Policy Convergence or Divergence? Evidence from U.S. House Roll Call Voting Records [*]

David S. Lee
UC Berkeley and NBER

Enrico Moretti
UCLA and NBER

Matthew J. Butler
UC Berkeley

July 2002

First Draft: March 2002

Abstract

Models of electoral competition typically predict that partisan politicians are implicitly compelled by voters to choose policies more moderate than their most-preferred positions. In a repeated election context, such policy convergence can be sustained by reputational mechanisms, but distinct possibility is complete policy *divergence* – whereby politicians are unable to credibly commit to any other policy than their most-preferred position. Using data on roll call voting records (1946-1994), this paper empirically evaluates *which* of these equilibria best describes the behavior of Representatives of the United States House. For our empirical tests, we exploit a quasi-experiment whereby party control over a seat is as good as randomly assigned among districts in which Representatives are involved in very close elections (as revealed by the observed vote share). Our regression discontinuity estimates of the impact of party control on subsequent voting records provide evidence inconsistent with both full and partial convergence, and most consistent with the full policy divergence equilibrium.

[*] This is a revised version of an earlier draft "Are Politicians Accountable to Voters? Evidence from U.S. House Roll Call Voting Records," Center for Labor Economics Working Paper #50. We thank David Card, John DiNardo, Hongbin Cai and Mel Hinich, for helpful discussions, and David Autor, Anne Case, and participants of workshops at UNC-Chapel Hill, UT-Austin, Chicago Economics and GSB, Princeton, and UCLA for comments and suggestions. We also thank Jim Snyder and Michael Ting for providing data for an earlier draft.

1 Introduction

A key prediction of traditional models of parties' political behavior is that electoral competition for political office compels opposing politicians and parties to moderate their policy positions.[1] The special case of the "median voter theorem" – whose main implication is that opposing parties adopt the *same* position – is a staple of virtually all economics and political science textbooks. More generally, models of electoral competition typically predict that politicians' policy positions converge toward each other, driven by the attempt to capture the vote of citizens in the "middle".[2]

However, as pointed out in Alesina (1988), a dynamic consistency problem arises with this notion when political parties care not only about the outcome of an election, but also what policy is implemented. That is, without binding pre-commitments, once elected, the winning party has every incentive to deviate from the announced platform, and implement its most preferred policy. Rational and forward-looking voters should expect this, so that the end result is that the parties' only credible, time-consistent promise is the extreme policy. The result is complete policy *divergence*. This possibility has been considered in more recent attempts to model representative democracy, for example, in Besley and Coate (1997, 1998).

Alesina (1988) suggests, however, that in a fully dynamic context, where opposing parties compete in repeated elections, reputational mechanisms may overcome the pre-commitment problem. This may lead to time-consistent equilibria that exhibit full policy convergence. Even when the complete convergence outcomes are not sustainable, as long as discount factors are nonzero, intermediate "partial convergence" equilibria – where parties' moderate their positions, but not completely – are sustainable. Of course, the complete divergent equilibrium remains sustainable in this dynamic context.

Given the multiplicity of sustainable equilibria, an important question is: Which of them is most empirically relevant? Two important problems stand in the way of any empirical analysis of this issue. First, at best only the *winning* politician's positions are typically measurable by the researcher. In fact, measuring

[1] Models of this type have its origins in Hotelling (1929) and Downs (1957). Other developments along these lines include Hinich, Ledyard, and Ordeshook (1972, 1973), McKelvey (1975), Wittman (1983), and Calvert (1985)

[2] See Osborne (1995) for a nice review of variations of spatial competition under plurality rule. Also see Persson and Tabellini (2000).

the difference between an implemented policy and the policy *that would have been implemented had the losing candidate actually prevailed* is impossible for any given time and place. Second, parties' most-preferred policy outcomes are also difficult to quantify and measure. This makes it difficult to empirically distinguish between the complete policy divergence (where parties choose their most-preferred policy – hereafter referred to as "bliss points") and partial policy convergence (where parties moderate their positions and deviate somewhat from those bliss points).

By directly addressing these problems, this paper attempts to assess *which* of the three types of equilibria is most empirically relevant for one of the most important legislative bodies of government in the United States – the House of Representatives. Our analysis takes advantage of a quasi-experiment that is embedded in the Congressional electoral system. That is, we argue that among elections decided by a very narrow margin (say, a fraction of a percent of the vote), which of the two parties won is virtually randomly assigned. This as-good-as random assignment provides a means for two tests.

First, average voting records of Republicans who are *barely* elected can credibly represent, on average, how Republicans *would have* voted in the districts that were in actuality, barely won by Democrats (and vice versa). Thus, the difference between barely-elected Democrats' records and barely-elected Republicans' records – our basic regression discontinuity (RD) esimator – represents the average magnitude of policy divergence for all of these districts, or the true "causal party effect". Complete policy convergence implies this difference should be zero.

Second, the key implication of full policy divergence is that *past* electoral outcomes (e.g. whether the Democrat wins) should not have its own direct impact on candidates' positions, because those positions are equal to the parties' pre-determined "bliss points". The only "reduced-form" causal effect that past electoral outcomes should have on today's legislators' policy positions is through its impact on *which party* is elected today. Thus, if 1) the equilibrium is complete policy divergence, 2) the indicator of past Democratic election victory is as good as randomly assigned (among close elections) and 3) it has a causal effect on whether the Democrat holds the seat today, then its use in instrumental variables estimation should yield the same "causal party effect" estimate as our regression discontinuity estimator. Since we provide empirical evidence that 2) and 3) are satisfied, we interpret a significant difference between the IV and regression discontinuity estimates as a rejection of the complete policy divergence hypothesis. As shown below, with a slight modification, this implication holds whether or not we assume homogeneous or heterogenous causal party effects across districts.[3]

Our analysis of roll call voting record data for the U.S. House of Representatives (1946-1995) yields the following findings. First, the data are strongly inconsistent with the complete policy convergence equilibrium. Republicans and Democrats in "otherwise equal" Congressional districts exhibit voting records that are extremely divergent. In fact, the barely elected Democrats (Republicans) have voting records that are about as extreme as those of Democrats (Republicans) who won by large electoral margins. We show that districts barely won by Democrats are similar to those barely won by Republicans along many predeterimined characteristics. This lends credibility to our main identifying assumption that party control of the seat is as-good-as randomly assigned among close elections.

Second, we find that, after conditioning on close electoral races in the previous period, (which generates as-good-as random assignment of past Democratic victory) our IV estimate is about the same magnitude as our main RD estimate. This implies that Republican and Democratic candidates' positions are *not* directly impacted by past electoral outcomes, as predicted by the full divergence equilibrium. We conclude that the data are more consistent with the complete policy divergence equilibrium, whereby parties' positions are at their respective "bliss" points, and where voters, as a whole, are unable to compel Democrats and Republicans to moderate their extreme positions.

The paper is organized as follows. Section 2 reviews the basic results of a model of policy convergence and divergence in the context of repeated two-party elections for office. Section 3 describes inference problems encountered, and our test for complete policy convergence, and how we propose to differentiate between complete divergence and partial policy convergence. We also discuss why the U.S. House is an appealing context to study. Section 4 contains the main results of the paper, where we begin by describing

[3] As shown below, the robustness of this test to general forms of heterogeneity is unique to this election context. That is because we can, to a first-order approximation, identify which districts are affected by the instrument because we directly observe the latent index that determines electoral victory – the vote share.

our measures of "liberalness"/"conservativeness" of roll call voting records. We then turn to our main test of full convergence using ADA scores, and also present evidence that lend support to the notion that among close elections, there is near random assignment of party control of the seat. Having rejected full convergence, we conduct our tests of complete policy divergence, first adopting the assumption of homogeneity of divergence across districts, and then under a more general heterogeneous environment. Finally, we show that our results do not change when we use alternative measures of roll call voting patterns. Section 6 concludes.

2 Background: Theory and Context

The goal of this paper is to assess the empirical relevance of the notion that politicians are compelled by voters to adopt more moderate positions than their (or their party's) most-preferred position. More specifically, we examine the "policy convergence" hypothesis within the context of one of the most important legislative bodies of government in the U.S. – the House of Representatives.

2.1 Theoretical Framework

In this section, we describe the theoretical framework that we will adopt for our empirical tests for complete convergence and full policy divergence. There are many ways in which the behavior of partisan politicians can be modelled (see Chapter 5 of Persson and Tabellini, 2000). However, we believe that the framework of Alesina (1988) is most directly motivated by the issue of sustaining policy convergence. Furthermore, we show that its level of parsimony makes it empirically tractable. Since we adopt that framework, we begin by briefly outlining its key features and the results that are most relevant to our analysis. Details of the model and justifications for its assumptions are found in Alesina (1988).[4]

For a given Congressional district, there are two political parties, party 1 and party 2. No distinction is made between the party and its nominee.[5] The party's preferences are defined over a single-dimensional

[4] For convenience to the interested reader, we also adopt identical notation to Alesina (1988).

[5] Alesina and Spear (1988) develop a model in which politicians are considered finite-lived, while parties are considered infinite-lived. They show that in an overlapping-generations model, partially convergent equilibria are dynamically sustainable.

policy space (e.g. characterizing how "liberal" the policy is), expressed as

$$\begin{aligned} U(l) &= -\sum_{t=0}^{\infty} \frac{1}{2} q^t (l_t - c_1)^2 \\ V(l) &= -\sum_{t=0}^{\infty} \frac{1}{2} q^t (l_t - c_2)^2 \end{aligned} \tag{1}$$

for party 1 and 2, respectively, with $c_1 > c_2$ and $0 < q < 1$. l_t is the chosen policy of the officeholder following election t, and q is the per-election-cycle discount factor. These quadratic-loss functions imply that party 1's and 2's most preferred policies are c_1, and c_2, respectively. They are party 1's and 2's "bliss" points.

Electoral outcomes themselves are *not* deterministic, and are characterized by a probability function

$$P_t = P(x_t^e, y_t^e, \delta_t) \tag{2}$$

which denotes the probability that party 1 will win the district in election t. This function can be interpreted as capturing voters' own preferences regarding policy, and other characteristics of the parties.[6] x_t^e and y_t^e are the voters' (rational) expectations of the policy that party 1's and 2's candidate, respectively, will adopt if elected. δ_t represents a non-policy determinant of this "vote production function", and parameterizes the popularity of a party's candidate, keeping expected positions constant.[7] Voters are forward-looking and have rational expectations; that is, in equilibrium, their expectations of legislators' actions are correct. Thus, in the discussion below, in equilibrium party 1 chooses the policy $x_t = x_t^e$ and party 2 chooses $y_t = y_t^e$.

The most important assumption regarding the function P is that a candidate's probability of winning the election rises, *ceteris paribus*, as her anticipated future policy choice moves "closer" to that of her opponent.[8] This captures the notion of the electoral (probabilistic) benefit resulting from moving to the "middle" in order to capture more of the vote.

An uncertain electoral outcome implies that the implemented policy is also uncertain. The welfare

---

[6] See Alesina (1988) for a more detailed discussion of how the function can be derived from voters' preferences.

[7] Alesina (1988) does not include δ_t, but refers to such a factor in the text, considering how an exogenous increase in the popularity of a particular party will alter the Nash bargaining solution. We formally introduce it here, for the exposition in a later section.

[8] This is Assumption (iv) of Alesina (1988).

of the party is assumed to be the expected utility

$$
\begin{aligned}
w^1 &= P(x_t, y_t, \delta_t) U(x_t) + (1 - P(x_t, y_t, \delta_t)) U(y_t) \\
w^2 &= P(x_t, y_t, \delta_t) V(x_t) + (1 - P(x_t, y_t, \delta_t)) V(y_t)
\end{aligned} \tag{3}
$$

for party 1 and 2, respectively.

The timing of elections is as follows. Before election t, candidates from each party announce how they will act (how they will vote on roll call votes), if elected. Voters' form expectations of how each candidate would act. The election is held, and the winning party's candidate chooses a position (x_t for party 1, and y_t for party 2). The voters' rational expectations of the candidates' choices turn out to be correct. The electoral cycle then repeats.

The key results from Alesina (1988) are:

1. The efficient outcome is one where $x_t = y_t$ – full convergence.[9] Because of the concavity of the preference functions, both parties would like to choose a moderate outcome with certainty than take a risky bet, even if the expected value of the outcome is the same.
2. In a one-period game, without the possibility of binding pre-commitments to policies, candidates' announcements are meaningless, and candidates, if elected, choose their bliss points, as expected by voters. The basic problem is that once elected, candidates have every incentive to move to the most-preferred policy, and face no recourse from deviating from an announced policy. Rational voters expect that and vote accordingly. The inability of candidates to make binding pre-commitments leads to the complete divergence result. Thus, policy convergence results require some way to overcome this inability to pre-commit; otherwise the equilibria will be time-inconsistent. This equilibrium is inefficient.
3. In a repeated game context, fully convergent equilibria can be sustained as long as the discount factor is sufficiently high. The proposed equilibrium is one in which parties agree to announcing, and carrying out a moderate outcome if elected. The expectation is that if the legislator deviates from the announced policy, the party (having lost its reputation) reverts to the bliss point forever, and the opposing party would also revert to its bliss point. As long as the discounted threat of "punishment" outweighs the short-term gains from "cheating", the full convergence sub-game perfect equilibrium can be sustained.
4. Even when the fully-convergent equilibria cannot be sustained, as long as the discount factor is not zero, *partially* convergent equilibria are still sustainable under the same kind of reputational mechanism. That is, the electoral benefit to capturing "middle voters" leads to both parties to moderate their positions, which is Pareto superior (from the perspective of the parties) to the fully divergent one-shot Nash outcome.

Thus, in an infinitely repeated game context, there are multiple equilibria. Fully-convergent, "median-voter-type" equilibria are sustainable, as are partially convergent equilibria. In addition, the one-

[9] By "efficient", Alesina (1988) refers exclusively to the welfare of the political parties, not the voters. We adopt his notion here. See Besley and Coate (1997, 1998) for a detailed discussion of the notion of "efficiency" in models of representative democracy.

shot Nash, full policy divergence result also remains an equilibrium of the dynamic game.

Remark 1. There is a multiplicity of Pareto-improving (relative to the one-shot Nash) points on the fully-convergent efficient frontier. Alesina (1988) uses Nash bargaining to choose one point, and proves an intuitive result: the Nash-bargaining solution moves towards party 1's bliss point with an exogenous change in party 1's popularity. That is, *ceteris paribus*, the equilibrium $x_t = y_t$ moves towards c_1 in response to an increase in P_t. The intuition is that party 1's bargaining position is strengthened by an exogenous increase in its popularity.

Remark 2. An analogous implication of the bargaining solution for partially convergent equilibria has not been established for the case when fully-convergent outcome cannot be credibly sustained, although Alesina (1988) proves existence of a Pareto-improving choices, when $q > 0$. The characterization of the efficient frontier within this smaller set of choices (when the fully convergent outcome is unsustainable) is not as straightforward as in the fully convergent case.

Remark 3. If c_1 and c_2 are unquantifiable/unobservable to the researcher *and* if there is no clear theoretical prediction about how an exogenous increase in P_t would impact the partially convergent equilibrium policy choices, then there is no empirical content to the partially convergent equilibrium. The partially convergent and fully divergent equilibria would be empirically indistinguishable.

In this paper, we are interested in empirically differentiating between the full divergence and partial convergence hypothesis. Since we are presuming that c_1 and c_2 are unobservable, such an attempt only makes sense if $\frac{\frac{\partial x_t}{\partial \delta_t}}{\frac{\partial \widehat{P}_t}{\partial \delta_t}} \neq 0$, where the $\widehat{P}_t$ denotes the equilibrium probability.

2.2 Context: the U.S. House

Given the wide range of possible equilibria, it would informative to obtain evidence on *which* of the three types of equilibria is most empirically relevant for describing the policy formation mechanism in a major, long-standing representative democracy, such as the United States.

In particular, the U.S. House of Representatives is an ideal context to explore the empirical content of these ideas for several reasons. First, the U.S. federal legislative body is virtually a two-party system;

when there are more than two candidates, the basic insight of the Hotelling (1929) and Downs (1957) approach to policy convergence is somewhat weakened (see Osborne 1995). Thus, the concept of policy-convergence is more well-defined in the context of a two-party system like that which exists in the U.S.

Second, elections to the U.S. House are of the plurality/winner-take-all type. This allows us to more plausibly think of the legislator in a Congressional district as a potential representative of voters in that district. Competition between opposing candidates is more plausibly thought of as happening at the district level, and hence there are many units available for a micro-econometric analysis. This should be contrasted to proportional representation systems, whereby seats are allocated in proportion to the national vote.

Third, there are reasons why reputational mechanisms are more likely to be relevant for elections to the U.S. House, compared to elections to other political offices. Elections are held every two years, and there are no term limits (as opposed to gubernatorial and presidential elections), meaning that the potential of political careers of many terms is possible. Furthermore, political tenure in the House is often a stepping-stone to participating in electoral races for higher offices. This may make politicians more mindful about the consequences of deviating from promised positions.

For these reasons, we believe that electoral competition for seats to the U.S. House is one of the most likely settings for finding empirical evidence of the ability of reputational mechanisms to sustain full or partial policy convergence.

3 Empirical Problems and Implications

In this section, we describe the two empirical problems which stand in the way of empirically distinguishing between fully convergent, partially convergent, and fully divergent equilibria. We also describe how we use a regression discontinuity design in this context to address these problems.

3.1 Two Empirical Problems

The first important problem is that although the announced and expected positions of both opposing candidates are known to voters in each district, as researchers, we can only systematically measure the actions of the legislator. We are unable to measure on the same "scale", the expected positions of the candidate who eventually lost the electoral race. More specifically, in our analysis, we focus on the roll call voting behavior of the legislator, which we can observe and quantify. But we do not observe and cannot quantify what the losing candidate's roll call voting behavior *would have been*, had he won the election instead.

More formally, adding the subscript i to denote the Congressional district, we only observe

$$RC_{it} = \begin{cases} x_{it} \text{ if party 1 wins election } t \\ y_{it} \text{ if party 2 wins election } t \end{cases} \tag{4}$$

which can be equivalently written as

$$RC_{it} = y_{it} + DEM_{it} \left(x_{it} - y_{it} \right) \tag{5}$$

where RC_{it} is a measure of district i's legislator's roll call voting behavior – for example, how liberal the voting record is – in the Congressional session that follows election t. $DEM_{it} = 1$ if the party 1 (e.g. Democrats) wins election t in district i, and 0 if party 2 (e.g. Republicans) prevails. As researchers, we cannot measure both positions simultaneously, so it is impossible to know, for a particular district, if x_{it} equals (e.g. full convergence) or substantially deviates (e.g. full divergence) from y_{it}.

The second problem is that it is difficult to obtain credible measures of the bliss points of parties in any given district (denote the district-time-specific bliss points for party 1 and 2 as c_{1it} and c_{2it}, respectively).[10] This makes it impossible to assess whether or not $x_{it} = c_{1it}$ ($y_{it} = c_{2it}$), which is, by definition, what differentiates a partially convergent equilibrium from the fully divergent case.

3.2 Identification Strategy: Regression Discontinuity Design

3.2.1 A Test of Full Policy Convergence

A regression discontinuity (RD) design inherent in the electoral system permits a direct test of the full convergence hypothesis. In particular, we argue that districts in which candidates for party 1 (e.g. Democrats) are *barely* elected (say, by a tiny fraction of the vote) are in all ways *ex ante* similar to districts in which candidates for party 2 (e.g. Republicans) are *barely* elected. In particular, if the regression discontinuity

[10] For example, leaders of Democrats in Alabama may have ideal positions quite different from the Democratic leadership in Massachusetts.

design is valid, the two groups of districts would be similar, along all pre-determined characteristics, including the voters' preferences, and the parties' district-specific bliss points. This virtual "random assignment" of which party wins a close district implies that the average voting records of Republicans that are barely elected can credibly represent, on average, how Republicans *would have* voted in the districts that were in reality, barely won by Democrats (and vice versa).

To see how the regression discontinuity design addresses the inference problem, first note that we can express $P_{it} = P\left(x_{it}^e, y_{it}^e, \delta_{it}\right)$ in terms of the vote share for party 1 (Democrats):

$$VS_{it} = vs\left(x_{it}^e, y_{it}^e, \delta_{it}, \varepsilon_{it}\right) \tag{6}$$

where ε_{it} is an unpredictable and unforecastable component of the vote share, is independent of all other factors. This could be interpreted as turn-out on voting day, or errors in polls, etc.[11] vs is a continuously differentiable function, and the framework presented in the previous section implies that the partial derivatives have the following sings: $vs_1 < 0$, $vs_2 < 0$; we normalize $vs_3 > 0$ and $vs_4 > 0$. In a two-party system, $DEM_{it} = 1$ if and only if $VS_{it} > \frac{1}{2}$.

In a fully convergent equilibrium, the positions of candidates in district i for election t are completely determined by the bliss points and the voting production function, so that $x_{it} = m^f\left(c_{1it}, c_{2it}, \delta_{it}\right) = y_{it}$. We assume throughout that m^f is continuously differentiable with respect to its arguments.

The simple difference between the voting records of Democratic and Republican legislators is uninformative about the full policy convergence hypothesis. This is because

$$\begin{aligned} & E\left[RC_{it}|DEM_{it}=1\right] - E\left[RC_{it}|DEM_{it}=0\right] \qquad (7) \\ = \; & E\left[m^f\left(c_{1it}, c_{2it}, \delta_{it}\right)|VS_{it} > \frac{1}{2}\right] - E\left[m^f\left(c_{1it}, c_{2it}, \delta_{it}\right)|VS_{it} < \frac{1}{2}\right]. \end{aligned}$$

There is potential for serious selection bias, as c_{1it}, c_{2it}, and δ_{it} all help determine VS_{it}, and hence the outcome of the election. Intuitively, the districts that were won by Democrats are likely to be systematically different from those won by Republicans. In particular, it is plausible (and likely) that voters in districts won by Democrats are, on *average*, more liberal than voters in districts won by Republicans; it is also

[11] The existence of this component is equivalent to Alesina's (1988) maintained assumption that electoral outcomes are uncertain.

plausible that *both* parties' bliss points are more liberal in relatively more liberal districts. The source of the problem is that the distribution of c_{1it}, c_{2it}, and δ_{it} within Democratic-won districts is quite likely to be very different from that within Republican-won districts.

Under a mild continuity assumption, if the attention is restricted to elections where the vote share margin of victory is slim, the Democrat and Republican districts will become arbitrarily similar in the distribution of these quantities.

Proposition 1 *If c_{1it}, c_{2it}, δ_{it}, and ε_{it} has continuous joint density, then the density of c_{1it}, c_{2it}, and δ_{it} conditional on $VS_{it} = \frac{1}{2} + \Delta$ equals the density conditional on $VS_{it} = \frac{1}{2} - \Delta$ in the limit, as $\Delta \to 0$.*

This is an important result for the empirical analysis in the paper, which focuses on the comparison of barely-elected Democrat and Republican districts. Essentially, it implies that when examining "close" elections, there is virtual "random assignment" of which party ultimately wins. In the closest of elections (e.g. decided by 1 vote), which party wins is determined as if by the "flip of a coin". This will result in the bare-Democrat and bare-Republican districts being on average similar in all the characteristics that determine the vote share.

It follows that

$$\begin{aligned} & E\left[RC_{it}|VS_{it} = \frac{1}{2} + \Delta\right] - E\left[RC_{it}|VS_{it} = \frac{1}{2} - \Delta\right] \\ \approx\ & E\left[x_{it} - y_{it}|VS_{it} = \frac{1}{2}\right] \end{aligned} \tag{8}$$

for Δ sufficiently small.

So the regression discontinuity (RD) estimand – the comparison of voting patterns between barely-elected Democrats and barely-elected Republicans – should equal the average difference in policies between opposing candidates in those districts. If parties are able to sustain full policy convergence, this quantity should be zero. A significant departure from zero indicates a rejection of the full convergence hypothesis.

3.2.2 Differentiating between Complete Divergence and Partial Convergence

The bliss points c_{1it} and c_{2it} are not easily measured by the researcher, which makes a simplistic test of full policy divergence infeasible. However, the theoretical framework described in Section 2 generates a strong prediction for the full policy divergence hypothesis. Specifically, in the fully divergent equilibria (where $x_{it} = c_{1it}$ and $y_{it} = c_{2it}$) an exogenous change in the probability of a Democrat (Republican) victory should *not* cause a change in the parties' positions, because those policy positions are completely determined by the exogenously determined "bliss points". An exogenous change in the relative popularity of a party in any given district should only have the effect of altering the relative odds of whether party 1's or 2's bliss point is ultimately chosen.

By contrast, this stark prediction does not hold for the partial convergence equilibrium. It is *possible* that the equilibrium would remain the same. But, in general, we would expect it to move since the constraints of sub-game perfection (that the gain from cheating not exceed the punishment) depend on expected utilities, which in turn, depend on the probability P_{it}.[12] As noted earlier, given that researchers cannot observe c_{1it} and c_{2it}, the partial convergence concept has no empirical content unless the partial convergence equilibrium systematically moves in response to an exogenous change in the

The sign of $\frac{\frac{\partial x_t}{\partial \delta_t}}{\frac{\partial \widehat{P}_t}{\partial \delta_t}}$ will likely depend on the particular bargaining solution concept used, as well as the functional form of the function $P(\cdot,\cdot,\cdot)$. However, we note that if the fully convergent equilibrium were to have the same *qualitative* features as the fully divergent equilibrium, we would expect both $\frac{\frac{\partial x_t}{\partial \delta_t}}{\frac{\partial \widehat{P}_t}{\partial \delta_t}}$ and $\frac{\frac{\partial y_t}{\partial \delta_t}}{\frac{\partial \widehat{P}_t}{\partial \delta_t}}$ to be positive.[13] That is, if there were an exogenous increase in the popularity of a Democrat, we might expect the relative bargaining power of the Democrats to rise, causing both candidates to shift their positions towards the Democratic "bliss point".

In our analysis, we use the notion that party incumbency *causes* an exogenous increase in the probability of winning the subsequent election for testing the full divergence hypothesis. Lee (2001, 2002) argues that the regression discontinuity estimate

$$E\left[DEM_{it}|VS_{it-1} = \frac{1}{2} + \Delta\right] - E\left[DEM_{it}|VS_{it-1} = \frac{1}{2} + \Delta\right] \tag{9}$$

of the incumbency advantage is a valid estimate of the incumbency effect. Lee (2001, 2002) finds that, winning an election causes an increase in the probability that the party will win the next election as much

[12] See, for example, equations 15 and 16 in Alesina (1988).
[13] This would be analogous to Proposition 3 (ii) in Alesina (1988).

as 0.45.

Thus, our test of full divergence amounts to assessing whether the party winning an election – by causing it have a greater probability of winning the next election – causes it to change its policy position for the *next* election, *all else equal*. The regression discontinuity design is helpful here because it arguably generates as-good-as randomized variation in whether or not a party wins an election, and hence keeps *all else equal*.

Proposition 2 *If DEM_{it-1}has an average causal effect on DEM_{it} (there exists a true electoral advantage to party incumbency) then DEM_{it-1} has an impact on δ_{it}.*

This follows immediately from the fact that DEM_{it} is a known, deterministic function ($DEM_{it} = 1$ if and only if $VS_{it} > \frac{1}{2}$) of VS_{it} , so anything that causally effects DEM_{it} must do so by impacting VS_{it}. With the theoretical framework of Section 2, the equilibrium values of VS_{it} are completely determined by c_{1it}, c_{2it}, δ_{it}, and ε_{it}. In the framework, the bliss points are exogenously determined, and ε_{it} is assumed to be the unpredictable component that generates the uncertain electoral outcome; hence DEM_{it-1} must induce an impact on DEM_{it} through δ_{it}.

While there are many interpretations of what δ_{it} could represent, one concrete example is that it represents the voters' (independent from any partisan preferences) valuation of experience in Congress. With this interpretation, an incumbent party has a higher probability of winning the seat again because the expected experience level of its candidate will be higher than the expected challenger (Lee 2001).

Test under Homogeneity Before turning to a more general model, we illustrate the basic intuition of the test by starting with the assumption that the *difference* between opposing parties' bliss points is *constant* across districts. In other words, even though c_{1it}, c_{2it} varies across districts, $c_{1it} - c_{2it} = k_0$.

Proposition 3 *Under homogeneity, if 1) whether or not a Democrat held the seat in election $t-1$ (DEM_{it-1}) is as-good-as randomly assigned, and 2) DEM_{it-1} has a nonzero impact on DEM_{it}, and 3) the complete divergence hypothesis is true, then DEM_{it-1} constitutes a valid instrument for estimating the impact of DEM_{it} on RC_{it}. This IV estimand should equal k_0, which is also consistently estimated by the RD estimate in Equation 8.*

To see this, note that under the null hypothesis that candidates always choose their parties' bliss points (complete divergence), Equation 5 can be re-written as

$$RC_{it} = c_{2it} + DEM_{it}k_0 \tag{10}$$

As mentioned earlier, the bliss point c_{2it} is determined outside the model, so DEM_{it-1} has no impact on c_{2it}. Therefore, if DEM_{it-1} is as good as randomly assigned, and has an effect on DEM_{it}, then DEM_{it-1} would be a valid instrument for estimating k_0, the causal impact of DEM_{it} on RC_{it}. Under homogeneity, our RD estimator from 8 is also consistent for k_0.

The basic intuition of the test is that under full divergence, positions are determined by and identical to the bliss points, which are considered fixed and unaffected by electoral outcomes. Therefore DEM_{it-1} should only affect period t's policy outcome through its impact on which party is elected in period t, DEM_{it}. Thus, the IV estimate is consistent for k_0, the degree of policy divergence, and should match the RD estimate of k_0.

We have already argued that the first condition of the above proposition holds (see Proposition 1) if we restrict our attention to close elections in period $t-1$, and Lee (2001, 2002) provides strong evidence that the second condition holds. Using Proposition 1, it is easy to show that under complete divergence

$$\frac{E\left[RC_{it}|VS_{it-1}=\frac{1}{2}+\Delta\right]-E\left[RC_{it}|VS_{it-1}=\frac{1}{2}+\Delta\right]}{E\left[DEM_{it}|VS_{it-1}=\frac{1}{2}+\Delta\right]-E\left[DEM_{it}|VS_{it-1}=\frac{1}{2}+\Delta\right]}=k_0 \tag{11}$$

This can be recognized as a "local" Wald estimator, or a the estimate from a "fuzzy" regression discontinuity design. We will refer to it below as the RD-IV estimate.[14] Thus, under complete divergence (and homogeneity) this IV-type estimand should equal the regression discontinuity (RD) estimand in Equation 8.

In our empirical analysis, we show that condition 1) and 2) in the above proposition is strongly supported by the data. Therefore, a substantial difference between the RD and RD-IV estimates constitutes a rejection of the complete divergence hypothesis. Furthermore, we interpret such a rejection as evidence more in favor of the alternative hypothesis – that is, the partial convergence equilibrium. In fact, under

[14] For a recent formalization of the use of the regression discontinuity design to estimate causal effects, see Hahn, Todd, and van der Klaauw (2001).

partial convergence, we would expect the RD-IV estimate to exceed the RD estimate.[15]

Test under Heterogeneity Even under a more general model where $c_{1it} - c_{2it}$ varies across districts, the basic intuition of the test remains, with a slight modification. In the discussion that follows, assume that we have conditioned on the districts involved in close elections in $t-1$, $\frac{1}{2} - \Delta < VS_{it-1} < \frac{1}{2} + \Delta$, with Δ small. We denote E_Δ as the expectation conditional on these close elections in $t-1$. Note that among this group of districts, the average causal effect of DEM_{it-1} on RC_{it} (estimated by the numerator in Equation 11) is a weighted average of the causal effects for three sub-populations[16]:

$$E_\Delta\left[x_{it}|DEM_{it-1} = 1, STRONG_t^{DEM}\right] - E_\Delta\left[x_{it}|DEM_{it-1} = 0, STRONG_t^{DEM}\right] \tag{12}$$

$$E_\Delta\left[y_{it}|DEM_{it-1} = 1, STRONG_t^{REP}\right] - E_\Delta\left[y_{it}|DEM_{it-1} = 0, STRONG_t^{REP}\right] \tag{13}$$

$$E_\Delta\left[x_{it}|DEM_{it-1} = 1, SWING_t\right] - E_\Delta\left[y_{it}|DEM_{it-1} = 0, SWING_t\right] \tag{14}$$

The first expression represents the average effect of DEM_{it-1} on the Democrats' positions for the sub-population of Democrats ($STRONG_t^{DEM}$) who would have won the election in period t irrespective of DEM_{it-1}. The second expression is the analogous effect for the sub-population of Republicans ($STRONG_t^{REP}$) who would have won election t irrespective of DEM_{it-1}. The final expression is the effect – or degree of policy divergence – among the sub-population of districts ($SWING_t$) that switched from Democratic to Republican control *because* of the incumbency advantage enjoyed by the Democrats in period $t-1$ (DEM_{t-1}). In the terminology of Angrist, Imbens, and Rubin (1996), the expressions represent the causal effects for the "always-takers", "never-takers" and "compliers", respectively.

There are two main implications of the complete divergence hypothesis – where positions are equal to pre-determined bliss points – when allowing for more general heterogeneity:

1. the first two effects should be zero; this is testable insofar the two effects can be estimated with data.
2. the third effect should be positive and equal to $E_\Delta\left[c_{1it} - c_{2it}|SWING_t\right]$. Strictly speaking, this is an untestable implication. The effect in Equation 14 can be estimated, but it will not be known whether

[15] That is, if it is assumed in the partially convergent case, that an exogenous increase in the relative popularity of party 1 induces candidates to move their positions towards party 1's bliss point (i.e. $\frac{\frac{\partial x_t}{\partial \delta_t}}{\frac{\partial \hat{P}_t}{\partial \delta_t}} > 0$ and $\frac{\frac{\partial y_t}{\partial \delta_t}}{\frac{\partial \hat{P}_t}{\partial \delta_t}} > 0$). This is the same qualitative result as the fully convergent case, as discussed in Alesina (1988).

[16] This assumes a monotonicity condition – incumbency cannot have a negative impact on the probability of election. See Hahn, Todd, and van der Klauuw (2000), which discusses the regression discontinuity design analogy to the local average treatment effect (LATE) of Imbens and Angrist (1994).

it equals $E_\Delta\left[c_{1it} - c_{2it}|SWING_t\right]$ given that c_{1it} and c_{2it} are unobservable to the researcher. However, in a relatively stationary environment, a good approximation to $E_\Delta\left[c_{1it} - c_{2it}|SWING_t\right]$ would be $E_\Delta\left[c_{1it-1} - c_{2it-1}|SWING_t\right]$, which can be independently estimated. Under stationarity, a substantial departure between estimates of the two quantities constitutes a rejection of the full divergence hypothesis.

Normally, with one instrument and one endogenous regressor, it is impossible to identify the three sub-populations described above, and hence estimation of expressions 12, 13, and 14 is infeasible. However, in our context, we can actually construct first-order approximations to these sub-populations in the data, because we observe the index VS_{it} which perfectly determines DEM_{it}.

Proposition 4 *Conditioning on the districts involved in close elections in* $t-1$, $\frac{1}{2} - \Delta < VS_{it-1} < \frac{1}{2} + \Delta$, *with* Δ *small, there exist* θ_1 *and* θ_2 *such that the three sub-populations can be, to the first-order, approximated as follows:*

$$\begin{aligned}
STRONG_t^{DEM} \textit{ if } DEM_{it-1} &= 0, DEM_{it} = 1 \\
\textit{or } DEM_{it-1} &= 1, VS_{it} > \frac{1}{2} + \theta_1 \\
STRONG_t^{REP} \textit{ if } DEM_{it-1} &= 1, DEM_{it} = 0 \\
\textit{or } DEM_{it-1} &= 0, VS_{it} < \frac{1}{2} - \theta_2 \\
SWING_t \textit{ if } DEM_{it-1} &= 0, \frac{1}{2} + \theta_1 \\
\textit{or } DEM_{it-1} &= 1, \frac{1}{2} - \theta_2 < VS_{it} > \frac{1}{2}
\end{aligned}$$

where θ_1 *and* θ_2 *are implicitly defined by* $\Pr[VS_{it} > \frac{1}{2} + \theta_1 | DEM_{it-1} = 1] = \Pr[DEM_{it} = 1| DEM_{it-1} = 0]$ *and* $\Pr[VS_{it} < \frac{1}{2} - \theta_2 | DEM_{it-1} = 0] = \Pr[DEM_{it} = 0| DEM_{it-1} = 1]$.

Thus, our testing procedure amounts to estimating the causal effects 12, 13, and 14 by dividing our sample of close elections in period $t-1$ into $STRONG_t^{DEM}$, $STRONG_t^{REP}$, and $SWING_t$ groups, and estimating the mean difference in roll call votes (RC_{it}) between Democrat and Republican districts in period $t-1$ (by DEM_{it-1}), for each group, respectively. Finding that the first two effects substantially differ from zero, and the third effect significantly departs from the benchmark $E_\Delta\left[x_{it-1}|DEM_{it-1} = 1, SWING_t\right] - E_\Delta\left[y_{it-1}|DEM_{it-1} = 0, SWING_t\right] \approx E\left[c_{1it-1} - c_{2it-1}|SWING_t\right]$, would constitutes evidence against the complete divergence hypothesis.

4 Empirical Analysis

4.1 Data Description: Roll Call Voting Records in the U.S. House

We begin by discussing the choice of the dependent variable. There is an excess of various ways one can measure politician's "behavior" in voting on legislation. A widely used measure is a voting score constructed by the liberal political organization, Americans for Democratic Action (ADA). For each Congress, the ADA chooses about 20 high-profile roll call votes, in order to construct an index that varies between 0 and 100 for each Representative of the House and member of the Senate. Higher scores correspond to a more "liberal" voting record. Throughout the paper, our preferred index of roll call votes is represented by ADA scores, although we show below that our results are robust to alternative interest groups scores and other voting record indices.

We utilize data on ADA scores for all Representatives in the U.S. House from 1946-1995, linked to election returns data during that period.[17] Contrary to what casual observers of the daily news may believe, there appears to be considerable variation in ADA scores *within* each party. This is illustrated in Figure 1, which provides the distribution of ADA scores for Democrat and Republican U.S. House of Representatives in the three most recent Congresses. To make the comparison across congresses possible, we follow the literature and use real ADA scores throughout the paper.[18] The figure shows significant overlapping in ADA scores between the parties, and it is not uncommon for Democrat representatives to vote more conservatively than Republican candidates, and vice versa.

One advantage of using ADA scores is that it is a widely used index in the literature. However, one limitation is that it includes only 20 votes per legislature, and the choice of what roll call vote to include and what weight to assign to each issue is necessarily arbitrary.[19] To assess how robust our results are to alternative measure of "liberalness" of roll call votes, we have re-estimated all our models using three sets of alternative measures.

First, for each Congress, we calculate the percent of a representative's votes that agree with the

[17] The data on roll call votes are from ICPSR xxxx. Data on electoral outcomes are from ICPSR yyyy.

[18] While nominal ADA scores are between 0 and 100, real ADA scores may be negative.

[19] Another disadvantage is that the choice of which vote to include is made at the end of the legislature, not at the beginning. (At the beginning of a legislature it is not obvious exactly which roll call will take place during a legislature.) This fact that the choice is made ex-post raises the possibility that the criteria use to assign ADA scores could be somewhat endogenous.

Democrat party leader or the Democrat party whip. These two measures have the advantage of including *all* votes, not an arbitrary subset of votes.

Second, we use ratings from interest groups other than ADA. We include both liberal and conservative ratings from groups such as the American Civil Liberties Union, the League Women Voters, the League Conservation Voters, the American Federation of Government Employees, the American Federation of State, County, Municipal Employees, the American Federation of Teachers, the AFL-CIO Building and Construction, the United Auto Workers, the Conservative Coalition, the US Chamber of Commerce, the American Conservative Union, the Christian Voters Victory Fund, the Christian Voice, Lower Federal Spending, and Taxation with Representation. Not all the ratings are available in all years, so sample size varies when using these alternative ratings.

A legitimate concern is that interest group may choose the criteria used to calculate their ratings based on partisan considerations. This could happen if partisan considerations enter interest groups decisions on which votes to include in the ratings, or the weight assigned to each vote. If this is the case, our finding could overstate the coefficient on party affiliation. Given that we use ratings from 15 different interest groups that are likely to use completely different criteria to create their ratings, we do not expect this to be a major problem.

However, we address this concern directly by creating our own ratings using roll call votes on abortion. We choose abortion because the classification of votes as pro-life or pro-choice is straightforward in most cases. For the three most recent Congresses for which we have data (100th to 103rd), we identify all roll call votes that contain the word "abortion" in the title or the description, and assign each candidate a zero if the candidate voted pro-life and one if the candidate vote pro-choice. We then calculate the percentage of pro-choice votes for each candidate. Such index is more objective than the interest group ratings, because it includes *all* votes on a specific issue, not just a selected sample, and assigns equal weight to all votes, not arbitrary weights.

As we show below, the qualitative results of our estimation are remarkably stable across alternative measures of roll call votes. This finding lends some credibility to the conclusion that our estimates are not driven by the unique characteristics of one particular measure. See the Data Appendix for a detailed discussion of our samples and data sources.

4.2 Graphical Analysis of the Discontinuity, and Tests of Random Assignment

In the remainder of this section we use the framework described in Section 3 to distinguish between full convergence, partial convergence and full divergence. We address the issues of unobserved parties' and legislators' preferences by using a regression discontinuity analysis based on close elections. We compare the roll call behavior of representatives from districts where the two-party vote share for the Democratic candidate is just below 50% with the roll call behavior of representatives from districts where the vote share for the Democratic candidate is just above 50%.

We begin by visually showing the discontinuity around the 50% threshold. We present two specification tests that lend some support to our assumption that close elections provide as good as randomized variation in party control of a seat. We then turn to the main results of the paper. First, we test for full convergence using ADA scores as a measure of roll call records. Having rejected full convergence, we then determine whether the data are more consistent with either complete divergence or partial divergence. We do this first under the assumption of homogeneity, and then under the more general assumption of heterogeneity. Finally, we show that our results do not change when we use alternative measures of roll call votes.

The top panel of Figure 2 plots ADA scores against the Democrat vote share. Data are for years 1946 to 1995. Throughout the paper, the unit of observation is the district at a given point in time. But to give a general picture of the data, each point in Figure 2 is an average of the ADA score within 0.001-wide intervals of the vote share. The vertical line marks 50% of the two-party vote share. Districts to the right of the vertical line are Democrat, districts to the left are Republican. The continuous line in Figure 2 is the predicted ADA scores from a regression that includes a 4th-order polynomial in vote share. A striking feature of the figure is that ADA scores appear to be a continuous and smooth function of vote shares everywhere, except at the threshold that determines party membership. There is a large discontinuous jump in ADA scores at the 50% threshold, indicating that representatives from districts with similar vote shares have very different roll call behavior depending on the party to which they belong.

Compare a district where the Democrat candidate barely lost (for example, vote share is 49%), with a district where the Democrat candidate barely won, (for example, vote share is 51%). As argued in Section 3, under the hypothesis of full convergence, we should observe little differences around the 50% threshold. On the contrary, it seems that representative from districts with almost identical vote shares have widely different roll call records. The difference at the 50% threshold appears quite large. From the Figure it seems that representatives from districts on the Democrat side of the 50% threshold have ADA scores that are almost 50 points higher than representatives from districts on the Republican side.[20]

The key identifying assumption in this paper is that, as one compares closer and closer elections, all pre-determined characteristics of Republican and Democrat districts (including the district-specific bliss points) become more and more similar. We provide two pieces of evidence to support this assumption. First, in the bottom panel of Figure 2, we plot once-lagged ADA scores against the Democrat vote share. If our identifying assumption is correct, we should observe no discontinuity in the relationship between lagged ADA scores (which are already determined before the assignment of the "treatment") and the vote share. If, on the contrary, the jump in the top panel is caused by some permanent characteristics of districts that has nothing to do with the party affiliation, then we should observed a jump. The lack of discontinuity lends some credibility to our identifying assumption.[21]

As a second specification check, we show that as we compare closer and closer elections, Republican and Democrat districts have other similar pre-determined characteristics. Consider, for example, geographical location. There are sizable geographical differences in the full sample. Democrats are significantly more likely to be elected in the South than in the North and the West. However, as we start restricting the sample to closer and closer elections, the geographical differences decrease. For elections that are only within two percentage points from the threshold, the differences are not statistically significant.

[20] A second striking feature of Figure 2 beside the jump at the threshold, is the apparent lack of correlation between vote share and roll call behavior.

[21] The estimated gap is 3.5 (5.6).

This is shown in Figures 3 and 4, which plot district characteristics against vote share. Beside geographical location, we consider the following pre-determined characteristics: real income, percentage with high school degree, percentage black, percentage eligible to vote, and size of the voting population.[22] Generally, the figures indicate that the magnitude of the discontinuity around the 50% threshold is limited.[23] Overall, we conclude that in a close neighborhood of 50% Republican and Democrat districts have similar pre-determined characteristics, lending credibility to our main identifying assumption.

4.3 Differentiating Between Convergence, Complete divergence and Partial Divergence under Homogeneity

The size of the discontinuity in ADA scores documented in Figure 2 is more precisely quantified in Table 1. All the entries in the table are based close elections. Specifically, we include only districts with Democrat vote share at time $t-1$ between 48% and 52%: $48\% < VS_{t-1} < 52\%$.

There are 915 such districts. We begin, in column 1, by reporting the coefficient on DEM_t in a regression of ADA_t on DEM_t, including only districts in which the vote share *in the previous election* is between 48% and 52%. The estimated coefficient is 48.8.

Column 2 reports an estimate of Equation 8. Specifically, it shows the RD estimate in a regression of ADA_{t-1} on DEM_{t-1}. The mean difference in ADA scores at time $t-1$ between Democrat and Republican representatives that were elected in close elections is 47.6. This estimate corresponds to the magnitude of the discontinuity documented in the top panel of Figure 2. In the framework proposed in Section 3, such large difference is a strong rejection of the full convergence hypothesis.

Having rejected full convergence, we now turn to the question of whether the equilibrium can be best described as complete divergence or partial divergence. In Section 3, we argue that one way to distinguish between full divergence and partial divergence is to analyze the effect of an exogenous change in the probability of victory of a candidate. Under full divergence, we should observe little change in

[22] Data on districts characteristics in each election year are from the last available Census of Population. Because the census takes place every ten years, standard errors allow for clustering at the district-decade level.

[23] One exception is the percentage blacks, for which the magnitude of the discontinuity is statistically significant. This is due to few outliers in the outer part of the vote share range. When the polynomial is estimated including only districts with vote share between 25% and 75%, the coefficients becomes insignificant.

the representative roll call behavior. Under partial divergence, the representative should move toward her party's bliss point. We empirically implement this test by using the party that wins the election at time $t-1$ as an instrument for the party that wins at time t, conditioning on close elections at time $t-1$. As shown in Proposition 3 above, under the null hypothesis of full divergence, RD-IV estimates are consistent and should be similar to the RD estimates shown in column 2 of Table 1. Under the alternative hypothesis of partial divergence, RD-IV estimates are not consistent and are likely to exceed RD estimates.

We begin by graphically showing the first stage in the top panel of Figure 5, which plots DEM_t against VS_{t-1}. The large discontinuity around the 50% threshold confirms that there exists an incumbency advantage, as documented in Lee (2001,2002). For representatives who were elected in close elections at time $t-1$, such an incumbency advantage translates into an exogenous change in the probability of election at time t, which we use to distinguish between full divergence and partial divergence. The effect is quantified in column 3 of Table 1, which suggests that the first stage coefficient is 0.48.

We now report the reduced-form estimates. In column 4, we show the coefficient on DEM_{t-1} in a regression of ADA_t on DEM_{t-1}. The estimated coefficient is 21.2. The same effect is shown graphically in the bottom panel of Figure 5, which plots ADA scores at time t against vote share in the previous election.

Finally, we turn to the RD-IV estimator described in Equation 11. The estimate in column 4 is 43.9, slightly smaller than the RD estimate in column 2. Comparing the RD-IV to the RD estimate, we cannot reject the hypothesis of full divergence.

4.4 Differentiating Between Convergence, Complete divergence and Partial Divergence Under Heterogeneity

We now turn to a more general framework where the *difference* between the parties' bliss points vary across districts and legislatures. Again, we focus on the 915 districts that experience close elections at time $t-1$.

We divide this group of districts in three sub-groups, based on the vote share at time t. *Strong Democrat districts* include districts where a Republican (barely) won at time $t-1$ but a Democrat regained the district at time t. Strong Democrat districts also include the districts with the largest Democrat vote share at time t, among the districts where a Democrat won at time $t-1$.[24] There are 224 such districts. In Angrist, Imbens and Rubin (1995) these districts would be called "always takers", because irrespective of the "random assignment" of DEM_{t-1}, they are Democrat at time t. Under complete divergence, we expect the ADA scores of representatives from strong Democrat districts at time t to be unaffected by who wins the election at time $t-1$.

Similarly, *strong Republican distrcits* include those where a Democrat (barely) won at time $t-1$ but a Republican regained the district at time t, as well as districts with the largest Republican vote share at time t, among the districts where a Republican won at time $t-1$.[25] There are 250 such districts. In Angrist, Imbens and Rubin (1995) these districts would be called "never takers", because irrespective of the "random assignment" of DEM_{t-1}, they are Republican at time t. Again, under complete divergence, we expect the ADA scores of representatives from strong Republican districts at time t to be unaffected by who win the elections at time $t-1$.

Finally, we call the remaining districts *swing districts*. The swing districts are marginal districts that are neither safe Republican nor safe Democrat. These are the districts where a Democrat or Republican victory at time $t-1$ is expected to have the largest effect on ADA scores at time t. There are 441 swing districts.

In the bottom part of Table 1 we report separate estimates for the three groups. Column 2 shows the coefficients from regressions of ADA_{t-1} on DEM_{t-1} for the three groups. The estimates are 52.1 for strong Democrat districts, 49.6 for swing districts and 44.4 for strong Republican districts. This suggests that the RD estimate of Equation 8 does not vary significantly across the three groups.

The key results are shown in column 4, where we regress ADA_t on DEM_{t-1}. The coefficients for strong Democrat districts and strong Republican districts are not statistically different from zero. The coefficient for swing districts is the only one different from zero. Importantly, it is very similar to the RD

[24] Specifically, we included the top 24.18% districts with the largest democrat vote share among the districts where a democrat won at time $t-1$. The reason why we included the top 24.18 districts is that among all the districts where the democrat lost at time $t-1$, the democrat candidate won at time t in 24.18% of the cases. Because there is "random assignment" at time t-1, the probability of winning the elections at time t for a democrat candidate *irrespective of the outcome at time t-1* should be the same.

[25] Specifically, we included the 27.39% districts with the largest republican vote share, since among all the districts where the republican lost at time $t-1$, the republican candidate won at time t in 27.39% of the cases.

coefficient in column 2. Based on the discussion in Section 3, these results strongly suggest the data ore more consistent with full divergence in this heterogenous case as well.

4.5 Alternative Measures of Roll Call Records

The results presented so far are not specific to the ADA scores. We now show that the findings presented so far can be replicated using alternatives measures of roll call records.

Figure 6 is analogous to Figure 2, but instead of using ADA scores, it is based on the percent of votes cast that are equal to the vote cast by the Democrat party leader. Table 2 quantifies the magnitude of the discontinuity and reports RD-IV estimates. All the qualitative results of Table 1 hold up using this measure. The RD estimate is 0.291, remarkably close to the RD-IV estimate in column 5, which is 0.294. A pattern similar to the results for the ADA scores emerges in the bottom panel of Table 2, when the heterogeneous model is estimated. The effects of RC_t on DEM_{t-1} are virtually zero for strong Democrat and Republican districts, as predicted in Section 3. The reduced-form coefficient for the swing districts is .30, remarkably similar to the RD estimate in column 2.

Ratings from different interest groups yield qualitatively similar results. This is demonstrated in Figures 7, which is based on ratings from several liberal and conservative interest groups. Liberal groups include: American Civil Liberties Union, the League Women Voters, the League Conservation Voters, the American Federation of Government Employees, the American Federation of State, County, Municipal Employees, the American Federation of Teachers , the AFL-CIO Building and Construction, the United Auto Workers. Conservative groups include: the Conservative Coalition , the US Chamber of Commerce, the American Conservative Union, and the Christian Voice. All the ratings range from 0 to 100. For liberal groups, low ratings correspond to conservative roll call votes, and high ratings correspond to liberal roll call votes. For conservative groups, the opposite is true. The figure plots the RD estimates against the RD-IV estimates for each group. The line is the 45^o degree line. Most estimates are on the line or close to the line, indicating again that across a variety of different interest groups scores, the results are highly consistent with the full divergence hypothesis.

Similar results are obtained from the ratings that we constructed based on roll call votes on abortion. Just to give an example, Figure 8 shows a pattern similar to the one uncovered for ADA scores in Figure 2. This suggests that the discontinuity documented in Figure 2 is not just a consequence of partisan criteria used by ADA.

5 Relation with Previous Studies

Previous empirical studies have examined the extent to which party affiliation influences roll call voting behavior. For example, Levitt (1996) argues that national party line is a key determinant of roll-call voting behavior, and that constituents' preferences are assigned only a small weight in senator utility functions. Levitt's conclusion is consistent with the position of many political scientists, who suggest that political parties are more important than voters' preferences in shaping legislation. Poole and Rosenthal (1984), for example, show that senators from the same state belonging to different parties have significantly different voting records.

More recently, Snyder and Groseclose (2000) estimates the effect of party affiliation on roll call votes using an innovative identification strategy based on lopsided votes. Their strategy depends on the assumption that on votes with 65 percent or more legislators on one side, legislators are not subject to party influence. They conclude that in the majority of cases, party affiliation is a significant determinant of roll calls. In a recent review of the literature, Snyder and Groseclose (2000) conclude that "virtually all studies of roll-call voting in the United States Congress in the political science literature find that political party affiliation is one of the best predictors of voting behavior".[26] Indeed, the agreement between these findings and our quasi-experimental results on the issue of *full convergence* suggest that the previous literature's findings are not simply artifacts of selection bias. Indeed, our own analysis (Figure 2) reveals *ex post*, that a simple difference in means would not have been particularly biased for the RD estimand.

On the other hand, the existing literature does not directly address the question of whether roll

[26] The literature is too large to be summarized here. Examples include, but are not limited to, Poole and Rosenthal (1984, 1996), Snyder and Ting (2001a), Snyder and Ting (2001b), Lott and Davis (1992), Canes-Wrone, Brady and Cogan (2000), Krehbiel (2000), Bender (1991), Lott (1990), McArthur and Marks (1988), Douglas and Sielberger (1987), McCarty, Poole and Rosenthal (2000).

call voting behavior in Congress is best described as a partial convergence or complete policy divergence equilibrium. For example, in Levitt (1996), it is *presumed* that the Senator is making a non-trivial trade-off between voters preferences and his own ideology, and the presumption that the marginal conditions are met for an "interior solution" is what justifies the fixed effects regression analysis. This is useful for estimating parameters of the implied utility function, but this approach is not informative about whether the equilibrium is an "interior" or "corner" solution in the first place. Indeed, Alesina (1988) shows that in a dynamic context, the per-period first-order conditions are *not* met in the fully divergent equilibrium.[27]

Other studies in the literature focus on the question of whether or not roll call votes are correlated with alternative measures of constituents' preferences. However, this correlation has little bearing on whether opposing candidates, *conditional* on the preferences of voters in the district, deviate from their most-preferred position, in response to pressure from voters. For example, districts where voters are conservative may attract relatively conservative candidates from *both parties*, so it may be unsurprising to observe this correlation. However, that correlation is not informative about whether the difference between opposing candidates *within* those districts is smaller because of the threat of losing the election (partial policy convergence), or whether it is as large as it could possibly be (complete policy divergence).

6 Conclusion

In this paper, we take advantage of quasi-random assignment of party control over seats to empirically assess the extent of policy convergence and divergence in the roll call voting patterns of Representatives in the U.S. House. This allows us to measure the average degree of divergence between opposing candidates, even though the outcome the losing candidate would have implemented is strictly unobserved. Our analysis reveals a striking divergence in policy, and is strongly inconsistent with the complete policy divergence hypothesis. In fact, the degree of divergence among closely contested elections is no smaller than among districts that experienced land-slide electoral victories.

The effective random assignment of party control also allows us to differentiate between complete

[27] When the first order conditions are met, Alesina's equation (6) holds, but the complete divergence equilibrium is in general different.

divergence and partial convergence. We find candidates' chosen positions do not respond to exogenous changes in the probability of winning the election, which is a direct implication of the complete policy divergence equilibrium. Thus our evidence is more consistent with full divergence. It may be possible to construct a partially convergent equilibrium whereby positions are not affected by exogenous changes in the relative popularity of the candidates. However, given that candidates' bliss points are unobservable to the researcher, that kind of equilibrium would eliminate any remaining empirical content of the partial convergence notion.

Appendix A.

Proof of Proposition 1 Denote the joint density of VS_{it}, c_{1it}, c_{2it}, and δ_{it} as $g(vs, c_1, c_2, \delta)$ and the density of VS_{it} as $h(vs)$. $g(vs, c_1, c_2, \delta)$ is continuous with respect to it arguments because vs is a continuous function with respect to its arguments, and c_{1it}, c_{2it}, δ_{it}, and ε_{it} has continuous joint density by assumption. For the same reasons, h is also continuous with respect to vs. Thus $\frac{g(vs,c_1,c_2,\delta)}{h(vs)}$ is continuous. See DiNardo and Lee (2002) and Lee (2002) to see how the result holds even if the heterogenous components $(c_{1it}, c_{2it}, \delta_{it})$ have finite and discrete support, and only ε_{it} has continuous density.

Proof of Proposition 4 Condition on "close elections" in $t-1$. Consider the sub-population $DEM_{it-1} = 0, DEM_{it} = 1$. Clearly they belong to the group $STRONG^{DEM}$, by definition ($DEM_{it-1} = 0, DEM_{it} = 0$ do not). We need to show that, to a first-order approximation, for the population $DEM_{it-1} = 1$, Democrats would have won even if $DEM_{it-1} = 0$, if and only if $VS_{it} > \frac{1}{2} + \theta_1$.

Consider the "vote production function", in equilibrium $VS_{it} = vs(x(c_{1it}, c_{2it}, \delta_{it}), y(c_{1it}, c_{2it}, \delta_{it}),$ $\delta_{it}, \varepsilon_{it})$. Taking a linear approximation to this function around the equilibrium yields (normalizing all coefficients to 1) $VS_{it} \approx -c_{1it} - c_{2it} + \delta_{it} + \varepsilon_{it}$. Furthermore, given Proposition 2, we can decompose δ_{it} so that $VS_{it} \approx -c_{1it} - c_{2it} + \gamma DEM_{it-1} + \delta^*_{it} + \varepsilon_{it}$. Now, θ_1 is defined so that $\Pr[-c_{1it} - c_{2it} + \delta^*_{it} + \varepsilon_{it} > \frac{1}{2}] = \Pr[-c_{1it} - c_{2it} + \gamma + \delta^*_{it} + \varepsilon_{it} > \frac{1}{2} + \theta_1]$ which implies that $\theta_1 = \gamma$. This implies that for $DEM_{it-1} = 1$, $-c_{1it} - c_{2it} + \delta^*_{it} + \varepsilon_{it} > \frac{1}{2}$ if and only if $VS_{it} > \frac{1}{2} + \theta_1$.

A similar argument holds for the $STRONG^{REP}$ and $SWING$ districts.

References

[1] Alesina, Alberto (1988). "Credibility and Policy Convergence in a Two-Party System with Rational Voters", *American Economic Review*, 78, 796-805.

[2] Alesina, Alberto and Spear, Stephen E. (1988). "An Overlapping Generations Model of Electoral Competition", *Journal of Public Economics*, 37, 359-379.

[3] Angrist, J., Imbens, G., and Rubin, D. (1996) "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444-445.

[4] Besley, Timothy, and Stephen Coate (1997). "An Economic Model of Representative Democracy", *Quarterly Journal of Economics*, 112, 85-114.

[5] Besley, Timothy, and Stephen Coate (1998). "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis", *American Economic Review*, 88, 139-156.

[6] Bullock and Brady "Party Constituency, and Roll-Call Voting in the US Senate" Legislative Studies Quarterly 8 (1983).

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

[8] Canes-Wrone, Brady and Cogan "Electoral Accountability and House Members' Voting" mimeo (2001)

[9] DiNardo, John, and David S. Lee (2002). "The Impact of Unionization on Establishment Closure: A Regression Discontinuity Analysis of Representation Elections," *NBER Working Paper #8993*, June 2002.

[10] Downs, Anthony (1957). *An Economic Theory of Democracy*, New York, NY: Harper and Row.

[11] Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw (2001). "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica* 69, 201-209.

[12] Heckman and Snyder "A Linear latent Factor Voting Model", Rand Journal of Economics 28-0 (1997).

[13] Hinich, Melvin, Ledyard, John and Ordeshook, Peter (1972). "Nonvoting and the Existence of Equilibrium Under Majority Rule," *Journal of Economic Theory*, 4, 144-53.

[14] Imbens, G., and Angrist, J. (1994). "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62 (4): 467-476.

[15] Kalt and Zupan "Capture and Ideology in the Economic Theory of Politics" American Economic Review 74-3 (1984)

[16] Kalt and Zupman "The Apparent Ideological Behavior of Legislators: Testing Principal-Agent Slack in Political Institutions" Journal of Law and Economics 33-1 (1990)

[17] Kau and Rubin "Self-Interest, Ideology and Log-rolling in Congressional Voting" Journal of Law and Economics 22-2 (1979)

[18] Lee, David S. (2001) "The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Close Elections", *UC Berkeley Center for Labor Economics, Working Paper #31*, April 2001.

[19] Lee, David S. (2002)."Regression Discontinuity Analysis of the Incumbency Advantage in the U.S. House", *UC Berkeley Manuscript*, June 2002.

[20] Levitt "How Do Senators Vote? Disentangling the Role of Voter Preferences, Party Affiliation, and Senator Ideology" American Economic Review, 86-3 (1996)

[21] McKelvey, Richard (1975). "Policy Related Voting and Electoral Equilibrium," *Econometrica*, 43, 815-63.

[22] Osborne, Martin J. (1995). "Spatial models of Political Competition under Plurality Rule: A survey of Some Explanations of the Number of Candidates and the Position they Take", Canadian Journal of Economics, 28, 261–301.

[23] Peltzman “An Economic Interpretation of the History of Congressional Voting in the Twentieth Century” American Economic Review 75-4 (1985)

[24] Persson, Torsten, and Guido Tabellini (2000). *Political Economics*. Cambridge, MA: MIT Press.

[25] Poole and Daniels “Ideology, Party and Voting in the US Congress 1959-1980” American Political Science Review 79-2 (1985)

[26] Poole and Rosenthal “The Polarization of American Politics” Journal of Politics 46-4 (1984)

[27] Poole and Rosenthal “A Spatial Model for Roll Call Analysis” American Journal of Political Science 29- 2 (1985)

[28] Poole and Rosenthal “Patterns of Congressional Voting” American Journal of Political Science 35-1 (1991)

[29] Snyder “Artificial Extremism in Interest Group Ratings” Legislative Studies Quarterly 17-3 (1992)

[30] J. Snyder and T. Grosecolse “Estimating Party Influence in Congressional Roll- Call Voting” American Journal of Political Science, 44-2, pp.187-205 (2000)

[31] J. Snyder and M. Ting “Party Labels, Roll-Call Votes, and Elections” mimeo (2001)

[32] J. Snyder and M. Ting “An Information Rational for Political parties” mimeo (2001)

[33] Wittman, Donald (1983). “Candidate Motivation: A synthesis of Alternative Theories”, *American Political Science Review*, 77, 142-157.

Table 1: The Difference in Republican and Democrat Representative's ADA Scores Conditioning on Close Elections at Time $t-1$

		RD	First Stage	Reduced Form	RD-IV	N
	ADA_t	ADA_{t-1}	DEM_t	ADA_t	ADA_t	
	(1)	(2)	(3)	(4)	(5)	(6)
Homogeneity						
Coefficient on DEM_t	48.8				43.9	915
	(1.3)				(2.7)	
Coefficient on DEM_{t-1}		47.6	0.48	21.2		915
		(1.3)	(0.02)	(1.9)		
Heterogeneity						
Strong Democrat Districts						
Coefficient on DEM_{t-1}		52.1		-1.7		224
		(3.1)		(3.0)		
Swing Districts						
Coefficient on DEM_{t-1}		49.6		47.4		441
		(1.8)		(1.8)		
Strong Repubblican Districts						
Coefficient on DEM_{t-1}		44.4		-4.6		250
		(2.5)		(2.3)		

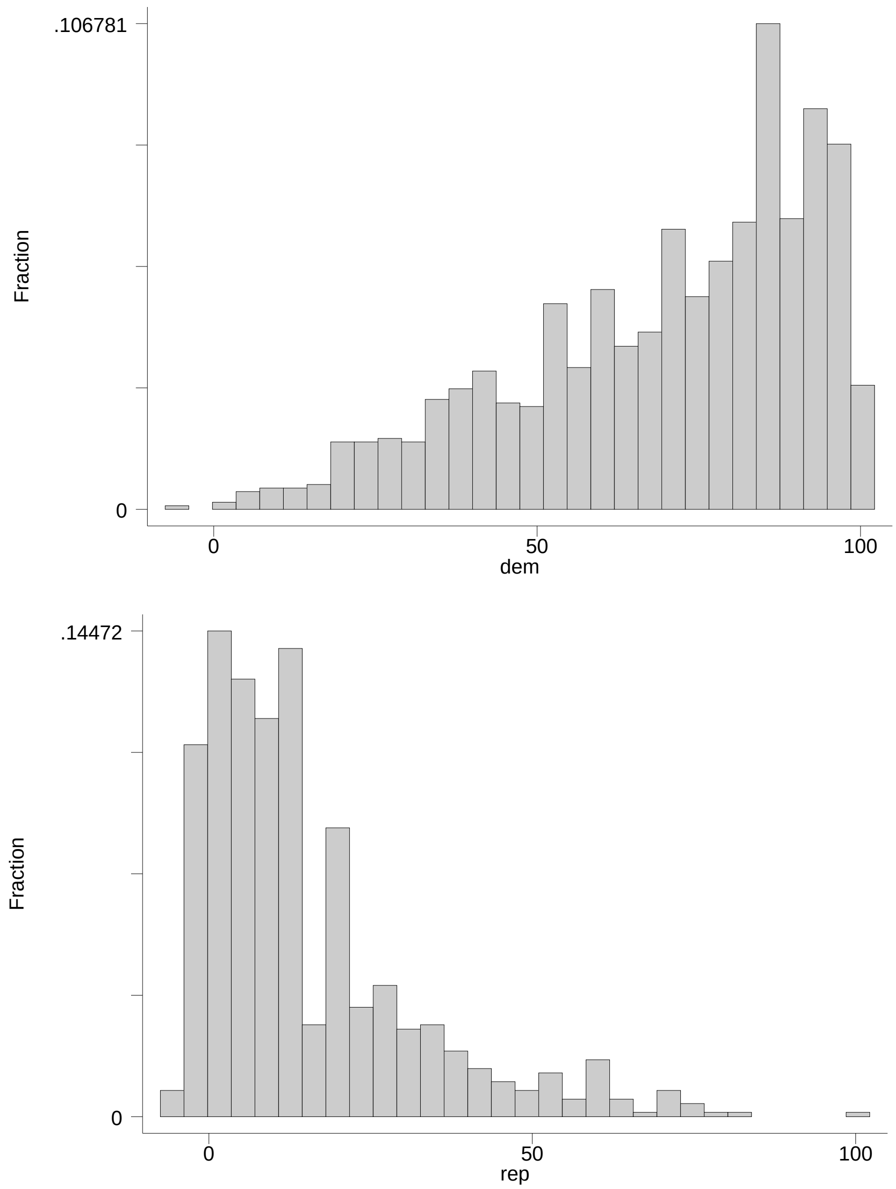
Notes: Standard errors in parenthesis. Only observations for which the democrat vote share at time $t-1$ is between 48% and 52% are included.

Table 2: The Difference in Republican and Democrat Representative's Probability of Voting like the Democrat Leader Conditioning on Close Elections at Time t-1

		RD	First Stage	Reduced Form	RD-IV	N
	P_t	P_{t-1}	DEM_t	P_t	P_t	
	(1)	(2)	(3)	(4)	(5)	(6)
Homogeneity						
Coefficient on DEM_t	.302				.294	1010
	(0.007)				(.015)	
Coefficient on DEM_{t-1}		.291	0.46	.138		1010
		(0.006)	(0.02)	(.011)		
Heterogeneity						
Strong Democrat Districts						
Coefficient on DEM_{t-1}		.290		.001		248
		(.016)		(.011)		
Swing Districts						
Coefficient on DEM_{t-1}		.290		.308		474
		(.009)		(.010)		
Strong Repubblican Districts						
Coefficient on DEM_{t-1}		.294		.024		288
		(.012)		(.016)		

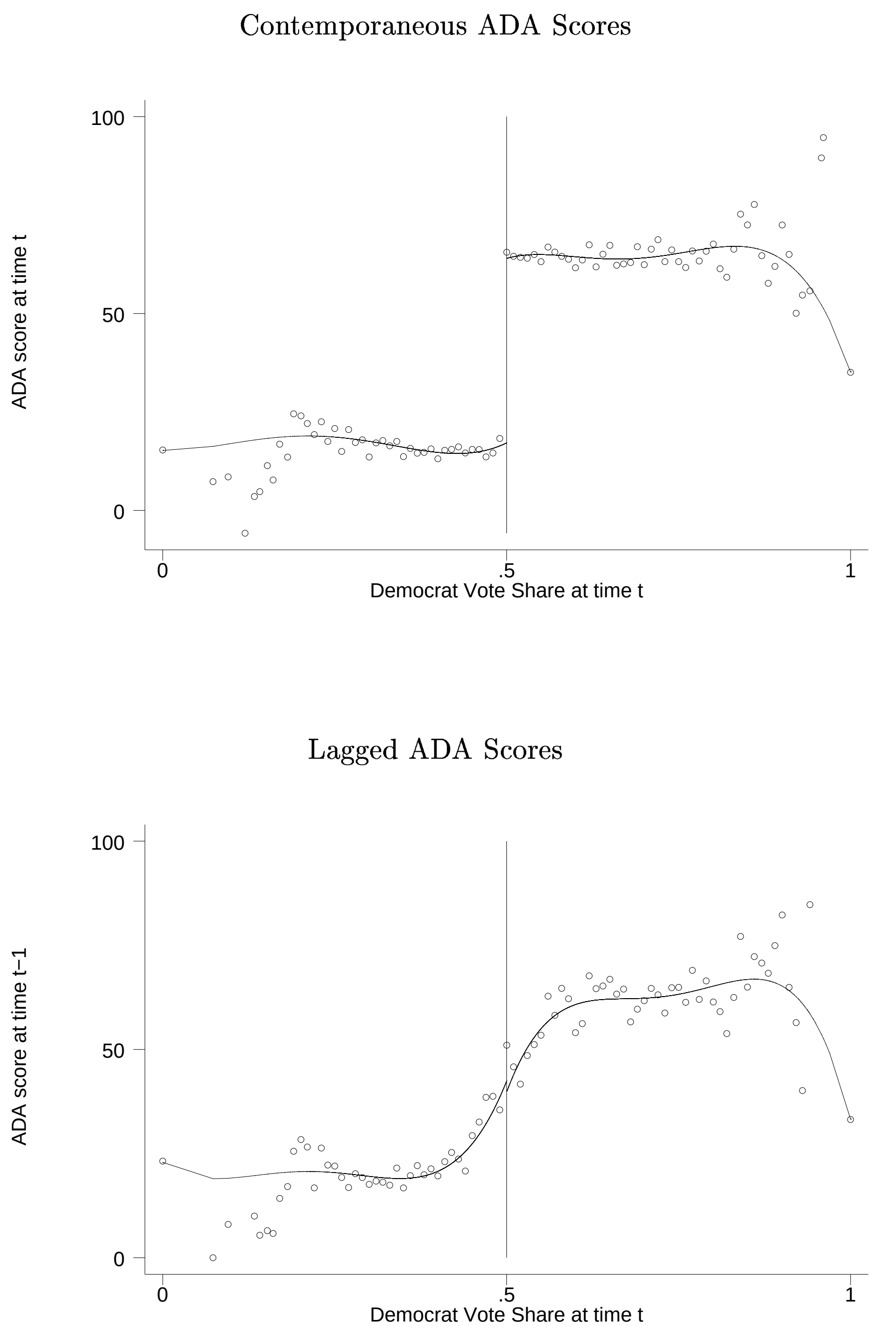
Notes: Standard errors in parenthesis. P_t is the percent roll call votes that a candidate vote is equal to the Democrat leader vote. Only observations for which the democrat vote share at time $t-1$ is between 48% and 52% are included.

Figure 1: Distribution of ADA Scores, by Party



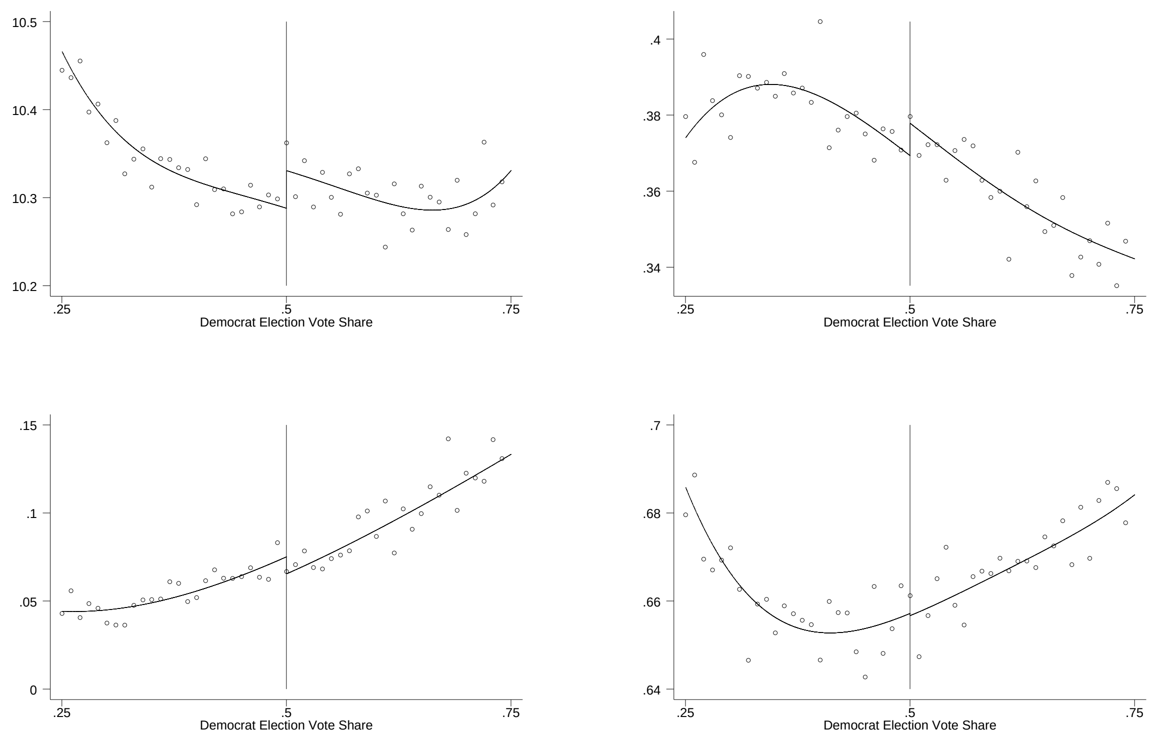
Notes: Real ADA scores for Congresses 102, 103 and 104.

Figure 2: The Relationship between ADA Scores and Democrat Vote Share



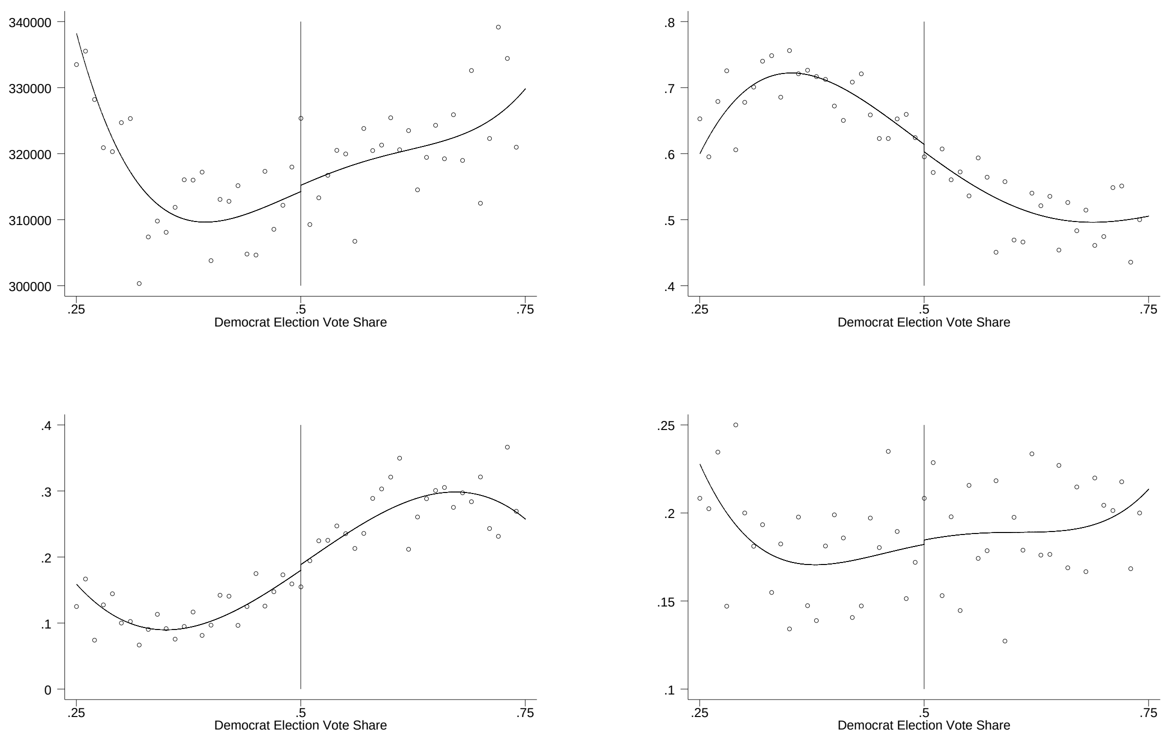
Notes: The top panel plots ADA scores against the democrat vote share. The bottom panel plots one time lagged ADA scores against the democrat vote share. Each point is the average lagged ADA score within intervals 0.001 wide.

Figure 3: Are Constituents Characteristics Different in Democrat and Republican Districts? Part 1



Notes: Panels refers to (from top left to bottom right) the following District Characteristics: real income, percentage with high-school degree, percentage black, percentage eligible to vote.

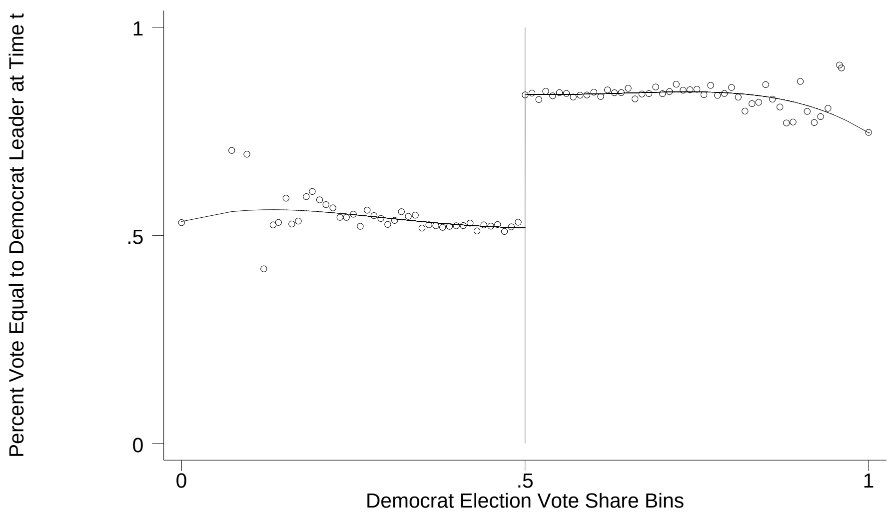
Figure 4: Are Constituents Characteristics Different in Democrat and Republican Districts? Part 2



Notes: Panels refers to (from top left to bottom right) the following District Characteristics: voting population, North, South, West.

Figure 5: The Relationship between Probability of Voting Along Party Lines and Vote Share

Percent Vote Equal to Vote of Democrat Party Leader



Percent Vote equal to Democrat Party Whip

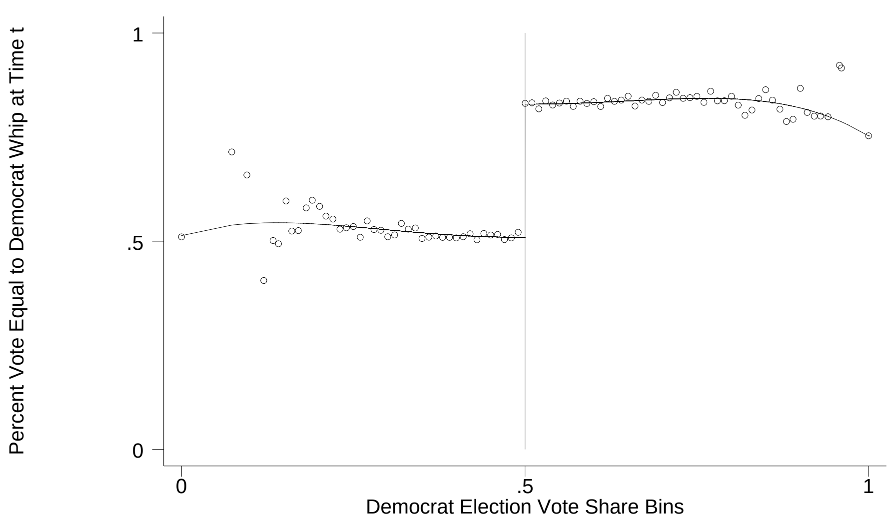
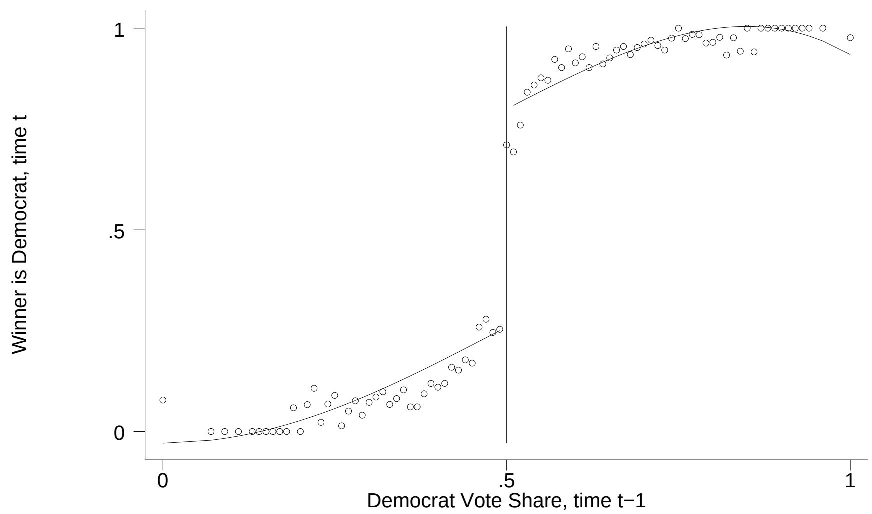


Figure 6: Probability of Democrat Win Election at time t, by Democrat Vote Share at time t-1; and ADA Scores at time t, by Democrat Vote Share at time t-1.

Probability of Democrat Win Election at time t



Probability of Democrat Win Election at time t

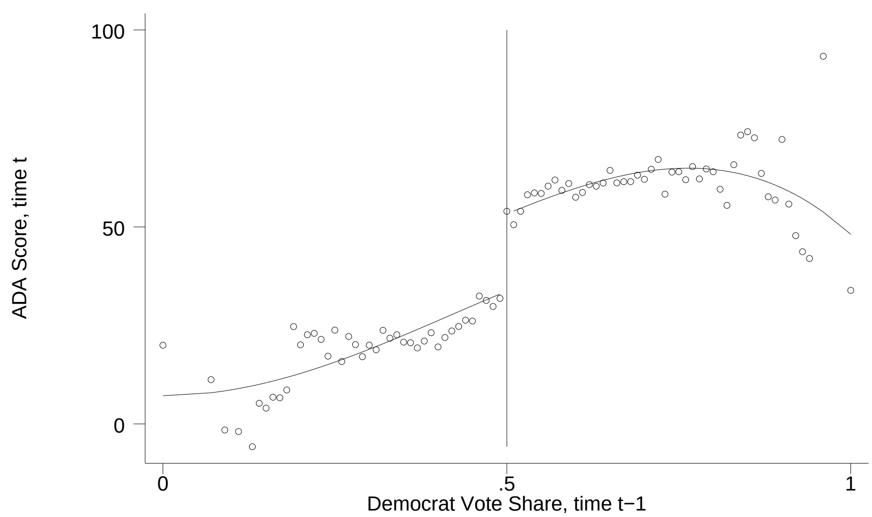
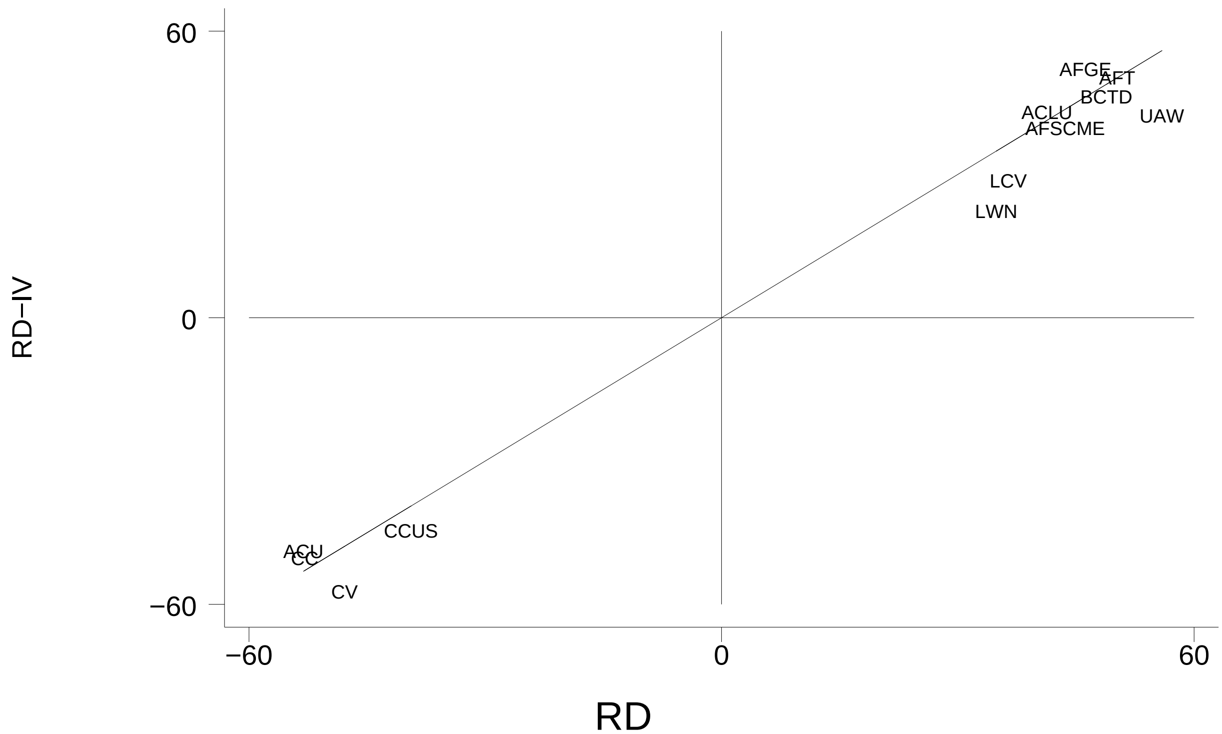
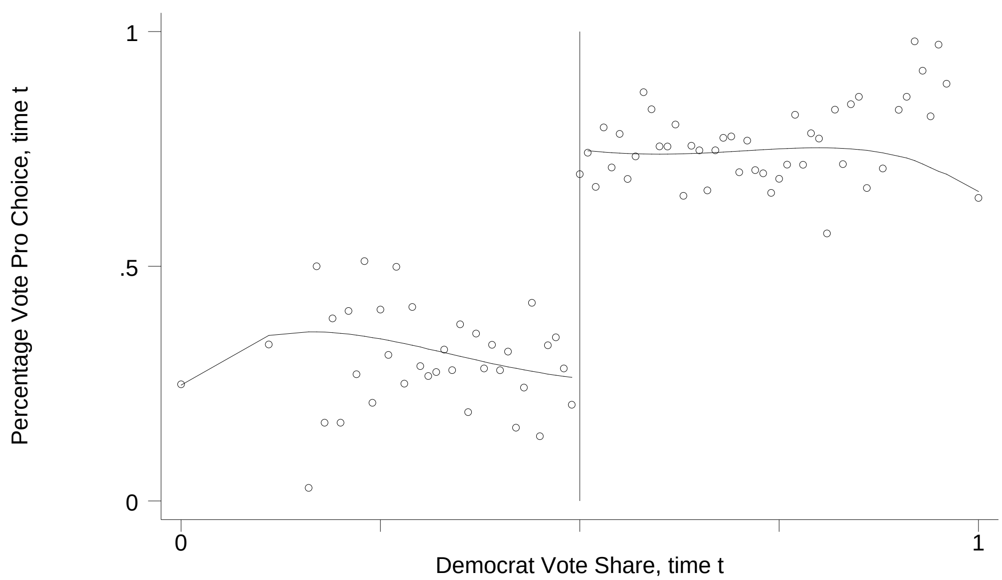


Figure 7: Regression Discontinuity Estimates versus Regression Discontinuity-IV Estimates, for Alternative Interest Group Scores



Note : ACLU is Am. Civil Liberties Union (N=16); LWV is League Women Voters (N=117); LCV is League Conservation Voters (N=111); AFGE is Am. Fed. Government Employees (N=60); AFSCME is Am. Fed. State, County, Mun. Emp. (N=131); AFT is American Federation of Teachers (N=137); BCTD is AFL-CIO Building and Construction ((=12); UAW is United Auto Workers (N=204); CC is Conservative Coalition (N=493); CCUS is US Chamber of Commerce (N=131); ACU is American Conservative Union (N=132); CV is Christian Voice (N=12).

Figure 8: The Effect of Party Affiliation on Roll Call Votes on Abortion



Appendix Table 1: The Difference in Republican and Democrat Representative's Propensities to Vote like the Democrat Party Leader, by Quarter, Conditional on Close Elections

	(1)
(1) First Quarter	.40
	(0.01)
(2) Second Quarter	.34
	(0.01)
(3) Third Quarter	.31
	(0.01)
(4) Fourth Quarter	.30
	(0.01)
(5) Fifth Quarter	.30
	(0.01)
(6) Sixth Quarter	.30
	(0.01)
(7) Seventh Quarter	.28
	(0.01)
(8) Seventh Quarter	.31
	(0.01)

Notes: Standard errors in parenthesis. The dependent variable is the percentage of roll call votes casted by a representative in a legislature that are equal to the vote of the democrat party leader. Only observations for which the democrat vote share is between 48% and 52% are included. All entries are from different regressions.

Appendix Table 2: The Difference in Republican and Democrat Representative's Propensities to Vote like the Democrat Party Leadership, by Region and Decade, Conditional on Close Elections

	North East	Mid West	South	West
	(1)	(2)	(3)	(4)
1940	0.37	0.39	0.42	0.40
	(0.01)	(0.01)	(0.01)	(0.01)
1950	0.27	0.37	0.19	0.31
	(0.01)	(0.01)	(0.01)	(0.01)
1960	0.27	0.36	0.26	0.36
	(0.01)	(0.01)	(0.01)	(0.01)
1970	0.14	0.21	0.18	0.27
	(0.01)	(0.01)	(0.01)	(0.01)
1980	0.18	0.30	0.34	0.37
	(0.01)	(0.01)	(0.01)	(0.01)
1990	0.41	0.44	0.46	0.47
	(0.01)	(0.01)	(0.01)	(0.01)

Notes: Standard errors in parenthesis. Only observations for which the democrat vote share is between 48% and 52% are included. All entries are from different regressions. The dependent variable is the percentage of roll call votes casted by a representative in a legislature that are equal to the vote of the democrat party leader.