potentially represent special interests.
- *Contributor-category*: This variable describes which industry/sub-industry or type of organization a contributor belongs to, according to CRP’s categorization into one of 464 different categories (e.g., “Defense”, “Agriculture”, or “Party-Committees”). Together with the contributor-type variable, we can thus employ this variable in order to filter our data for donations made by organizations (PACs) that are either close to a political party, corporations, or unions.

The samples we use in our analyses of donation flows (Subsection 5.3) are constructed as follows. For all samples, we first filter out contributor-categories that either contain rare special cases of donations, or contain contributors that are not reasonable with respect to the hypotheses we want to test (i.e., self-contributions), or remainders that could not have been properly categorized. Table A3 lists the contributor-categories that are generally excluded.

Then, in order to generate the sub-samples for corporate/organization, individual, and party donations, we further filter the data according to the contributor-type and contributor-category. For the sub-sample of individual donations, we select all remaining donations made by individuals. For the party sub-sample, we select all donations made by committees belonging to the categories listed in Table A4. For the corporate/organization

sub-sample, we select all remaining donations made by committees (which covers all industry- and union-related organizations, such as PACs directly connected to a corporation or a particular industrial sector).

Table A3: Generally excluded campaign finance contribution categories

Code	Category	Description
Z7777	non-contribution	office use only
Z8888	non-contribution	catcode error
Z9000	candidate self-finance	candidate contribution to his/her own campaign
Z9010	candidate self-finance	Republican candidate contributions to own campaign
Z9020	candidate self-finance	Democratic candidate contributions to own campaign
Z9030	candidate self-finance	third-party candidate contributions to own campaign
Z9040	candidate self-finance	nonpartisan candidate contributions to own campaign
Z9100	party committee transfer	transfer between national party committees
Z9500	non-contribution	transfer from intermediary (type 24i or 24t)
Z9600	non-contribution	non-contribution, miscellaneous
Z9700	non-contribution	unitemized (small) contributions
Z9800	public subsidy	campaign funding from public sources
Z9999	non-contribution	internal transfer and other non-contributions

Notes: CRP campaign finance categories (and respective codes) of campaign finance transactions that are generally excluded from our analyses.

Table A4: Selected donation categories for party donations

Code	Category	Description
J2000	leadership pacs	leadership committees
J2100	leadership pacs	Democratic leadership pac
J2200	leadership pacs	Republican leadership pac
J2300	leadership pacs	Democratic officials, candidates & former members
J2400	leadership pacs	Republican officials, candidates & former members
J2500	leadership pacs	non-federal candidate committees
J2510	leadership pacs	pacs operated by Republican state politicians
J2520	leadership pacs	pacs operated by Democratic state politicians
Z1000	candidate committees	candidate committees
Z1100	candidate committees	Republican candidate committees
Z1200	candidate committees	Democratic candidate committees
Z1300	candidate committees	third-party candidate committees
Z1400	candidate committees	unknown-party candidate committees
Z2100	Republican unitemized contributions	conservative/Republican under reporting threshold
Z2200	Democrat unitemized contributions	liberal/Democrat under reporting threshold
Z2300	third-party unitemized contributions	third-party under reporting threshold
Z2400	nonpartisan unitemized contributions	nonpartisan contributions under reporting threshold
Z4100	joint candidate cmte	Republican joint candidate committee
Z4200	joint candidate cmte	Democratic joint candidate committee
Z4300	joint candidate cmte	third-party joint candidate committee
Z5000	party committees	party committees
Z5100	party committees	Republican party committees
Z5200	party committees	Democratic party committees
Z5300	party committees	third-party party committees

Notes: CRP campaign finance categories (and respective codes) of campaign finance transactions by legal entities which are considered as ‘party donations’ in our analyses.