Noisy Retrospection: The Effect of Party Control on Policy Outcomes

John B. Holbein
Assistant Professor
Brigham Young University
john_holbein@byu.edu

Adam Dynes
Assistant Professor
Brigham Young University
adamdynes@byu.edu *

February 9, 2018

Retrospective voting is thought to be vital for democracy. But, are the objective performance metrics widely thought to be relevant for retrospection—such as the performance of the economy, crime, and the performance of schools, to name a few—valid criteria for evaluating government performance? That is, do political coalitions actually have the power to influence the performance metrics used for retrospection on the timeline introduced by elections? In this paper, we use difference-in-difference and regression discontinuity techniques to explore whether states governed by Democrats or those governed by Republicans offer better returns on economic, education, crime, family, social, environmental, and health outcomes on the timeline introduced by elections (2 and 4 years downstream). We find that states controlled by Democrats perform equally to states controlled by Republicans. Our results suggest that voters may struggle to truly hold government coalitions accountable, as objective performance metrics appear to be largely out of the immediate control of political coalitions.

Abstract Word Count: 150

Document Word Count: 11,394

*We wish to thank Dan Butler, Daniel Carpenter, Alexander Coppock, Albert Fang, Vladimir Kogan, Stéphane Lavertu, Matthew Lebo, Cecilia Mo, Jerome Schafer, and Christopher Warshaw for their invaluable feedback; Matt Grossman, Marty Jordan, James J. Feigenbaum, Alexander Fouirnaies, and Andrew B. Hall for sharing data; and the participants of the American Political Science Association (APSA), Association for Public Policy Analysis and Management (APPAM), Midwest Political Science Association (MPSA), and Brigham Young University for their contributions to this project.

"Proponents of retrospective voting have simply assumed that there are real, persistent differences in ... competence between competing teams of political elites."

–Achen and Bartels (2016, 158)

Whether citizens are able to hold government officials accountable is a fundamental question for democracy. Indeed, theories of political accountability argue that citizen retrospection—the capacity of citizens to electorally punish and reward policymakers based on performance—is vital for democratic well-being and prosperity (Fearon 1999; Ferejohn 1986; Grant and Keohane 2005). As a result of the importance of retrospective behavior, an abundant and ever growing literature explores whether citizens respond when performance deteriorates. In recent years, scholars have been critical of citizens' capacity to fulfill their retrospective duty given their propensity to react to forces truly orthogonal to the control of politicians—such as sporting events (Healy, Malhotra and Mo 2010; Busby, Druckman and Fredendall 2017), shark attacks (Achen and Bartels 2016), natural disasters (Healy and Malhotra 2009), and policy decisions outside of the realm of a given official (Sances 2017).

Others, however, take hope in the fact that voters also respond to metrics (seemingly) more directly in the control of elected officials. For example, there is evidence that citizens respond to dips in economic performance (Achen and Bartels 2016; Fiorina 1978; Healy and Lenz 2014, 2017)[1], spikes in crime (Arnold and Carnes 2012; Bateson 2012), increases in military deaths (Grose and Oppenheimer 2007), decreases in school performance (Berry and Howell 2007; Holbein 2016; Kogan, Lavertu and Peskowitz 2016), and changes to distributive spending (Chen 2013), to name a few. While the literature is somewhat mixed—with many studies showing signs of retrospective behavior, but others offering a moderating voice—an underlying normative assumption is that even if retrospective voting were not happening with these objective performance metrics, public welfare would increase if it were (Key 1966; Kramer 1971; Fiorina 1981).

Much of the literature takes as given that policy outcomes—like crime rates or the performance of the economy and schools—are appropriate measures of elected officials' performance. But, are they? In this paper, we re-examine whether elected officials are able to truly influence the metrics by which some scholars argue they are (or, at minimum, should be) judged at elections.

[1]But, see also Hall, Yoder and Karandikar (2017).

That is, we flip the direction of causality; instead of asking, "does performance affect political control?" (as the retrospective voting literature does), we ask, "how much does political control affect performance?" While having received some sporadic attention (e.g. Achen and Bartels 2016; Bartels 2009; Wolfers 2002), we bring this question into full focus. Our objective is to explore the extent to which political party control influences barometers of performance in the short-term, focusing on 2 to 4 years after changes in party control[2] since this is the timeline introduced by elections—the primary vehicle by which candidates and political parties are judged.

To do so, we estimate the effects of the party in power on a number of proxies of societal well-being. We choose to explore the role of party control, rather than of the election of individual candidates, given that changes in party control are more likely to push a specific policy agenda—and hence move policy outcomes—in a different direction (Caughey, Xu and Warshaw 2017). Our approach uses historical data from state legislatures in the United States matched to information on a host of policy outcomes across six different sectors measuring economic, education, crime, social, environmental, and health/family outcomes. With these data, we show simple observational evidence consistent with popular wisdom and previous research—with Republican and Democratic states being quite different in the policy outcomes they both realize on the timeline introduced by elections.

To rule out the possibility that these simple patterns are not a reflection of other broader social forces, we use difference-in-difference and regression discontinuity models that leverage changes in party control and scenarios where one party holds marginal control. These show that observational comparisons (and popular narratives) paint an oversimplified picture of the effects of party control. From these, it is clear that the party in power has almost no effects on economic, health/family, educational, crime, civic, and environmental outcomes on the 2 to 4 year timeline introduced by elections. This lack of an effect holds true through a host of robustness checks. Moreover, it persists regardless of whether government is split or unified, whether the party has persistent power over time, and regardless of the bundle of policy outcomes we explore. Simply put, there do not appear to be causal differences in policy outcomes over which retrospective vot-

[2]In the Online Appendix, we find that our results hold even when we examine policy effects up to T+8 years out.

ing does (or, according to some, should) occur between states controlled by either the Republican or Democratic party in the short-term.[3]

Our results make several important contributions. While political parties are often judged by immediate performance on the policy outcomes in the window that we explore, our work suggests that party majorities may actually have little control over these performance metrics in this timeframe. Thus, voters who retrospectively vote political parties out of office based on the performance of the economy, schools, or crime, may actually be simply responding to noise or factors that are largely out of the short-term control of elected officials. This presents an important new wrinkle for understanding democratic accountability (at least at the state level) and calls into question the quality of retrospective voting on the current, shortened timeline introduced by elections. This implies that in order for citizens to truly hold elected officials accountable, there may have to be more time allocated between when one party is in power and the time when accountability decisions are made. Alternatively, our findings suggest that elected officials may be best judged by prospective judgments or other performance metrics, both of which place higher demands on voters in terms of political knowledge. For example, it may be more useful for elected officials to be judged by the policies they pass, rather than what comes out as a result—an approach that is currently underutilized by voters in state legislative elections (Rogers 2017). In lieu of either of these potential solutions, the immediate effects of political parties on policy outcomes is highly relevant.

1 Background and Conceptual Framework

Retrospective voting involves "an attempt by voters to select the best available team of political leaders" (Achen and Bartels 2016, 98). But how should voters choose the best team or party to lead? Among political scientists, there is a general view that certain outcomes are good candidates for retrospection and others are not. The criteria for judging a useful metric has tended to revolve around whether that metric was directly influenced by an elected official or was "clearly beyond

[3]This does not necessarily mean that party control cannot result in changes in policy outcomes. Rather, among U.S. state governments, we do not find evidence of a systematic, partisan effect.

the leaders' control" (Achen and Bartels 2016, 142). Hence, scholars are concerned when voters punish elected officials for undesired sporting outcomes (Healy, Malhotra and Mo 2010; Busby, Druckman and Fredendall 2017)[4] the weather (Gomez, Hansford and Krause 2007; Henderson and Brooks 2016), or other events (including those listed above) that are orthogonal to politicians' control. In contrast, scholars judge a democracy to be well-functioning when voters respond to policy outcomes directly influenced by the policy decisions made by elected officials (Healy and Malhotra 2013). Because of their direct connection to policy outputs or levers within the control of elected officials, the overall performance of the economy, schools, criminal justice, health, and environment are thought to serve as useful metrics for incumbents' performance.

Social scientists from various disciplines and subfields have long argued that political parties influence the overall health of the economy and other policy outcomes relevant to societal well–being (e.g., in American politics: Bartels 2009, Faricy 2011, and Hacker and Pierson 2010; in comparative politics: Alvarez, Garrett and Lange 1991 and Alesina and Roubini 1992; and in political economy more generally: Hibbs 1977, Chappell and Keech 1986, and Alesina and Rosenthal 1995). As Nordhaus put it, "All... aspects of our economic life, and many more, are influenced by government policies" (Nordhaus 1975, 169). Hacker and Pierson (2016) similarly argue: "we should be worried, whatever our partisan tilt, that leading conservatives promote a ... model so disconnected from the true sources of prosperity."[5] Interest in the effect of party control on policy outcomes is not restricted to academics; citizens, the media, interest groups, and various other vested organizations make key decisions based on which party they believe promotes better economic, education, crime, family, social, environmental, and health outcomes. Indeed, the question of which party produces better policy outcomes is often the central issue of campaigns and political discourse more generally. In this way, the effect of political parties on policy outcomes is a foundational question to the health of democracy.

Despite its importance, empirical research on the effect of party control on policy outcomes has been sparse. As Achen and Bartels (2016, 158) note:

[4]But, see also Fowler and Montagnes (2015).
[5]See "The Path to Prosperity is Blue"; New York Times Op-Ed; July 30, 2016.

> [Retrospective voting] requires us to assume that there are real differences in economic competence between competing political teams...However, in light of the crucial importance of this assumption for the whole notion of retrospective selection, it is striking that it has never, as far as we know, been subjected to any systematic empirical examination. Proponents of retrospective voting have simply *assumed* that there are real, persistent differences in economic competence between competing teams of political elites.

Achen and Bartels (2016) go on to explore the observational relationship between the political party, or the "team of political elites," in power and the performance of the economy. They argue that political coalitions may have *too much* control of policy outcomes—being able to precisely manipulate these through quick-fix policy solutions in the lead up to elections that lead to immediate (but short-lived) downstream improvements (Achen and Bartels 2016, Ch. 5 & 6). This result is consistent with earlier work done by Bartels (2009, ch. 2), who argues that economic performance "in contemporary America is profoundly shaped by partisan politics" [34].

However, Achen and Bartels (2016) leave us wanting in several areas. First, Achen and Bartels (2016) and Bartels (2009) only focus on the performance of the economy. In neglecting the effect of political parties on other relevant policy outcomes, they ignore the multitude of policy domains where party control could have a meaningful influence and where retrospective voting is occurring. Second, their identification strategy makes strong assumptions about the distribution of unobservable characteristics that might bias the relationship of interest. Their approach uses times series data from the federal level and merely conditions on a few observable characteristics. As a result, we are left not knowing how robust the strong associations between party control and the economy they uncover are to more stringent specifications. Despite strong language to the contrary, this identification strategy does not approach causality given the multitude of complex factors that likely co-determine party control and policy outcomes. This is consistent with the other limited work on this topic: causal studies of the effect of parties on policy outcomes are lacking.

1.1 Timeline of Effects

Theories linking party control and policy outcomes often assume a long time horizon. That is, that political parties in power will have many years for their policy agendas to develop and for their policy outcomes to mature (Nordhaus 1975). However, when considering the retrospective implications of party control for policy outcomes, the frequent nature of elections complicates this calculation. In order for democracy to function—the story goes—citizens need to be able to respond to politically generated policy outcomes during elections. Thus, while political coalitions may have long-term effects, political parties are judged on a short-term timeline of just two or four years at nearly all levels of government in the U.S.

There are both reasons to suspect and to be skeptical that party control influences policy outcomes on the timetable introduced by elections. On the one hand, as we outline below, several studies have shown that party control moves the policy outputs—i.e., the legislation passed and implemented by government—that predate policy outcomes over the short-run (Caughey, Xu and Warshaw 2017). Moreover, previous policy evaluations have shown generally that changes in these policy domains can have short-term effects; and some, sporadic, analyses of the effect of party control on policy outcomes have shown immediate effects. On the other hand, however, some legislation may only have long-term effects. There are also a multitude of other forces in play that may swamp the influence of the party in power. Though scholars argue that citizens can, do, and should judge elected officials based on policy outcomes, this type of retrospective behavior may actually not be a desirable feature of electoral systems. If elected officials can't effectively move policy outcomes on the timeline introduced by elections, citizens may simply be responding to factors outside the control of these representatives. Under this scenario, retrospective responses to objective performance metrics may not actually be a savior to democratic systems (perhaps) plagued by seemingly irrational retrospective responses (Achen and Bartels 2016). They may be, in fact, just as bad.

Ex-ante, it is not clear which of the two scenarios we find ourselves: with parties either being in control of the objective performance metrics by which they are judged or not. However, this much is clear: in order for party control of branches of state government to affect *policy outcomes*, two conditions must hold. First, party control must influence *policy outputs*—i.e., the legislation passed and implemented by government—and, second, those different outputs must also have different effects on policy outcomes on the timeline introduced by elections.

1.2 Past Work on Party Effects and Policy *Outputs*

The extent to which the Republican and Democratic parties implement distinct agendas at the state level, given voter preferences and economic conditions, has been an open question in the literature. Basic spatial models of partisan competition predict that the parties' platforms should converge on the preferences of the median voter on each issue. At the same time, if elected officials are policy-motivated and cannot perfectly predict constituents' preferences, then candidates should diverge somewhat in their policy preferences. Consistent with this latter view, Shor and McCarty (2011) show that state legislatures have become more ideologically polarized over time.[6]

Scholars of state politics have long hypothesized that Democratic state legislatures and governors since the New Deal era should pursue more liberal policies than Republican ones, especially in terms of redistribution (Dye 1966; Jennings 1979; Erikson, Wright and McIver 1989). Despite these theoretical expectations, the empirical record provides mixed evidence of the effect of partisan control on state policy outputs (e.g., Besley and Case 2003; Lax and Phillips 2012).

Caughey, Xu and Warshaw (2017) stand out for its strong causal identification strategy. Like the current study, Caughey, Xu and Warshaw (2017) use difference-in-difference and regression discontinuity models to isolate the effect of party control of a state legislative chamber/governorship on the legislative agenda implemented at the state level. In addition, the authors use a sophisticated measure of the overall ideological bent of the policies passed each year at the state level across 148 policies examined from 1936 through 2014 (see Caughey and Warshaw (2015) for more details on this measure). This approach allows Caughey, Xu and Warshaw (2017) to examine the effect of partisan control on the overall tilt of a state's policy agenda[7] and avoid the file drawer

[6]See Grofman (2004) for a review of this literature and Broockman and Skovron (2015) for evidence of state legislators' misperceptions of constituent opinion.

[7]At the same time, this approach might also give more weight to less consequential or less salient policies where partisan differences are less meaningful (a fact Caughey, Xu and Warshaw (2017) readily acknowledge).

problem of examining only one or a few policy outputs at a time.[8]

In their analysis, Caughey, Xu and Warshaw (2017) find that when Democrats win a legislative chamber or the governor's office, policy *immediately* moves in a more liberal direction overall.[9] While the effects are modest, they are statistically significant.[10] This work suggests that the majority party, even if they are only in control marginally for a short period of time, is able to implement a more liberal (in the case of Democrats) or conservative (in the case of Republicans) policy agenda even when controlling for all other factors that could outweigh the effects of partisan majority control such as public opinion, economic conditions, and federal policies (Rose 1973; Winters 1976; Kemp 1978). This separation of parties on policy outputs is consistent with studies from local contexts.[11] Simply put, party control seems to change policy outputs over the short term—well within the window introduced by elections.

1.3 Past Work on Party Control and Policy *Outcomes*

But do these partisan changes in policy outputs trickle down to the policy outcomes of critical import? Most work on the effects of partisan control on policy outcomes examines national governments and macroeconomic outcomes. This literature comes to the conclusion that party control plays a large role on policy outcomes realized. For example, Bartels (2009, ch. 2) highly visible work studies the role of the party in power on economic inequality at the national level, finding that Republican presidents are associated with higher degrees of economic inequality. This result

[8] The "file drawer" problem refers to a bias that can occur in a literature due to researchers' publishing statistically significant findings but not null findings. This prevents the literature from identifying spurious findings. If research on the effects of party control only ever focuses on a single policy output or outcome, it could be possible that the literature includes many spurious findings, overstating the effects of party control.

[9] This effect primarily exists in the final two decades of their dataset, from 1990 through 2014. This likely explains why earlier work on this topic often failed to find an effect.

[10] Caughey, Xu and Warshaw (2017) estimate that Democrats' winning a majority in a legislative chamber will increase per capita welfare benefits by about $12 to $24 annually and increase the number of liberal policies passed by about a half percentage point.

[11] For example, Gerber and Hopkins (2011) find that Republican mayors spend more on policing but do not affect a host of other outputs. de Benedictis-Kessner and Warshaw (2016) find that Democratic mayors, at least in larger cities and over a longer time series, spend more and issue more debt to pay for it.

is consistent with a broader literature, beginning with Hibbs (1977), that argues that left-leaning parties (such as Democrats in the U.S. case) represent low-income voters and therefore are more likely to pursue macroeconomic policies that raise inflation and lower unemployment. Indeed, a meta-analysis on the massive literature examining Hibbs' theory finds evidence, overall, that partisan control is associated with macroeconomic outcomes (Franzese 2002).

However, the vast majority of this work simply relates party control with policy outcomes, perhaps conditioning on a few observable characteristics. However, the party in power and policy outcomes are affected by a multitude of factors, many of them unobserved. Given this, we argue that the current literature exploring the role of party control for policy outcomes does not use empirical methods that can isolate a party effects from other factors. Our objective in this paper is to significantly improve upon previous research on this dimension.

In so doing, we choose to focus on the role of party control at the (U.S.) state level. We think this is justified for three important reasons. First, doing so provides us with a greater degree of statistical power. Because there are more state units than federal units over time, this provides with a greater degree of precision to detect policy effects if they are, indeed, present. Second, this approach allows for better causal identification. Partially because of the larger sample size, at the state level we are able to better isolate exogenous variation in party control at the state level. States also give us a source of variation beyond time, thus allowing for stronger comparisons than are allowed with simple time series. In this way, states offer us a more robust counterfactual; one that is not present at the Federal level. Third, states are a vitally important arena for public policy. Because of its federalist system, the U.S. leaves much of its policymaking power to the states. States control many policy levers that influence the policy outcomes we study—determining policy inputs like tax rates, education spending, rules governing healthcare, policing policies, and environmental regulations, to name a few. Echoing many scholars of state politics, Rogers (2017) concludes that "State legislators have considerable authority over American lives. They determine who has the opportunity to vote, go to college, and even get married, and elections are the primary instrument by which citizens can exert control over those who govern them" (570).

Despite states' important role in the policy process, very little previous work has examined the effects of party control of state government on policy outcomes. The studies that have been done focus on either gubernatorial partisanship or only a small set of policy outcomes. For example, Cummins (2011) finds that Republican governors increase health insurance coverage more than Democratic ones. Yates and Fording (2005) find that Republican state legislatures are associated with a larger disparity between the imprisonment rates of blacks over whites while Republican governors are associated with higher imprisonment rates across both races. Leigh (2008) provides the most extensive and methodologically robust analysis of the effects of gubernatorial partisanship on policy outputs and outcomes. Using a regression discontinuity design, Leigh (2008) fails to find evidence that the partisanship of the governor matters for 32 outputs and outcomes, including measures of income and employment. However, Leigh (2008) does not consider the effects of party control of legislatures—a critical body for policymaking, perhaps much more so than governors. As such, it remains an open question as to whether Democratic or Republican controlled states produce better policy outcomes for their citizens.

Given Leigh's (2008) results and the finding that a legislature's partisanship has a modest effect on policy outputs (Caughey et al. 2016), we might not expect to find that the party in control of the legislature and governorship matters for policy outcomes. However, Leigh's (2008) lack of findings could be due to governors' wielding less influence over legislative outputs than legislatures. In addition, Caughey, Xu and Warshaw's (2017) analysis of an aggregate measure of legislations' liberalism could hide a partisan effect that is pronounced on a few key issues that have large effects on policy outcomes. Finally, there is strong evidence using a regression discontinuity design that partisan control affects policy outcomes in other countries outside the United States. For example, in Swedish municipalities Pettersson-Lidbom (2008)—that share many similarities to U.S. states[12]—Pettersson-Lidbom (2008) finds that party has a sizable impact on unemployment over the short-term. These findings suggest that party control could truly matter for policy outcomes over the timeline introduced by elections, but leave many questions unanswered.

[12]Like U.S. states, Swedish municipalities have a right to self-government and no limits on borrowing. They also freely set their tax rates. In addition, they account for 20-25% of spending and provide many of Sweden's social programs, including education, elderly care, and welfare. They also operate in a federal system and are affected by fiscal and macroeconomic policies set at the national level.

2 Data

In this paper, our independent variables of interest capture whether state legislative bodies and the governorship are controlled by Democrats.[13] This data comes from the the Correlates of State Policy Project Database (CSPPD). The CSPPD is housed at the Institute for Public Policy and Social Research (IPPSR) at Michigan State University (Jordan and Grossmann 2016). This data provides the party in power for both chambers in all states from 1960–2010. Importantly for our identification strategy (outlined in the next section), it also provides a measure of how close the state is to being controlled by Democrats/Republicans. We quantify this as the fraction of seats needed to take the majority in the legislature and the percent above or below the next closest candidate for the governorship (though in robustness checks we also use the alternate running variables suggested by (Hall, Feigenbaum and Fouirnaies Forthcoming)).

Figure 1 shows the distribution of the running variables across states over the time period of study. Points marked in blue are states controlled by Democrats, and points colored in red are states controlled by Republicans. Each of the figures also has a local non-parametric regression model superimposed on the figure—showing the trend towards Republican domination of state legislatures and governorships in recent years.

Our dependent variables capture states' performance on economic, health/family, civic, crime, education, and environment outcomes. In examining the effect of party control on policy outcomes, we are trying to strike a delicate balance. On the one hand, our objective is to be as thorough as possible to avoid any potential "file drawer" problems that could result by examining only a few policy outcomes. At the same time, however, we do not wish to introduce policy outcomes that are irrelevant to the party in power. This might skew our conclusions in the opposite direction and mute party control's effects, making them seem to be more noisy than they really are. As such, we have sought very carefully to focus on outcomes that could plausibly be linked to changes made by party coalitions. This approach balances our desire to be thorough in our anal-

[13]Because most of the chambers and gubernatorial offices that we study are either controlled by Republicans or Democrats, the effect estimates are approximately the inverse of the Republican party control effect.

Figure 1: Proximity to Democratic Majorities Over Time

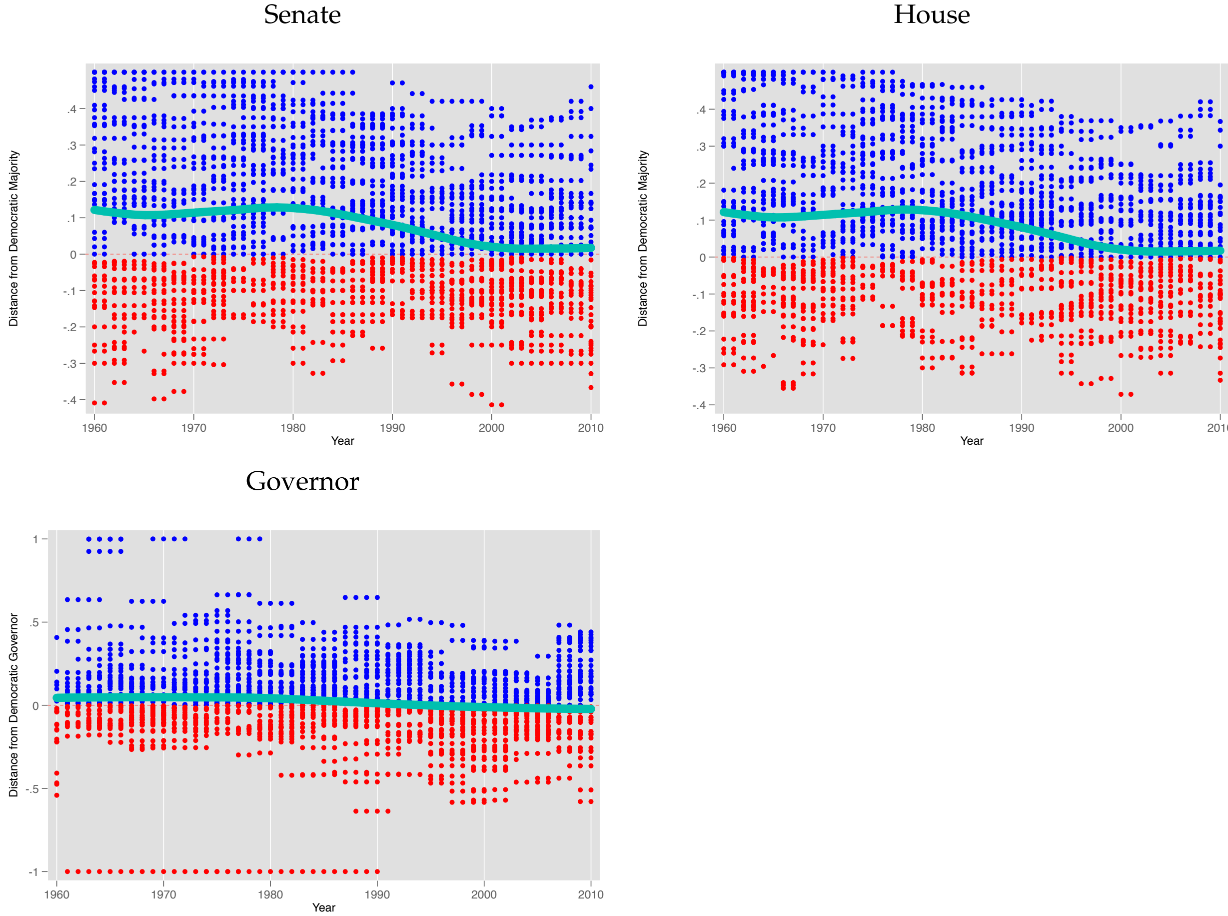


Figure 1 plots the proximity to Democratic control of the two legislative chambers and the governorship. Each point represents a state–year observation. Blue points represent states controlled by Democrats; whereas red points represent states controlled by Republicans. The grey line overlaid plots a local non-parametric regression.

ysis with our recognition that not all policy outcomes are theoretically connected to party control. Even if the reader is skeptical of a few of our outcomes and we set aside these, the story that we would tell would remain the same. As we show below, the results (null effects) are remarkably consistent across outcome types.

In this paper, we examine 27 outcomes capturing states' economic, health/family, civic, criminal, educational, and environmental well-being. The variables that we use vary in terms of their availability over time (see Table A1 in the Online Appendix). The economic outcomes available in the CSPPD include standard measures of average income (real, per-capita), inflation (CPI), unemployment, growth (GSP), quarterly housing prices, population growth, the performance of the agriculture sector, and as measures of economic inequality—the fraction of income held by the top 1% and 0.1%. The available health/family outcomes include measures of health spending (per capita), the number of new immigrants, the abortion rate, divorce rate, and the birth rate. The civic outcomes include voter turnout and the number of felons ineligible to vote. Crime-related outcomes include measures of the auto theft, murder, property crime, rape, robbery, and violent crime rates. Education measures available in the CSPPD include school attendance and graduation rates.[14] Finally, our environmental outcomes capture the amount of CO2 emissions, energy usage by the business sector, and the price of residential energy.[15] These provide a thorough picture of the effect of party control on the multiple dimensions of societal well-being.[16]

As we mention above, our primary interest is in examining the effects of party control on the timeline introduced by elections. As a result, we place our primary focus on the our 27 policy outcomes only up to four years downstream. (For space considerations, we plot the results from 2 and 4 years downstream in the paper. As we show in the Online Appendix, the results from one and three years downstream lead to the same conclusion.) A potential criticism of this research design is that state legislative/gubernatorial terms (usually 2 years in length, but 4 years, at most, in a few state chambers) do not provide sufficient time for policies passed under a given regime to have noticeable effects on policy outcomes. In our view, this is not a criticism of our work; but of retrospective voting systems built on policy outcome performance metrics and short election timetables. Indeed, we agree that it could feasibly be the case that party effects need time to mature. However, as we discussed at the beginning of the paper, if effects are not observed on the timetable of elections, then we might seriously question whether retrospective voting based on noisy policy outcomes is desirable. Before policy effects arrive, electoral decisions based on

[14] Standardized test scores, like those from the NAEP, are available only sporadically, making them hard to use in our analyses below.

[15] Included in table A1 in the Online Appendix is a list of the years for which each of these outcomes is available in the CSPPD files.

[16] On many of the outcomes we explore—such as individual income—the normative implications for directional changes is clear. Indeed, most would agree that increases in income are normatively desirable, and, likewise, decrease undesirable. For other outcomes, however—such as the abortion rate—the normative implications are less clear. With these measures, what is clear is that they are often used as policy outcomes regardless of the normative implications.

performance on policy outcomes have to be and are often made. Hence, the short-term (non) effects of party control are highly relevant to the quality of the democratic system.

Moreover, this criticism is muted by two other factors. First, in the Online Appendix we set aside our election timeline and also examine outcomes on a much longer timetable (up toT+8 years downstream).[17] Second, we note that though it is likely true that some effects may take a while, there are many policies that we should expect to have immediate effects. Indeed, in selecting our dependent variables, we chose outcomes known to be malleable to short-term forces, such as a liberal or conservative legislative agenda. For example, examining effects one year down stream is the norm in education (e.g., Chiang 2009; Angrist, Pathak and Walters 2013), health (e.g., Finkelstein et al. 2012; Newhouse and Group 1993), social welfare (e.g., Jardim et al. 2017), and criminal justice (e.g., Yokum, Ravishankar and Coppock 2017; Ludwig and Cook 2000), to name a few. State-level agencies and legislative analysts regularly predict that changes in fiscal and other policies will have short-term effects, which our research design should pick up. Further, as mentioned earlier, Pettersson-Lidbom (2008) find evidence of party control effects in Swedish municipalities on a limited number of policy outcomes, which have 3 year terms and play similar role as states within their federal system. Together these results suggest that policy effects can take place in the window that we use. Based on all of this, it is not totally unlikely that we might observe party effects on the timeline we use.

2.1 Naive Comparisons

Before outlining our identification strategy, we think it useful to discuss what purely observational models show. While not causal, this comparison gives us a way to benchmark to those frequently made in previous research at the federal level and to public discussion about the performance of the two political parties. That is, if voters are, indeed, engaging in retrospective voting, this is how they are doing it: making simple comparisons across the two parties. This helps us to see what voters are seeing and to make our case for a need of a compelling causal identification strategy.

At first blush, the empirics tend to support the idea that the party in power affects short-term

[17] We do not use this as our main results as the data restrict our ability to satisfactorily model dynamic party effects across party control transitions that occur across elections.

societal well–being. Figure 2 displays the observational relationship between party control and several policy outcomes, controlling only for the party in control in the other respective units. As can be seen, 33 out of the 48 policy outcomes show some signs of a meaningful party difference. Turning to the bottom right panel, we can see that state observations with unified Democratic governments have lower income, higher unemployment, lower voter turnout, higher murder rates, and lower graduation rates.[18] In short, the observational patterns here suggest that there may be something important about the political party in power in determining policy outcomes.

Of course, these observational patterns do not mean that the party in control *caused* these outcomes. Yet still, these simple trends have significance in the real world. Abundant research has shown that citizens make their voting decisions based on economic (Fiorina 1978; Bagues and Esteve-Volart 2013), education (Berry and Howell 2007; Holbein 2016; Kogan, Lavertu and Peskowitz 2016), crime (Arnold and Carnes 2012), and environmental (Stokes 2015) performance metrics. Hence, different patterns in the performance of policy outcomes may feed into future electoral results.

But therein lies one of the fundamental problems with observational relationships between the party in power and policy outcomes that have dominated previous research. The likely presence of retrospective voting implies endogeneity: muddying the ability to draw a clear causal conclusion over which party is producing better policy outcomes. That is, citizens may choose the party in power *because of* their perceptions of the overall performance of the state. If this type of retrospective voting occurs—and a large literature suggests that it often does—any simple relationship between the party in power and policy outcomes may simply be endogenous. Hence, in order to fully understand whether the party in power causally affects policy outcomes of substance, a

[18]States with lower chambers controlled by Democrats have, on average, noticeably higher unemployment, lower birth rates, lower voter turnout, more felons ineligible to vote, higher CO2 emissions, higher crime, and lower high school graduation rates than those controlled by Republicans. States with upper chambers controlled by Democrats have lower income, birth rates, felons ineligible to vote, lower CO2 emissions, robbery rates, and graduation rates, and have higher commercial energy consumption, murder rates, and school attendance. States with Democratic governors have descriptively lower income, birth rates, felons ineligible to vote, CO2 emissions, robbery rates, graduation rates, and attendance, and have higher unemployment, commercial energy consumption, and murder rates.

Figure 2: Simple Relationship Between Democratic Control & Policy Outcomes

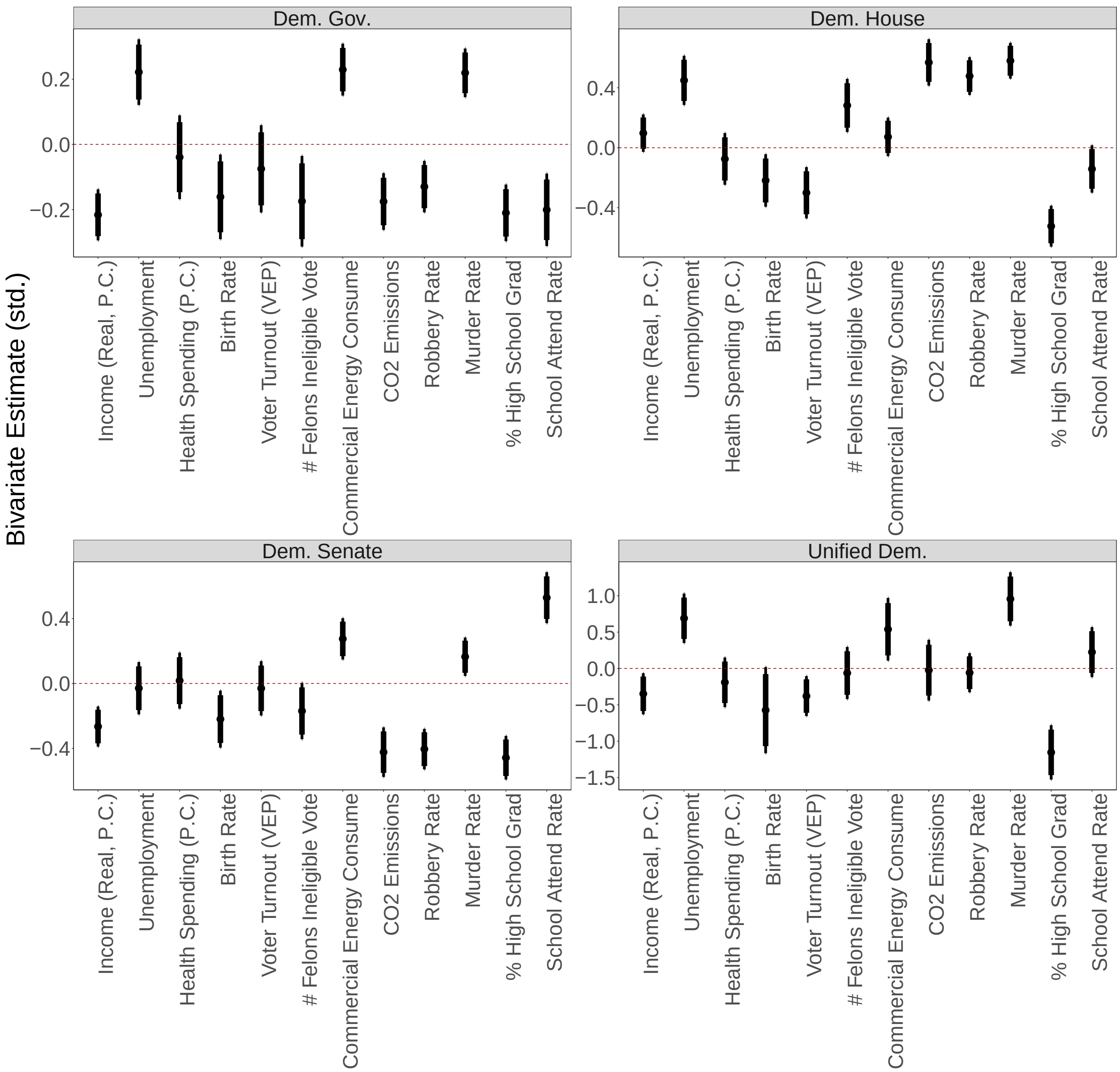


Figure 2 displays coefficient plots of the simple estimates between party control in the three bodies (upper, lower, governor) and 12 select policy outcomes in the first year after a given party is in power. Point estimates are shown with dots and 90/95% confidence intervals with bars. The outcomes are standardized simply to allow for a similar scale in the figure.

causal identification strategy is required.

3 Methods

Clearly, randomly assigning which party is in power is not feasible. Hence, we are forced to make quasi experimental comparisons. Here we use two complementary comparisons that allow us to approach causality as much as possible. The first uses the panel nature of our data to estimate difference-in-differences models. The second uses naturally–occurring, as-good-as random assignment of party control to estimate regression discontinuity models.

Our difference-in-difference models consist of a relatively straightforward estimation of a model that includes state and year fixed effects and flexible time trends for each state. These absorb all observed and unobserved factors that remain constant within states (e.g. political culture, social capital, rigid political institutions, etc.) and are shared within certain years (e.g. recessions, campaigns, etc.), and trends that vary across states (e.g. the natural trajectory on our outcomes). To mirror previous research on the effects of party control (Caughey, Xu and Warshaw 2017), in some difference-in-difference specifications we also include lagged measures of our independent and dependent variables—but, this does not change our results.[19] The inclusion of state-specific time trends allows us to relax the, sometimes tenuous, parallel trends assumption key to difference-in-difference specifications. Here our identifying assumption is that our outcomes in states separate by treatment deviate from common year effects by following the linear trend captured by the interaction term. Under this assumption, identification comes from sharp deviations from otherwise smooth state-specific trends. The assumption behind this approach is inherently untestable, although it is considered to be stronger than one required in a model with just state and year fixed effects (Angrist and Pischke 2008, 2014). While this approach may be slightly less internally valid than our second approach (the RDD), it offers the advantage of more statistical power and, perhaps, of having broader generalizability.

To complement this first approach, we estimate models that go as far as possible towards

[19]We do not use these as our primary results as, in general, including lagged dependent variables in difference in difference models is not advised because of the strong assumption one must make about the relationship between the errors over periods (Angrist and Pischke 2008). In this case, this appears to have little influence on the outcomes—we still get null results with lagged measures of the dependent variables included.

causality within our constrained quasi-experimental setup. This approach leverages exogenous variation around the cutoffs for which party is in power of the two legislative chambers and the governorship to estimate regression discontinuity models. Under a regression discontinuity framework, observations that are sufficiently close to an arbitrary discontinuity are separated primarily by exogenous shocks (Butler and Butler 2006; Imbens and Lemieux 2008; Lee 2008; Lee and Lemieux 2010; Lemieux and Milligan 2008). Regression discontinuity models benefit from continuity in potential outcomes around the cutoff. Put differently, this approach uses data on either side of the cutoff to establish treatment and control groups that are similar on observables and unobservables. Given modest assumptions, RDD models produce unbiased local average treatment effects that benchmark well with causal estimates from randomized–control trials (Buddelmeyer and Skoufias 2004; Green et al. 2009; Lee and Lemieux 2010). While this is a great advantage, it does come at the cost of reduced statistical power and, perhaps, in generalizability. Given that neither of our approaches is perfect, we rely on both the difference-in-difference and the RDD models below to make our conclusions. Ultimately, we are reassured that both give us the same answer.

The specific party control cutoff allows us to estimate the causal effect of legislatures and governorships being controlled by Democrats compared to those controlled by Republicans. This approach is similar to the increasingly common electoral discontinuities used by political scientists that come from close-race elections (e.g., Butler 2009; Eggers et al. 2015; Hainmueller and Kern 2008; Hall 2015; Lee 2008). Rather than being about individual races, however, this discontinuity leverages differences in aggregated party control.[20]

Following previous work estimating the effects of party control on policy outcomes (Caughey, Xu and Warshaw 2017), we start by estimating a standard regression discontinuity models for each of three cutoffs (upper chamber, lower chamber, and governorship) individually. These models take the form shown in Equation (1).

[20] Hall, Feigenbaum and Fouirnaies (Forthcoming) argue that the party control cutoff can be conceptualized as a multidimensional version of the close elections cutoff. We use this approach in robustness checks, but not as our primary specification as we are more optimistic (given the results form a host of additional balance checks we provide in the Online Appendix).

$$O_{st} = \beta_0 + \beta_1 D_{st} + g(P_{st}) + \epsilon_{st} \quad (1)$$

In Equation (1), the unit of observation is the state–year. The policy outcomes are denoted by O_{st}. These outcomes were outlined in the last section and include 27 commonly examined policy outcomes that measure states' economic, health/family, civic, criminal, educational, and environmental performance in the first and second years after a given election. The key input (D_{st}) describes which party is in power in a given year for that respective body. In this case, the variable takes the value of 1 when a state legislature or governorship was controlled by Democrats and 0 otherwise. The key parameter of interest is β_1, which shows the causal effect of a Democratic legislative body or governorship on our policy outcomes. The running variable—how close the state was to being controlled by a democratic majority legislative body—is denoted as P_{st}. In our RDD models, we specify this with the proportion of seats above the party control threshold for Democrats. However, in the Online Appendix, we also use the three alternate specifications of the running variable suggested by Hall, Feigenbaum and Fouirnaies (Forthcoming).[21] In these single-dimension RDD models we specify the running variable with a local kernel-smoothed function, $g(\cdot)$, and use the optimal bandwidth suggested by Calonico, Cattaneo and Titiunik (2014). Following previous practice in this domain (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), we cluster our standard errors at the state level to allow for correlations between legislative bodies within a given state and for underestimation of standard errors when there are few clusters.

If the party power discontinuity sorts legislatures in an as-good-as random manner, model (1) will provide the causal effect of political party power in state legislatures (β_1). This estimate will be unbiased by confounders or simultaneity because legislatures fall on either side of the party control cutoff as-good-as randomly within a narrow bandwidth. To examine whether our discontinuity satisfies the implications of local randomization, we conduct the standard RDD specification checks in the Online Appendix, finding evidence that the legislative power discontinuity

[21] Their approach uses data from the individual state legislative races and chooses one or a combination of the races to proxy for how close the legislature is to Democratic control. Ibn the Online Appendix, we show that these alternative running variable specifications do not change our results.

sorting our observations in an as–good–as random manner.[22]

That said, this approach to exploring the party cutoffs in the upper chamber, lower chamber, and governorship doesn't fully capture the potential effects of unified government. Specifying a regression discontinuity model for unified party control is a bit challenging. No approach to so doing is perfect and all come with different assumptions. The complicating factor is in how to specify how close to unified control one is. With three running variables (that for the house, senate, and governor), this is by no means straightforward. We take two complementary approaches to so doing. The first, and one we focus on in the paper text, specifies proximity to democratic control with the minimum of the house, senate, or governor scores.[23] This approach follows that in education policy by Ahn and Vigdor (2014) and in political science by Holbein (2016). The logic here is that in a situation where multiple running variables determine a single treatment, the minimum score shows how far the unit has to go to either be pulled over the threshold for treatment (if it is below the treatment cutoff) or how far it has to deteriorate to fall back into the control (if it is above the treatment cutoff). This approach assumes that the three running variables move in an order-preserving manner.

Recognizing the limitations of this approach, we leverage a second technique that conceptualizes treatment as truly multi-dimensional. That is, it conceptualizes treatment as being comprised of the interaction of three treatment variables and three running variables. This approach includes all of these (and their interactions) into one multiple-regression discontinuity model. This follows the suggested approach of Papay, Willett and Murnane (2011). The one drawback of this approach is that it stretches the state-level panel to its limits in terms of common support. With a six-way interaction (and its various sub-components) the resulting MRDD models come with inflated standard errors. While neither of these approaches is perfect, together these allow us to wrap our arms

[22] While our specification checks are consistent with an RDD framework where state legislatures and governorships on the margin are sorted as-good-as randomly to either Democratic or Republican control, we add an important robustness check to make sure that our findings are robust to any imbalances at the cutoffs. These models include state and year fixed effects. These are identified based on states that switch party control at least once over the period studied. These amount to a combination of our RDD and difference-in-differences approaches. This approach absorbs constant observed and unobserved factors that may remain imbalanced around the cutoffs.

[23] For republicans it is the maximum score on the democratic running variable.

around a regression discontinuity approach for looking at the effects of unified party control on policy outcomes.

Before we proceed to our results, we again note that by design, our research question requires that we examine multiple outcomes. Scholars have long noted that when examining multiple outcomes, the rules of probability state that some estimates will be statistically significant simply by chance (e.g., Shaffer 1995; Dudoit, Shaffer and Boldrick 2003). Given the results that we present below—overwhelmingly, coefficients not statistically different from zero—this criticism is of little concern. Still, we are careful to not over–interpret the presence of some coefficients that are statistically significant by simple random chance. We make an effort to note when effects do cross standard significance levels and whether this is robust to the standard multiple hypothesis testing adjustments (i.e., Bonferroni and Sidak corrections).

4 Results

For the most part, we find that the party in power has little to no immediate effect on policy outcomes in the economy, education, environment, health/family, crime, and civic sectors. Simply put, Democrats and Republicans appear to be equal in terms of their ability to produce the policy outcomes associated with overall well–being or social prosperity on the timeline introduced by elections. This holds across all of our difference in difference and RDD setups as well as across a multitude of other robustness checks.

In the figures below, we focus on effects in the second and fourth years downstream to allow for as much time as possible for a policy agenda to be implemented and potential downstream effects on policy outcomes to be observed. In the Online Appendix can be found figures displaying effects for the other years up to T+8 years downstream. In our figures, the outcomes are standardized to allow for greater ease in visualizing the estimates, but this too does not change the substantive results.

4.1 Difference-in-Difference Estimates

We start by exploring the effect of having a single chamber switching from republican to democratic party control. These estimates find meaning given the intense interest (by citizens) and effort (from vested interest groups) placed on individual chamber switches. These estimates are shown in Figure 3. This figure focuses on the four year downstream estimates to allow for longer for effects to materialize (within the window of elections); the analogous shorter-term effects from the second year can be found in the Online Appendix (see Figure A7). As can be seen, for the most part, the effect estimates are *not* statistically significant at the unadjusted 5% significance level—77/81 (95.1%) of the coefficients estimated have p-values greater than 0.05. This is about what we would expect by chance. The four exceptions (birthrate, Senate; abortion rate, Governor; CPI, Senate; income, House) are evenly spread around the distribution centered around 0—there are two positive and two negative effects. These exceptions are all small, averaging a meager -0.6% of a standard deviation. Moreover, none of these four coefficients clear the Bonferoni or Sidak multiple comparison levels.[24] Among the remaining 77/81 coefficients that are not significant at traditional unadjusted levels, the effects are small. The average effect size is 0.5% of a standard deviation (median = 0.6%σ). Moreover, most of these coefficients are not close to being statistically significant at traditional levels, with the average p-value being 0.5 (median = 0.45). There appear to be no systematic effects by chamber, timing, type of policy outcome (be they in the civic, crime, economic, education, environmental, or health/family domains), or individual outcomes themselves. Regardless of what set of outcomes you include in an overall evaluation of the effect of party control, the story is the same: switching from one political party in control (Democrats) to the other (Republicans) has surprisingly small to non-existent causal effects on policy outcomes on the timeline introduced by elections. (Though not our primary focus, we also find that effects are not present as far as t+8 years downstream. The average effect size when we expand to this window is 0.7% of a standard deviation with an average p-value of 0.51. See Figures A8 & A9 in

[24]For Bonferroni adjustments, the critical p-value when looking at k dependent variables is p/k, which equals 0.00185 in this case. For Sidak adjustments, the critical p-value is $1 - (1 - p)^{(1/k)}$, which equals 0.00190.

the Online Appendix.[25])

[25] These longer-term results should be taken with a grain of salt as they do not fully account for the dynamics of party control across elections: something we attempt to do in our dynamic difference-in-difference models below. For this reason, we place our highest degree of confidence on estimates that fall within the timeline introduced by elections.

Figure 3: Difference in Difference Estimates of Single Chamber Changes to Democratic Control (Fourth Year)

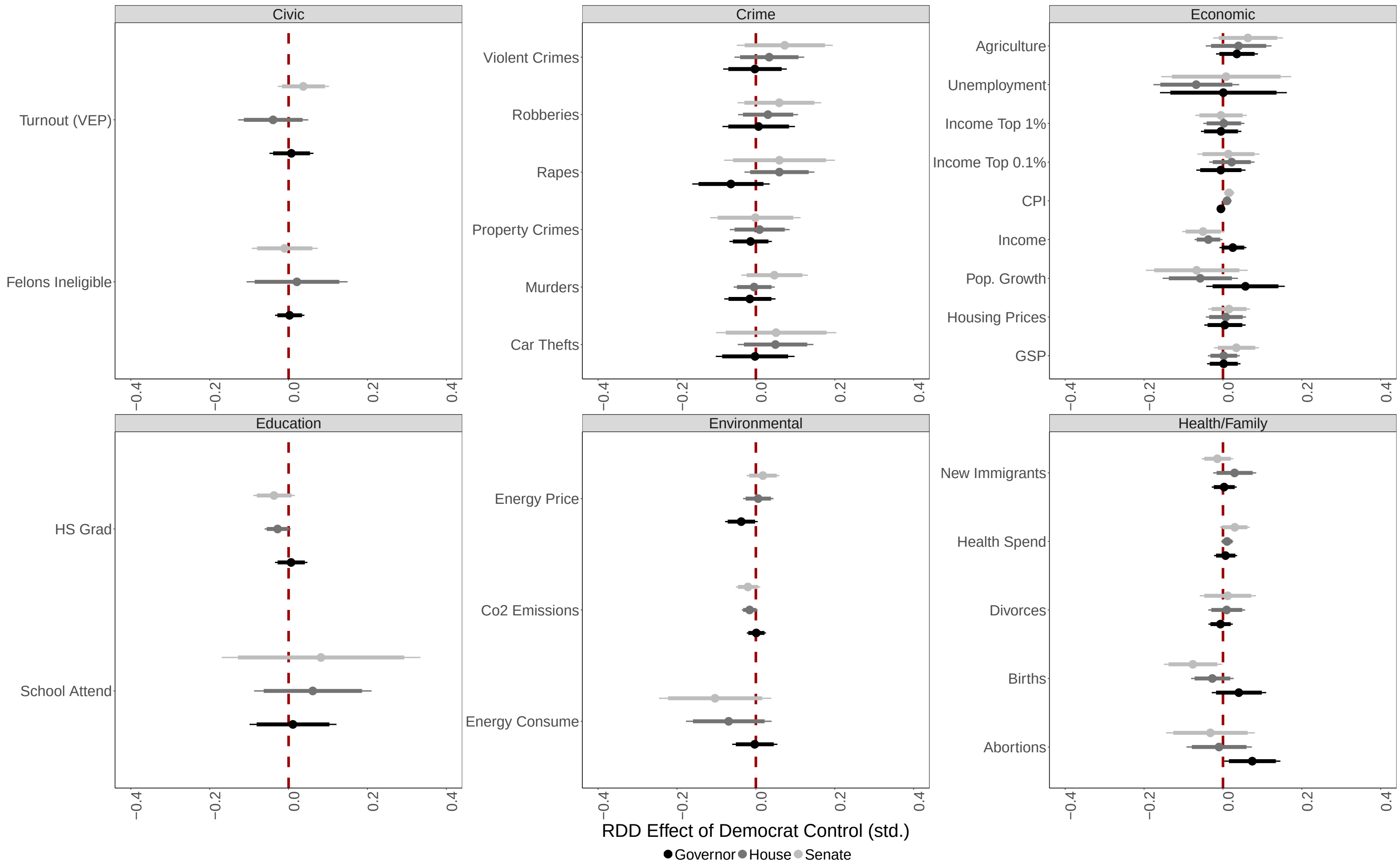


Figure 3 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for the effects of each individual chamber. Coefficients are faceted by policy area and broken by individual chamber within facets. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

We can run an additional check that uses a separate treatment measure. Instead of using a dummy variable to categorize *whether* democrats are in control, we can include the continuous running variable measure to capture *by how much* they do (or do not) have control. This allows us to see whether having more dominant control in a chamber/governor's office influences policy outcomes. When we run this check, the results remain the same. Our of the 162 coefficients estimated (27 outcomes * 3 chambers * 2 time periods) only 9 (5.5%) are statistically significant at the unadjusted 5% level, with only 1 of these 9 coming close to clearing the multiple comparisons threshold. Moreover, the coefficient estimates are small: being tightly centered around zero (mean = -.011 σ; median = -.002 σ). Simply put, holding more of a chamber (or a stronger position in the governor's chair) has no effect on policy outcomes in the window introduced by elections.

But, perhaps having *unified* party control across all three chambers is what provides parties with the ability to produce meaningful differences in policy outcomes. In practice, this is done by creating three indicators—one for whether state government is unified Democrat, the next for whether it is unified Republican, and third for whether it is divided government. Given the substantive interest between Republicans and Democrats, in the figures below we present the results for a model with the unified Republican category is the left out value. However, we present comparisons between divided government and unified Republicans and unified Democrats and divided government in the Online Appendix (see Figures A10 & A11 respectively).

Figure 4 shows our difference-in-difference estimates for the effect of unified democratic control compared to unified republican control. (For estimates that break the treatment apart by the various combinations of chamber/gubernatorial control, see the Online Appendix Figures A15–A26). As can be seen, none of the effect estimates is statistically significant at the unadjusted 5% significance level, much less at the Bonferoni or Sidak levels. In fact, the average p-value for these estimates is 0.60—far from statistical significance. As can be seen in Figure A12 in the Online Appendix, most of the unadjusted p-values are quite large; the distribution of p-values is skewed towards higher values. Moreover, our effect estimates are all small substantively. The average effect size is a paltry 0.9% of a standard deviation. Even the largest effect estimates (-0.071 σ and 0.105σ) are small. Our 95% confidence intervals are tight; allowing us to rule out effects of 30%

of a standard deviation (at absolute most), with most intervals being much tighter (our average lower bound is -5% standard deviation and our average upper bound is 6.5% of a standard deviation). In short, unified democratic governments produce policy outcomes that are statistically and substantively indistinguishable from unified republican governments on the timeline introduced by elections.[26]

Another way to see this is to look at the R^2 for our models. In models with just our party control variables, the average R^2 is a meager 0.031—meaning, we explain about 3.1% of the variance in policy outcomes. When we estimate the same models with our fixed effects and time trends, the R^2 jumps to an average of 0.90—meaning, we explain about 90% of the variance in our outcomes. This suggests that a large portion of the variation in policy outcomes can be ascribed to factors that remain constant within states over time, are shared by states within the same time period, or that vary linearly with time within states. Political parties play a minuscule role relative to these other more weighty factors, at least over the important time period of study.

Our focus here is on the short- to moderate-term effects. We do this intentionally to see if any effects emerge on the timeline of elections—the relevant barometer if one is interested in whether policy outcomes are useful separating signals for voters. However, some may still be interested to see whether effects materialize further down the road. If we expand our vision to longer-term effects, the results remain the same. When we estimate the same models from one year to eight years downstream, the results remain the same—the mean effect size remains about 0.9% of a standard deviation, with an average p-value of 0.60 (see Figure A12 in the Online Appendix). Moreover, there do not appear to be effects that emerge only after the longer-term. Simply put, according to our difference-in-difference models, there does not seem to be large causal differences on policy outcomes between the two parties. All of this implies that political control plays a small, virtually indistinguishable, role on the timeline introduced by elections.

[26]These results do not change if we use population weights.

Figure 4: Difference in Difference Estimates of Unified Democratic Control Compared to Unified Republican Control

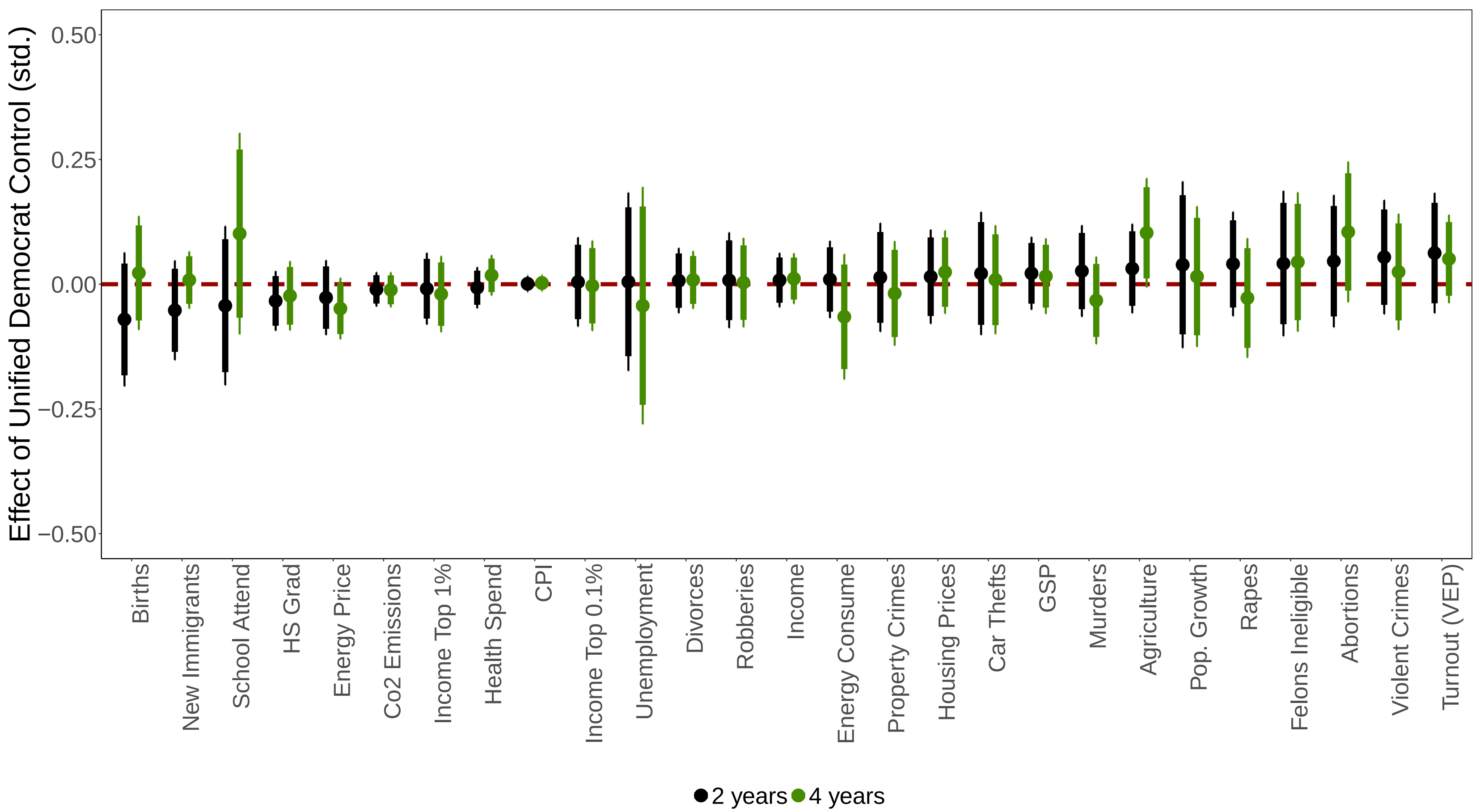


Figure 4 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for unified democratic control compared to unified republican control. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

Perhaps these nulls are masking time-based heterogeneities; that is, that party differences emerge in some time periods, but not others. This may be likely, as Caughey, Xu and Warshaw (2017) show that party effects on policy *outputs* vary over time. Figure A13 in the Online Appendix shows that our null effect holds when we allow for the estimates to vary over time. To do so, it interacts our measure of unified control with a continuous measure of time (the same conclusion holds if we make arbitrary decisions about where to split the sample along the time dimension). This suggests that our nulls are not a product of differences in time periods.

Perhaps, though, effects emerge when a party has control not only in a single-shot period—as we test with our difference-in-difference models above—but has *persistent* control over multiple periods. If party coalitions take multiple sessions to truly pass their agendas, persistent party control may have meaningful effect on policy outcomes.

To test this, we take two approaches: first, we change our treatment variable slightly to become a continuous measure of how long the party has been unified Democrat and second, we estimate dynamic difference-in-difference models. The first approach is relatively straightforward. In our difference-in-difference models, we include the same fixed effects as before, but them simply include a continuous measure that captures how long Democrats have been in power. These models also include how long the government has been divided and uses the number of years in Republican power as the left out category. With this approach, we look at effects up to four years downstream.

The dynamic difference-in-difference models include the same fixed effects as before (state, year, and state specific time trends), but add party control in the previous sessions to the models. In practice, this is done by adding the three- (to account for chambers that switch every two years) and five-year (to account for chambers that switch every five years) lags of unified Democratic control. These lagged treatment variables are then interacted with one another to allow us to estimate the combined effects over the period of study. Below we report the results for the triple interaction between control in the current period, three years previous, and five years previous. In practice, this allows us to estimate the effect of party control in two sequential sessions—capturing sequential control over as long as an eight year window.[27] Because we are using lags, we look at outcomes in the current year (T).

The results for these estimates are shown in Figure 5 below. For space constraints we focus on the continuous variable measure; however, the triple interaction can be found in the Online Appendix (see Figure A14). As can be seen, both methods give us a similar answer—that persistent unified control has little to no effect on policy outcomes on the timeline introduced by elections. The difference-in-difference estimates are all small substantively. The average effect size is mere 0.08% of a standard deviation (median = 0.04% σ). Moreover, these small estimates are statistically precise and centered around zero. Across the two methods, only 3/54 (income to the top 0.1% 2 years downstream, income to the top 1% 2 years downstream, and the abortion rate 4 years downstream) are statistically significant at the unadjusted 5% level—about what we would expect by chance. Moreover, none of these three exceptions clear the adjusted, multiple-hypothesis levels and all of them are small (with all being less than 1.2% of a standard deviate). Most of the p-values are large, with an average unadjusted p-value being 0.466 (median=0.459). With the second approach (the triple interaction), the estimates are slightly less precise—probably due to the strain two sets of triple interactions (for the democratic control and divided government) place on our state panel (see Figure A14). Still, we get the same answer—the estimates are spread almost normally around zero, with a dominant majority of the coefficients (20/27) being less than |15%| of a standard deviation. Moreover, none of the coefficients reach statistical significance at the 5% level. The average unadjusted p-value is 0.52 (median = 0.4). The standard errors for these models are larger; however, the estimates are, for the most part, substantively small. The mean estimated effect of persistent unified democratic control is a small -3.9% of a standard deviation (median = -4.6% σ). These results suggest that even persistent unified party control has little to no effect on policy outcomes.

These results show that *even when* Democrats have control of all three chambers for an extended period of time, they provide little separation from Republicans on the policy outcomes

[27] As in the difference-in-difference models above, we include the parallel lagged controls (and interactions) of divided government, to make the comparison of interest between persistent unified democratic control and persistent unified republican control.

that capture societal well-being and are thought to be useful metrics for retrospective voting.

Figure 5: Difference in Difference Estimate of Persistent Unified Democratic Control

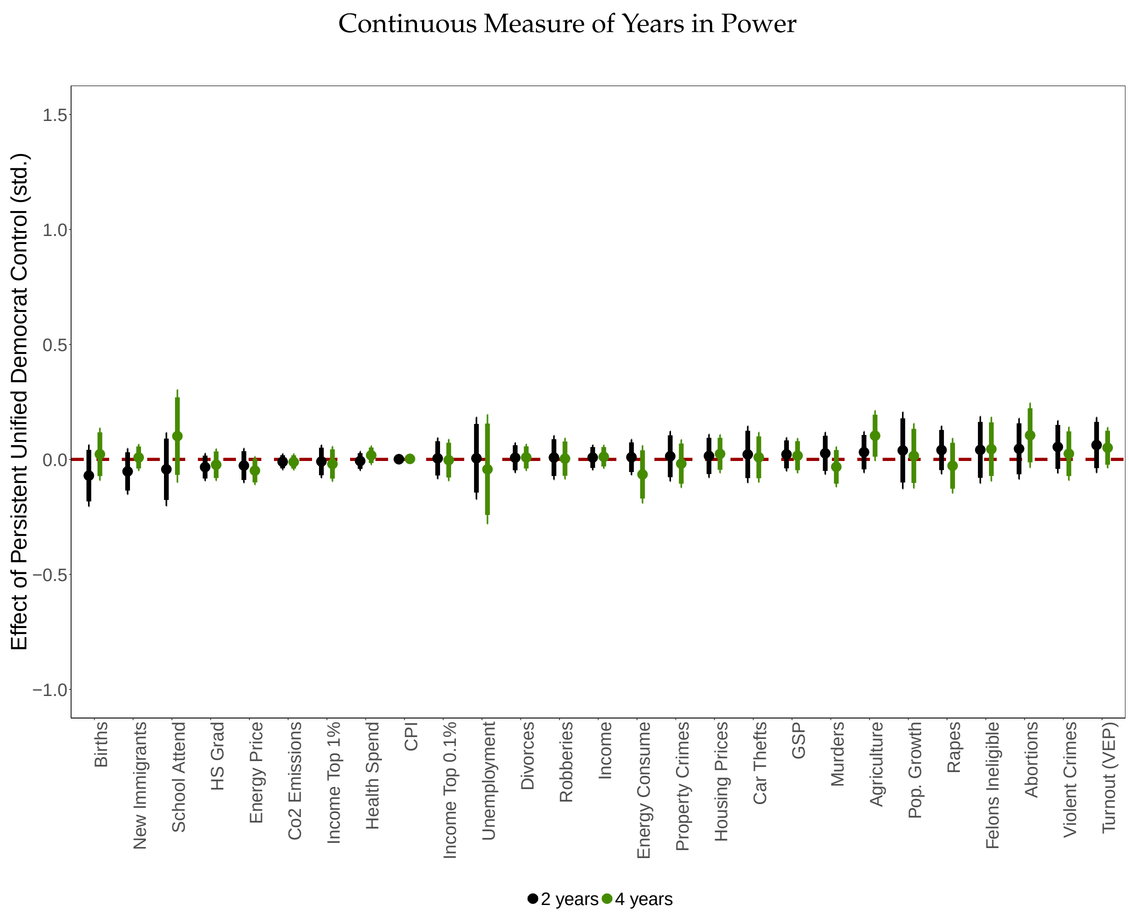


Figure 5 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for persistent (i.e. control in periods t, $t-3$, and $t-5$) unified democratic control compared to persistent unified republican control. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. The Figure displays the continuous treatment measure described in the text; for the triple interaction, see Figure A14 in the Online Appendix.

4.2 Regression Discontinuity Estimates

It's possible that the results just explored undersell the effect of party control on policy outcomes. While having the advantage of generalizability beyond any arbitrary cut-point, our difference-in-difference models may suffer from unobserved bias that attenuates our effects towards zero. To increase the internal validity of our estimates, in this section we transition to a regression discontinuity design. We follow the approach used in the last section—first examining the effects of flipping individual chambers, then considering the effects of unified power.

Figures 6 show the single cutoff RDD plots for four years downstream. Figure A28 in the Online Appendix shows the analogous graph for two years downstream. As in our difference-in-difference figure for individual chamber switches, the results are faceted by policy outcome type and broken by individual chambers. These estimates use the optimal bandwidth approach of Calonico, Cattaneo and Titiunik (2014).

As can be seen, out of the 81 models run (3 chambers by 27 policy outcomes) only 4 estimates (4.9%) are significant at the unadjusted 5% level—about what we would expect by random chance. None of 4 these exceptions, however, clear the multiple-hypothesis levels and all 4 of these exceptions are small to modest in size. Among all the coefficients, the average effect size is 0.5% of a standard deviation (median = 0.6% σ).

Figure 6: Single Cutoff RDD Effect of Democratic Control on Policy Outcomes (Fourth Year)

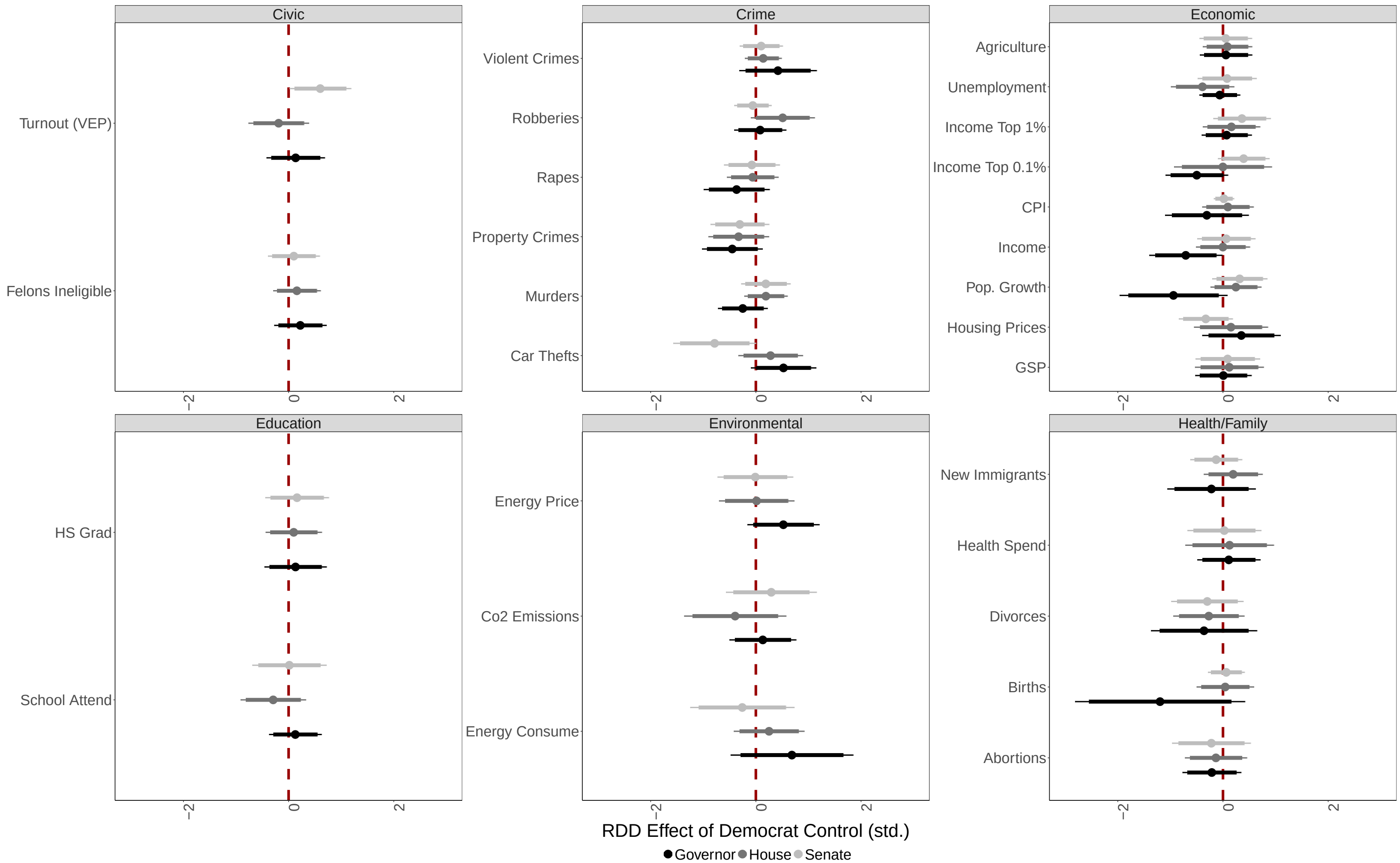


Figure 6 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the regression discontinuity estimates. The estimates are broken by the chamber that switches power. The running variable is modeled with a local kernel smoothed function. The estimates use the optimal bandwidth as specified by the `rdrobust` command in `STATA` created by Calonico, Cattaneo and Titiunik (2014). Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

These null effects persist and perhaps even become smaller and more precise when we use the alternate running variables suggested by Hall, Feigenbaum and Fouirnaies (Forthcoming), wherein the proximity to party control is conceptualized as a multi-dimensional space determined by the underlying close races. Figure A56 in the Online Appendix shows this visually by providing coefficient plots for the various regression discontinuity estimates across the various types of running variables. As can be seen, the estimates are remarkably similar across the various specifications of the running variable. Overall, we find that across our 108 RDD models (27 outcomes by 4 alternate measures of the running variable), that the estimates are all small to modest (median effect size = 6.1% of a standard deviation) only 3 (2.8%) show signs of a significant effect—slightly less than we would expect by chance.[28] All of the three exceptional cases come from instances where we use the number of seats as a running variable.[29] In these cases, the Hall, Feigenbaum and Fouirnaies (Forthcoming) running variables produce estimates that are closer to zero. It would seem then, if anything at all, the seat measure that we use produces similar results and in some cases produces estimates that probably overstate the effect of party control. But again, these different specifications of the running variable lead us to similar conclusions overall.

Figure A29 in the Online Appendix shows our first approach to trying to explore the RDD effect of unified control (the minimum running variable approach). Figure A30 in the Online Appendix shows the analogous figure for unified Republican control. As above, our regression discontinuity estimates use the optimal bandwidth created by Calonico, Cattaneo and Titiunik (2014). Figure A29 shows that even when we explore the RDD effect of unified control, the estimates are small, centered around zero as we might expect in a random draw of coefficients, and are (for the most part) not statistically significant at traditional levels.

These null effects also hold if we look at the MRDD estimates. Figures A31—A42 in the Online Appendix provide coefficient estimates for these models. Overall, when we adjust for multiple comparisons, only 10 (5.3%) out of the 189 tests are statistically significant.[30] Of these remaining

[28]If we take into account the direction of the effects, the overall median effect size is -6.1% of a standard deviation, with the overall mean being -7.3%.

[29]Number of Immigrants, Number of Seats RV, $p = 0.044$; CO_2 Emissions, Number of Seats RV, $p = 0.00$; Robbery Rate, Number of Seats RV, $p = 0.0003$.

[30]In the first year, the pattern is similar—only 5 (2.7%) clear the Bonferroni and Sidak significance

10, 6 belong to the amount of income belonging to the top 1% and 0.1%. These reveal that when democrats have unified control of the legislature, it may cause less income inequality. However, when Democrats only have access to one chamber or one chamber and the governorship, the reverse is true. This suggests that party control might effect the distribution of income once we account for other factors absorbed by our regression discontinuity design and the multi-cutoff nature of our data. That said, when we subject our models to an even more stringent comparison by adding state and year fixed effects to our MRDD models, the differences in income inequality lose their significance (see Figures A43–A54 in the Online Appendix for the results of this robustness check). With state and year fixed effects in the second year only 2 out of 189 estimates (1.1%) clear the Bonferroni and Sidak adjustments for multiple hypothesis testing.[31] This suggests that differences in policy outcomes at the cutoff may reflect idiosyncratic differences at the cutoffs we use.[32]

These null effects persist across a host of additional checks not shown here. We observe systematic nulls of small to modest size when we look at regression kink designs—where we look for differences in slopes across the cutoff (Card et al. 2015)[33], donut regression discontinuity designs—where we hold out observations close to the cutoff to try to account for any precise sorting around the cutoff (Barreca, Lindo and Waddell 2016)[34], or look across elections to explore the effects on longer-term policy outcomes (in years 3–9).[35]

Our analyses imply that any significant estimates appear to be the exception rather than the rule. For the most part, the Democratic party and the Republican party perform equally well on

levels.

[31]The same pattern exists in the first year only 1 coefficient (0.5%) clears the Bonferroni and Sidak significance levels.

[32]In general, the literature on MRDD suggests the inclusion of the full set of interactions between the treatment and running variables. If we remove the interactions for the running variables, our results become more precise; however, our conclusions do not change.

[33]With this approach, we observe 1 statistical significant effect in 108 tests (0.9%).

[34]With these, 4.9% of the effect estimates are significant at the 5% level—almost exactly what we would expect by chance.

[35]This last check allows us to explore the effect of legislative bodies in the upper chamber that have 4 year terms. With this check, we find that 7.2% of our RDD models are significant (only slightly more than we would expect by chance). Here the average effect size is 2% of a standard deviation (absolute value median = 8%).

a number of dimensions of societal well-being over the timeline introduced by elections. Simply put, economic, school, health/family, environmental, civic, and criminal justice outcomes are remarkably similar regardless of who is in control. Party control matters much less than what previous work has suggested.

5 Discussion

Do government coalitions influence the performance metrics by which they are (or, by many accounts, should be) retrospectively evaluated? There are many reasons to believe that they do; after all, large observational differences in policy outcomes exist (see Figure 2), scholars have argued that these observational differences are causal (e.g., Achen and Bartels 2016; Bartels 2009), immediate causal differences in policy *outputs* exist (Caughey, Xu and Warshaw 2017), a long literature on retrospective voting assumes differences, and the units which we explore (states) have a great degree of influence over public policy (Rogers 2017). In this paper, we have explored whether there is, indeed, a causal effect of party control on policy outcomes across a number of dimensions. We have focused our attention on the timeline introduced by elections. Using several different difference-in-difference and regression discontinuity models, we show that over this time period, political parties perform at roughly an equal level. That is, there is little separation between state governments controlled by Republicans and those controlled by Democrats.

Our work makes several important contributions that span both academic and contemporary political debates. First, the analysis conducted here lends itself to an important question in the broader study of democratic representation. Though retrospective voting on policy outcomes is thought to be widespread (Arnold and Carnes 2012; Bagues and Esteve-Volart 2013; Berry and Howell 2007; Fiorina 1978; Healy and Malhotra 2013; Holbein 2016; Kogan, Lavertu and Peskowitz 2016; Stokes 2015), our work suggests that party majorities may actually have much less control over the short-term metrics by which they are judged in elections. Though public policies may take time to have their largest effects, voters are not left with this luxury: being forced by rigid election schedules to make decisions about which party to support. Our analyses suggest that policy outcomes appear to be too noisy of metrics to evaluate public officials; if we are struggling to find effects, voters too may also struggle to finding meaningful variation in impacts directly related to elected officials and political parties. Given that party control has little effect during the timeline that voters have to decide who to vote for, our results call into question some of the broader ideas related to voter retrospection. The natural implication thus becomes: if voters respond to performance signals that are largely just noise—being out of the control of elected officials—perhaps retrospective voting on policy outcomes is not as desirable as many have argued. Perhaps voting on the economy is just as bad as voting on shark attacks, the weather, sporting events, or policy decisions outside the realm of control of policymakers.

Based on this information, it remains unclear just how elected officials should be judged. Our results could be taken in two directions in future work. First, our results imply that in order for citizens to truly hold elected officials accountable, there may need to be more time allocated between when one party is in power and the time when accountability decisions are made. The contemporary schedule of elections on 2- and 4-year windows in the United States may not long enough for policy differences to emerge. Future work might do well to consider the effects of party control in contexts where election windows are wider. Second, our results suggest that elected officials may be best judged by alternate performance metrics. For example, it may be more useful for retrospective systems to be built around the inputs elected officials put into the policy system, rather than what comes out. This approach mirrors the public policy literature that argues for performance evaluations of other public officials (i.e. teachers, principals, etc.) to be build around inputs-based rather than outputs-based metrics (e.g. Horsford 2010; Podgursky and Springer 2007).

From a practical perspective, our work provides scholars, citizens, and practitioners with a clear (but somewhat constrained) picture of how Democrats perform relative to Republicans. This question, beyond capturing the attention of scholars of retrospective voting, political parties, and democratic representation, also plays a central role of many political campaigns. Previous research, observational patterns, and conventional wisdom all suggest that the party in control influences policy outcomes. Our work should give pause to those seeking to misinterpret simplistic, conditional on observables or, even worse, raw difference comparisons between political parties.

These provide a markedly different view of the performance of political parties than the more rigorous methods we have employed.

Our study also speaks to the literature on political credit claiming and blame avoidance (e.g. Weaver 1986; Grimmer, Messing and Westwood 2012; Samuels 2002). Studies, and practical experience, has shown that politicians are more than willing to claim credit for a well-performing economy, or good outcomes on other policy dimensions, and to shun responsibility when things go badly. Our work provides compelling evidence that many political efforts to claim credit for immediate improvements in policy outcomes may be misleading. In many cases, politicians are simply claiming credit for happy accidents that went in their favor, rather than returns on their policy initiatives (which do not seem to appear within the window most elected officials hold office). The reverse also holds true: many politicians who are blamed for poor performance may actually just be woefully unlucky: hence the problems for retrospection we have discussed in this paper.

Despite our study's strengths, our approach does have some important limitations. As we discussed, our data is inherently limited in its ability to look far downstream. Fully mapping the dynamic effects of party control over multiple periods is difficult due to too few years where data is available. With this limitation in mind, we note that this is a reasonable window for observing policy effects. Policy outputs (i.e. legislation) move in such a time period (Caughey, Xu and Warshaw 2017) and policy evaluations from other contexts have shown that we can expect immediate movement in the outcomes we explore if the treatment is, indeed, effective (e.g., Chiang 2009; Dee and Jacob 2011; Cook et al. 2014; Angrist, Pathak and Walters 2013).

Future work would do well to consider the effects of party control at the local level on a host of policy outcomes. We have intentionally focused on the state level—where much of contemporary policy debates occur. However, it is possible that partisan control of school boards, mayoral offices, and city councils has larger, more pronounced effects on policy outcomes. In expanding this literature, we emphasize the need for scholars to be thorough in exploring than just one or two cherry-picked policy outcomes.

As long as voters are tasked with the job of evaluating the policy performance of elected officials, studying the effects of party control on policy outcomes is vitally important. Our work suggests that policy outcomes may be too noisy of a foundation on which to build retrospective voting systems.

References

Achen, Christopher H and Larry M Bartels. 2016. *Democracy for realists: Why elections do not produce responsive government*. Princeton, NJ: Princeton University Press.

Ahn, Thomas and Jacob Vigdor. 2014. "The Impact of No Child Left Behind's Accountability Sanctions on School Performance: Regression Discontinuity Evidence from North Carolina." *National Bureau of Economic Research* (w20511).

Alesina, Alberto and Howard Rosenthal. 1995. *Partisan Politics, Divided Government, and the Economy*. Cambridge University Press.

Alesina, Alberto and Nouriel Roubini. 1992. "Political Cycles in OECD Economies." *The Review of Economic Studies* 59(4):663–688.

Alvarez, R. Michael, Geoffrey Garrett and Peter Lange. 1991. "Government Partisanship, Labor Organization, and Macroeconomic Performance." *The American Political Science Review* 85(2):539–556.

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Angrist, Joshua D and Jörn-Steffen Pischke. 2014. *Mastering'metrics: The path from cause to effect*. Princeton University Press.

Angrist, Joshua D, Parag A Pathak and Christopher R Walters. 2013. "Explaining charter school effectiveness." *American Economic Journal: Applied Economics* 5(4):1–27.

Arnold, R Douglas and Nicholas Carnes. 2012. "Holding mayors accountable: New York's executives from Koch to Bloomberg." *American Journal of Political Science* 56(4):949–963.

Bagues, Manuel and Berta Esteve-Volart. 2013. "PoliticiansÕ luck of the draw: Evidence from the Spanish christmas lottery.".

Barreca, Alan I, Jason M Lindo and Glen R Waddell. 2016. "Heaping-Induced Bias in Regression-Discontinuity Designs." *Economic Inquiry* 54(1):268–293.

Bartels, Larry M. 2009. *Unequal democracy: The political economy of the new gilded age*. Princeton University Press.

Bateson, Regina. 2012. "Crime victimization and political participation." *American Political Science Review* 106(3):570–587.

Berry, Christopher R and William G Howell. 2007. "Accountability and local elections: Rethinking retrospective voting." *Journal of Politics* 69(3):844–858.

Besley, Timothy and Anne Case. 2003. "Political Institutions and Policy Choices: Evidence from the United States." *Journal of Economic Literature* 41(1):7–73.
URL: *http://www.jstor.org/stable/3217387*

Broockman, David E. and Christopher Skovron. 2015. "What Politicians Believe About Their Constituents: Asymmetric Misperceptions and Prospects for Constituency Control.".

Buddelmeyer, Hielke and Emmanuel Skoufias. 2004. *An evaluation of the performance of regression discontinuity design on PROGRESA*. Vol. 827 World Bank Publications.

Busby, Ethan C, James N Druckman and Alexandria Fredendall. 2017. "The Political Relevance of Irrelevant Events." *The Journal of Politics* 79(1):346–350.

Butler, Daniel M. 2009. "A regression discontinuity design analysis of the incumbency advantage and tenure in the US House." *Electoral Studies* 28(1):123–128.

Butler, Daniel M. and Matthew J. Butler. 2006. "Splitting the difference? Causal inference and theories of split-party delegations." *Political Analysis* 14(4):439–455.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.

Card, David, David S Lee, Zhuan Pei and Andrea Weber. 2015. "Inference on causal effects in a generalized regression kink design." *Econometrica* 83(6):2453–2483.

Caughey, Devin and Christopher Warshaw. 2015. "The Dynamics of State Policy Liberalism, 1936Ű2014." *American Journal of Political Science* pp. n/a–n/a.

Caughey, Devin, Yiqing Xu and Christopher Warshaw. 2017. "Incremental democracy: The policy effects of partisan control of state government." *The Journal of Politics* 79(4):000–000.

Chappell, Henry W. and William R. Keech. 1986. "Party Differences in Macroeconomic Policies and Outcomes." *The American Economic Review* 76(2):71–74.

Chen, Jowei. 2013. "Voter partisanship and the effect of distributive spending on political participation." *American Journal of Political Science* 57(1):200–217.

Chiang, Hanley. 2009. "How accountability pressure on failing schools affects student achievement." *Journal of Public Economics* 93(9):1045–1057.

Cook, Philip J, Kenneth Dodge, George Farkas, Roland G Fryer Jr, Jonathan Guryan, Jens Ludwig, Susan Mayer, Harold Pollack and Laurence Steinberg. 2014. The (surprising) efficacy of academic and behavioral intervention with disadvantaged youth: Results from a randomized experiment in chicago. Technical report National Bureau of Economic Research.

Cummins, Jeff. 2011. "Party Control, Policy Reforms, and the Impact on Health Insurance Coverage in the U.S. States*." *Social Science Quarterly* 92(1):246–267.
URL: *http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6237.2011.00766.x/abstract*

de Benedictis-Kessner, Justin and Christopher Warshaw. 2016. "Mayoral partisanship and municipal fiscal policy." *The Journal of Politics* 78(4):1124–1138.

Dee, Thomas S and Brian Jacob. 2011. "The impact of No Child Left Behind on student achievement." *Journal of Policy Analysis and management* 30(3):418–446.

Dudoit, Sandrine, Juliet Popper Shaffer and Jennifer C Boldrick. 2003. "Multiple hypothesis testing in microarray experiments." *Statistical Science* pp. 71–103.

Dye, Thomas R. 1966. *Politics, Economics, and the Public: Policy Outcomes in the American States*. Rand McNally. Google-Books-ID: MYdCAAAAIAAJ.

Eggers, Andrew C, Anthony , Jens Hainmueller, Andrew B Hall and James M Snyder. 2015. "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races." *American Journal of Political Science* 59(1):259–274.

Erikson, Robert S., Gerald C. Jr. Wright and John P. McIver. 1989. "Political Parties, Public Opinion, and State Policy in the United States." *American Political Science Review* 83(03):729–750.

Faricy, Christopher. 2011. "The politics of social policy in America: The causes and effects of indirect versus direct social spending." *The Journal of Politics* 73(01):74–83.

Fearon, James D. 1999. "Electoral accountability and the control of politicians: selecting good types versus sanctioning poor performance." *Democracy, accountability, and representation* 55:61.

Ferejohn, John. 1986. "Incumbent performance and electoral control." *Public choice* 50(1):5–25.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P Newhouse, Heidi Allen, Katherine Baicker and Oregon Health Study Group. 2012. "The Oregon health insurance experiment: evidence from the first year." *The Quarterly journal of economics* 127(3):1057–1106.

Fiorina, Morris P. 1978. "Economic Retrospective Voting in American National Elections: AM icro-A nalysis." *American Journal of Political Science* 22(2):426–443.

Fiorina, Morris P. 1981. "Retrospective voting in American national elections.".

Fowler, Anthony and B Pablo Montagnes. 2015. "College football, elections, and false-positive results in observational research." *Proceedings of the National Academy of Sciences* 112(45):13800–13804.

Franzese, Robert J. 2002. "Electoral and Partisan Cycles in Economic Policies and Outcomes." *Annual Review of Political Science* 5(1):369–421.

Gerber, Elisabeth R and Daniel J Hopkins. 2011. "When mayors matter: estimating the impact of mayoral partisanship on city policy." *American Journal of Political Science* 55(2):326–339.

Gomez, Brad T, Thomas G Hansford and George A Krause. 2007. "The Republicans should pray for rain: Weather, turnout, and voting in US presidential elections." *Journal of Politics* 69(3):649–663.

Grant, Ruth W and Robert O Keohane. 2005. "Accountability and abuses of power in world politics." *American political science review* 99(1):29–43.

Green, Donald P, Terence Y Leong, Holger L Kern, Alan S Gerber and Christopher W Larimer. 2009. "Testing the accuracy of regression discontinuity analysis using experimental benchmarks." *Political Analysis* 17(4):400–417.

Grimmer, Justin, Solomon Messing and Sean J Westwood. 2012. "How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation." *American Political Science Review* 106(4):703–719.

Grofman, Bernard. 2004. "Downs and Two-Party Convergence." *Annual Review of Political Science* 7(1):25–46.

Grose, Christian R and Bruce I Oppenheimer. 2007. "The Iraq War, Partisanship, and Candidate Attributes: Variation in Partisan Swing in the 2006 US House Elections." *Legislative Studies Quarterly* 32(4):531–557.

Hacker, Jacob S and Paul Pierson. 2010. "Winner-take-all politics: Public policy, political organization, and the precipitous rise of top incomes in the United States." *Politics & Society* 38(2):152–204.

Hainmueller, Jens and Holger Lutz Kern. 2008. "Incumbency as a source of spillover effects in mixed electoral systems: Evidence from a regression-discontinuity design." *Electoral Studies* 27(2):213–227.

Hall, Andrew B. 2015. "What Happens When Extremists Win Primaries?" *American Political Science Review* 109(01):18–42.

Hall, Andrew B, James J. Feigenbaum and Alexader Fouirnaies. Forthcoming. "The Majority-Party Disadvantage: Revising Theories of Legislative Organization." *Quarterly Journal of Political Science* .

Hall, Andrew B, Jesse Yoder and Nishant Karandikar. 2017. "Economic Distress and Voting: Evidence from the Subprime Mortgage Crisis.".

Healy, Andrew and Gabriel S Lenz. 2014. "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy." *American Journal of Political Science* 58(1):31–47.

Healy, Andrew and Gabriel S Lenz. 2017. "Presidential Voting and the Local Economy: Evidence from Two Population-Based Data Sets." *The Journal of Politics* 79(4):1419–1432.

Healy, Andrew J, Neil Malhotra and Cecilia Hyunjung Mo. 2010. "Irrelevant events affect voters' evaluations of government performance." *Proceedings of the National Academy of Sciences* 107(29):12804–12809.

Healy, Andrew and Neil Malhotra. 2009. "Myopic voters and natural disaster policy." *American Political Science Review* 103(3):387–406.

Healy, Andrew and Neil Malhotra. 2013. "Retrospective voting reconsidered." *Annual Review of Political Science* 16:285–306.

Henderson, John and John Brooks. 2016. "Mediating the Electoral Connection: The Information Effects of Voter Signals on Legislative Behavior." *The Journal of Politics* 78(3):653–669.

Hibbs, Douglas A. 1977. "Political Parties and Macroeconomic Policy." *The American Political Science Review* 71(4):1467–1487.

Holbein, John. 2016. "Left behind? Citizen responsiveness to government performance information." *American Political Science Review* 110(2):353–368.

Horsford, Sonya Douglass. 2010. *New perspectives in educational leadership: Exploring social, political, and community contexts and meaning*. Vol. 1 Peter Lang.

Imbens, Guido W and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of econometrics* 142(2):615–635.

Jardim, Ekaterina, Mark C Long, Robert Plotnick, Emma van Inwegen, Jacob Vigdor and Hilary Wething. 2017. "Minimum Wage Increases, Wages, and Low-Wage Employment: Evidence from Seattle.".

Jennings, Edward T. 1979. "Competition, Constituencies, and Welfare Policies in American States." *The American Political Science Review* 73(2):414–429.

Jordan, Marty P. and Matt Grossmann. 2016. "The Correlates of State Policy Project v.1.0.".
URL: *https://www.ippsr.msu.edu/public-policy/correlates-state-policy*

Kemp, Kathleen A. 1978. "Nationalization of the American States A Test of the Thesis." *American Politics Quarterly* 6(2):237–247.

Key, Valdimer Orlando. 1966. *The responsible electorate*. Belknap Press of Harvard University Press.

Kogan, Vladimir, Stephane Lavertu and Zachary Peskowitz. 2016. "Performance Federalism and Local Democracy: Theory and Evidence from School Tax Referenda." *American Journal of Political Science* 60(2):418–435.

Kramer, Gerald H. 1971. "Short-term fluctuations in US voting behavior, 1896–1964." *American political science review* 65(1):131–143.

Lax, Jeffrey R. and Justin H. Phillips. 2012. "The Democratic Deficit in the States." *American Journal of Political Science* 56(1):148–166.

Lee, David S. 2008. "Randomized experiments from non-random selection in US House elections." *Journal of Econometrics* 142(2):675–697.

Lee, David S and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48:281–355.

Leigh, Andrew. 2008. "Estimating the impact of gubernatorial partisanship on policy settings and economic outcomes: A regression discontinuity approach." *European Journal of Political Economy* 24(1):256–268.

Lemieux, Thomas and Kevin Milligan. 2008. "Incentive effects of social assistance: A regression discontinuity approach." *Journal of Econometrics* 142(2):807–828.

Ludwig, Jens and Philip J Cook. 2000. "Homicide and suicide rates associated with implementation of the Brady Handgun Violence Prevention Act." *Jama* 284(5):585–591.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142(2):698–714.

Newhouse, Joseph P. and Rand Corporation. Insurance Experiment Group. 1993. *Free for all?: lessons from the RAND health insurance experiment*. Harvard University Press.

Nordhaus, William D. 1975. "The political business cycle." *The review of economic studies* 42(2):169–190.

Papay, John P, John B Willett and Richard J Murnane. 2011. "Extending the regression-discontinuity approach to multiple assignment variables." *Journal of Econometrics* 161(2):203–207.

Pettersson-Lidbom, Per. 2008. "Do Parties Matter for Economic Outcomes? A Regression-Discontinuity Approach." *Journal of the European Economic Association* 6(5):1037–1056.

Podgursky, Michael J and Matthew G Springer. 2007. "Teacher performance pay: A review." *Journal of Policy Analysis and Management* 26(4):909–949.

Rogers, Steven. 2017. "Electoral Accountability for State Legislative Roll Calls and Ideological Representation." *American Political Science Review* 111(3):555–571.

Rose, Douglas D. 1973. "National and Local Forces in State Politics: The Implications of Multi-Level Policy Analysis." *The American Political Science Review* 67(4):1162–1173.

Samuels, David J. 2002. "Pork barreling is not credit claiming or advertising: Campaign finance and the sources of the personal vote in Brazil." *The journal of Politics* 64(3):845–863.

Sances, Michael W. 2017. "Attribution Errors in Federalist Systems: When Voters Punish the President for Local Tax Increases." *The Journal of Politics* 79(4):1286–1301.

Shaffer, Juliet Popper. 1995. "Multiple hypothesis testing." *Annual review of psychology* 46:561.

Shor, Boris and Nolan McCarty. 2011. "The Ideological Mapping of American Legislatures." *American Political Science Review* 105(03):530–551.

Stokes, Leah C. 2015. "Electoral backlash against climate policy: a natural experiment on retrospective voting and local resistance to public policy." *American Journal of Political Science* .

Weaver, R Kent. 1986. "The politics of blame avoidance." *Journal of public policy* 6(4):371–398.

Winters, Richard. 1976. "Party Control and Policy Change." *American Journal of Political Science* 20(4):597–636.

Wolfers, Justin. 2002. *Are voters rational?: Evidence from gubernatorial elections*.
URL: *http://users.nber.org/ jwolfers/papers/Voterrationality(latest).pdf*

Yokum, David, Anita Ravishankar and Alexander Coppock. 2017. *Evaluating the Effects of Police Body-Worn Cameras: A Randomized Controlled Trial*. Working Paper.

Online Appendix:
Noisy Retrospection: The Effect of Party Control on Policy Outcomes

Not intended for publication in printed versions

6 Online Appendix Overview

In the Online Appendix, we provide supporting information that:

- shows states' party control dynamics over time,
- shows alternate specifications for the difference-in-difference models,
- provides the standard RDD specification checks,
- shows the robustness of our conclusions to a combination of our RDD model with a difference–in–difference that leverages states that switch party control,
- shows the robustness of our conclusions to alternate specifications of the running variable suggested by Hall, Feigenbaum and Fouirnaies (Forthcoming).

7 States' Party Control Over Time

One may wonder what party control looks like within states over time. Or, put differently, which states are close to switching over time and, hence, help identify the diff/diff and regression discontinuity models estimated in the text. Figures A1–A6 show how close legislative chambers and governors are to being controlled by a Democratic majority from 1960–2010.[36] These shows that some states switch between party control (like Illinois or Nevada), while others stay more constantly under Democratic (Massachusetts) or Republican (Utah).

[36]Nebraska is omitted, given the state's unicameral nonpartisan legislature.

Figure A1: States' Proximity to Democratic Majorities Over Time (Lower, 1)

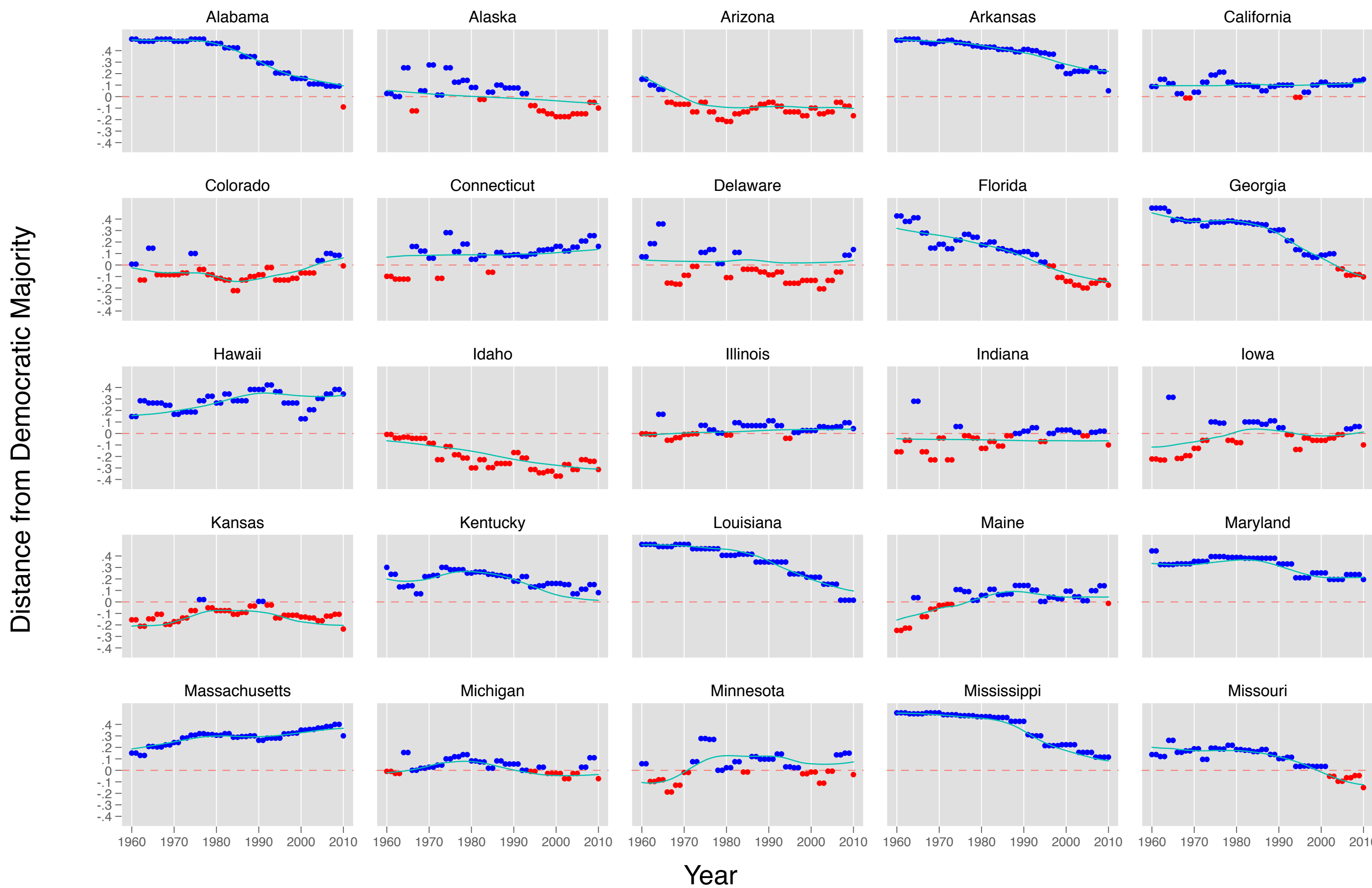


Figure A2: States' Proximity to Democratic Majorities Over Time (Lower, 2)

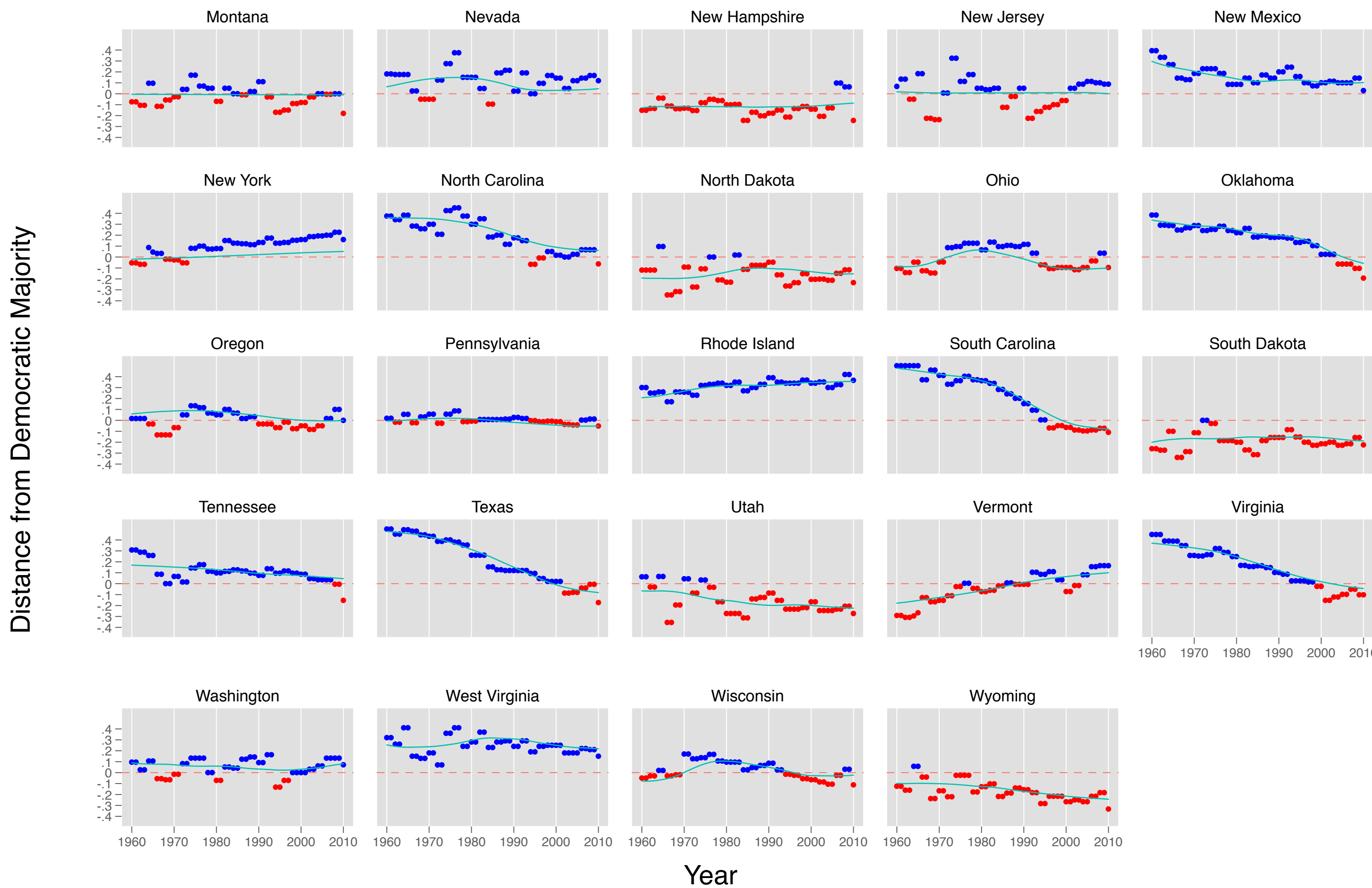


Figure A3: States' Proximity to Democratic Majorities Over Time (Upper, 1)

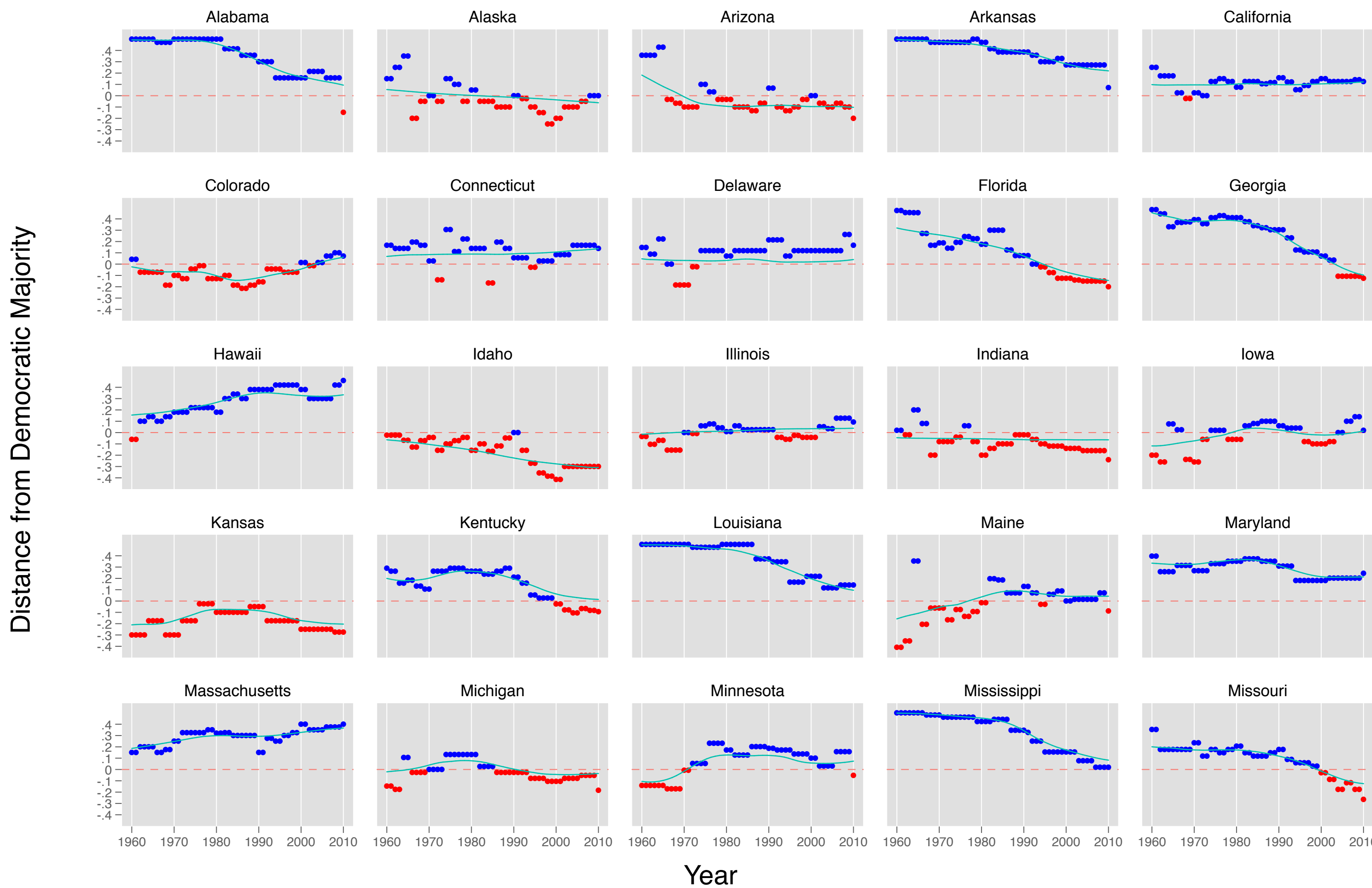


Figure A4: States' Proximity to Democratic Majorities Over Time (Upper, 2)

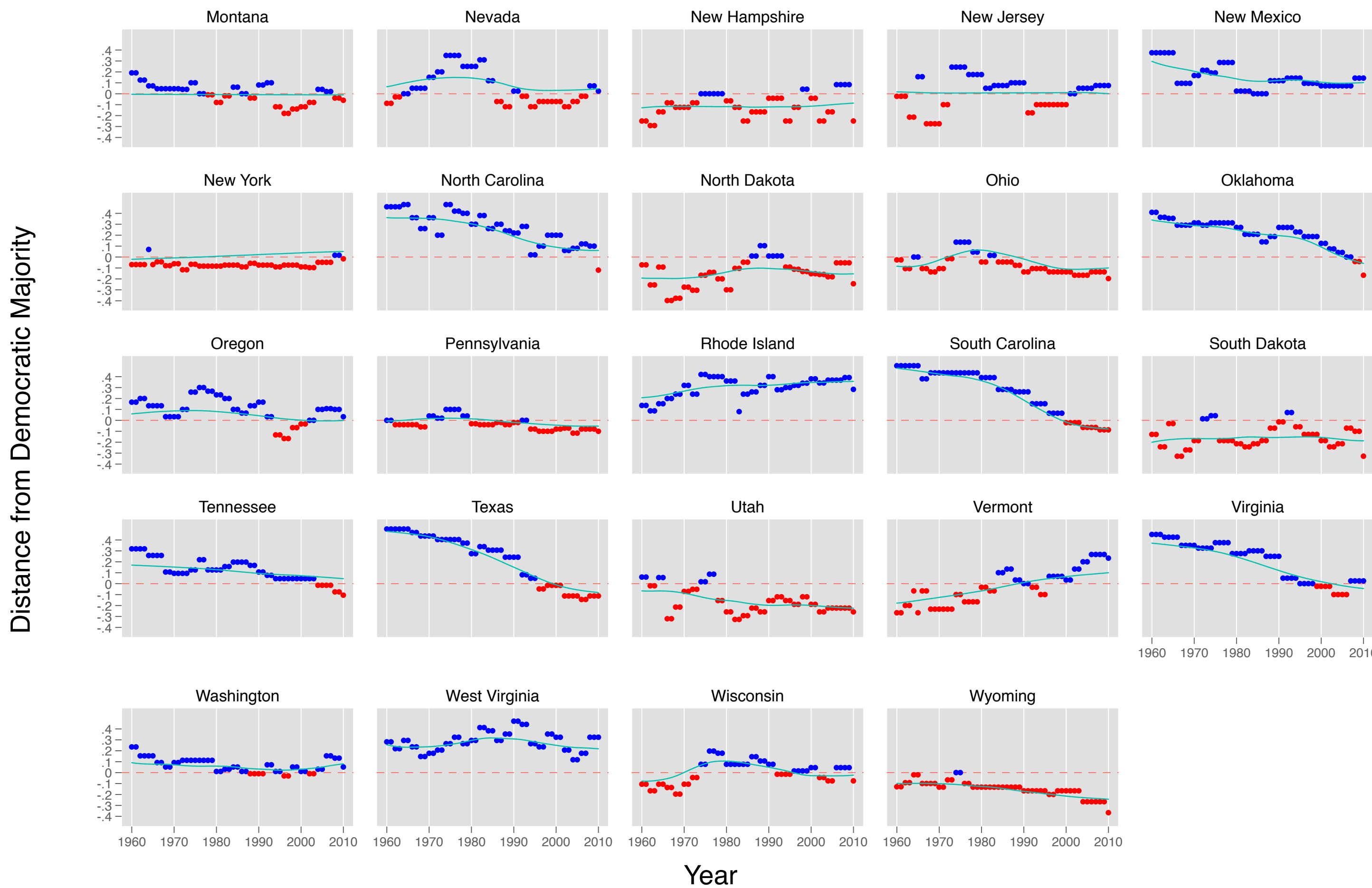


Figure A5: States' Proximity to Democratic Majorities Over Time (Governor, 1)

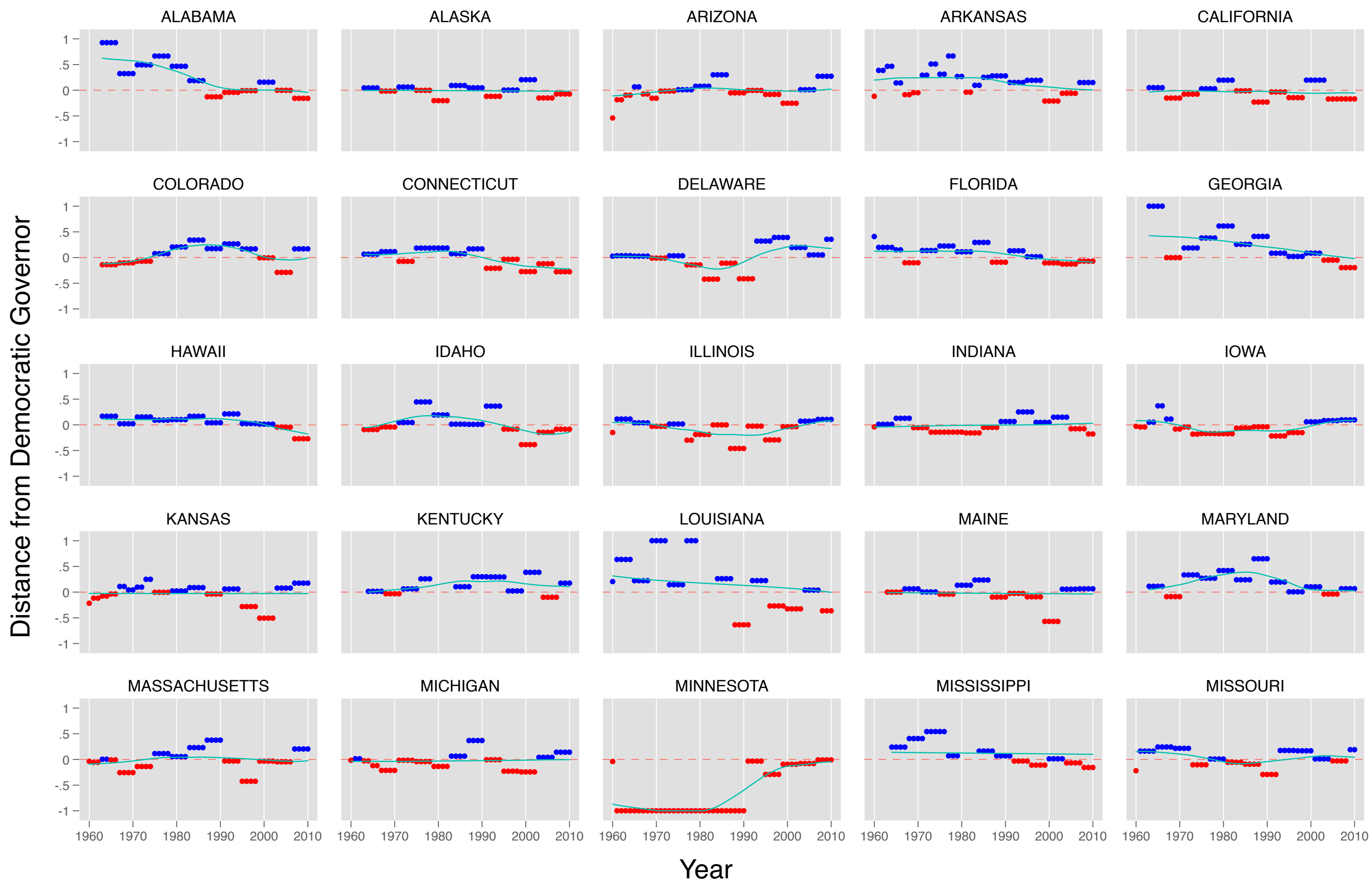
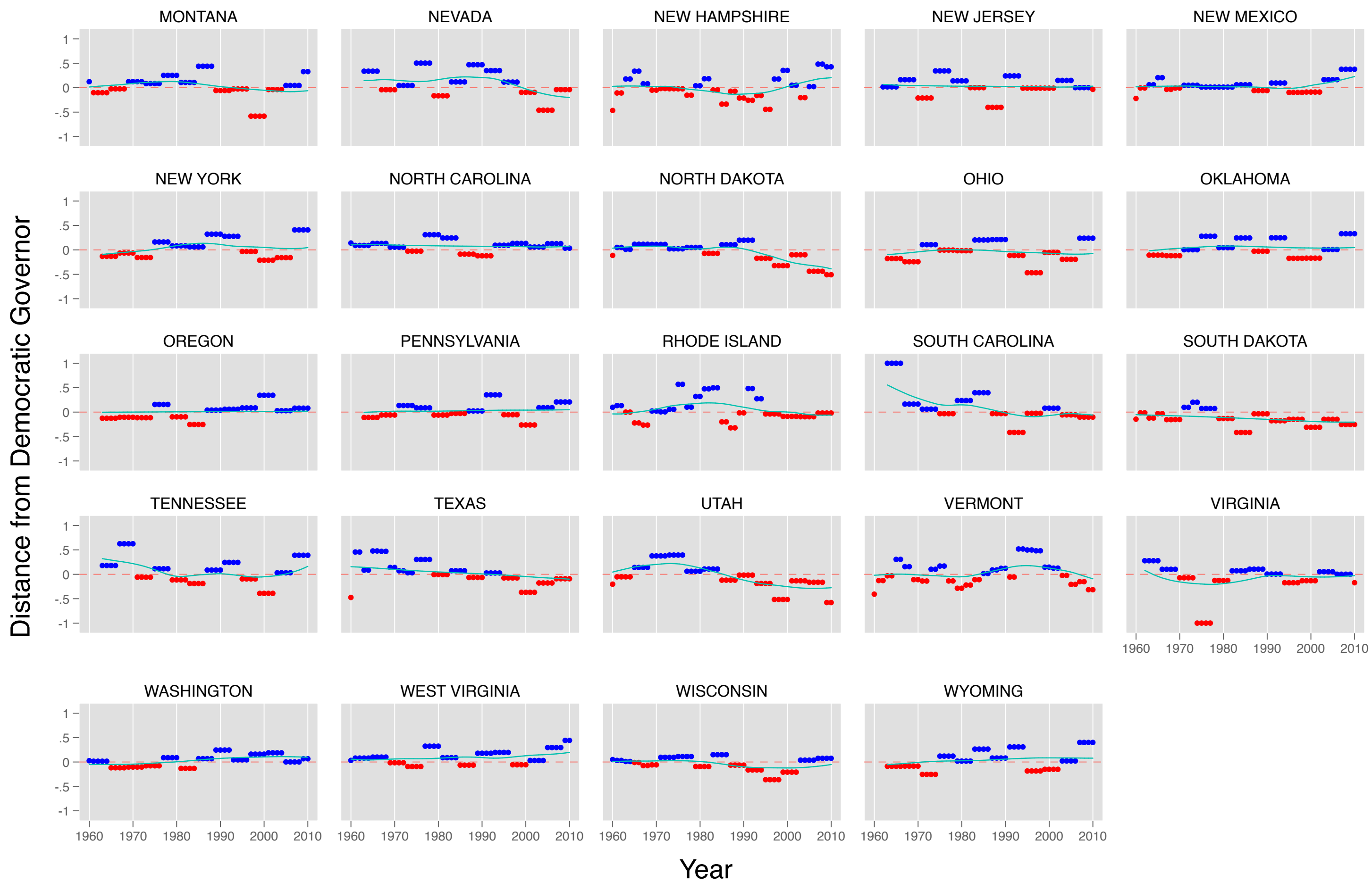


Figure A6: States' Proximity to Democratic Majorities Over Time (Governor, 2)



8 Alternate Difference-in-Difference Specifications

8.1 Difference-in-Difference Estimates for Single Chamber Switches

Figure A7: Difference in Difference Estimates of Single Chamber Changes to Democratic Control (Second Year)

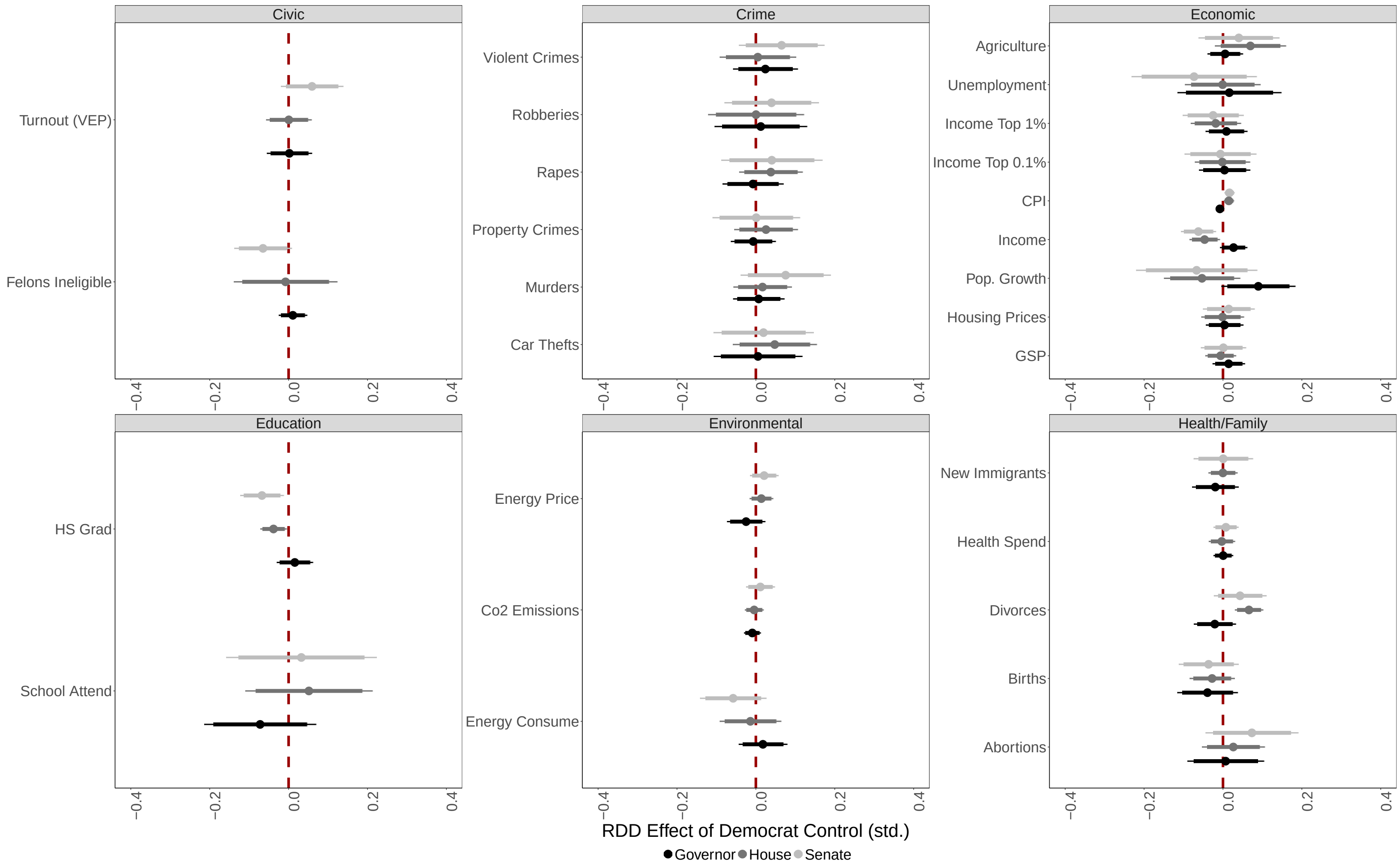


Figure A7 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for the effects of each individual chamber. Coefficients are faceted by policy area and broken by individual chamber within facets. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

Figure A8: Distribution of Coefficient Estimates for Single Chamber Changes to Democratic Control (Diff-Diff, Years T–T+8)

(a) House (b) Governor (c) Senate

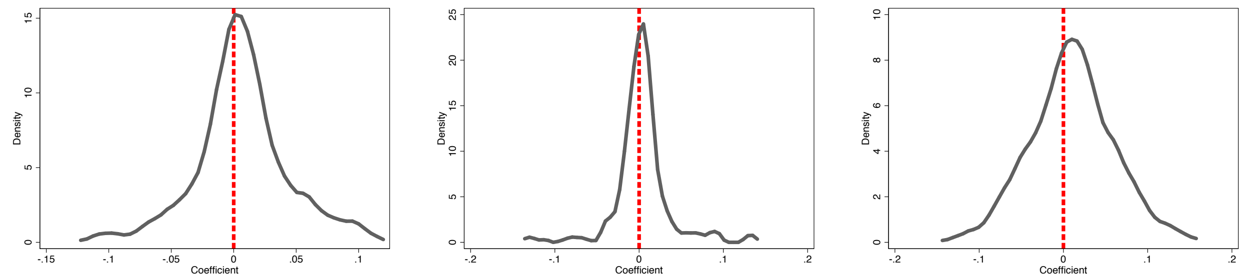


Figure A8 plots the distribution of coefficients from the single chamber switches difference in difference estimates for years t–t+8 downstream.

Figure A9: Distribution of P-Value Estimates for Single Chamber Changes to Democratic Control (Diff-Diff, Years T–T+8)

(a) House (b) Governor (c) Senate

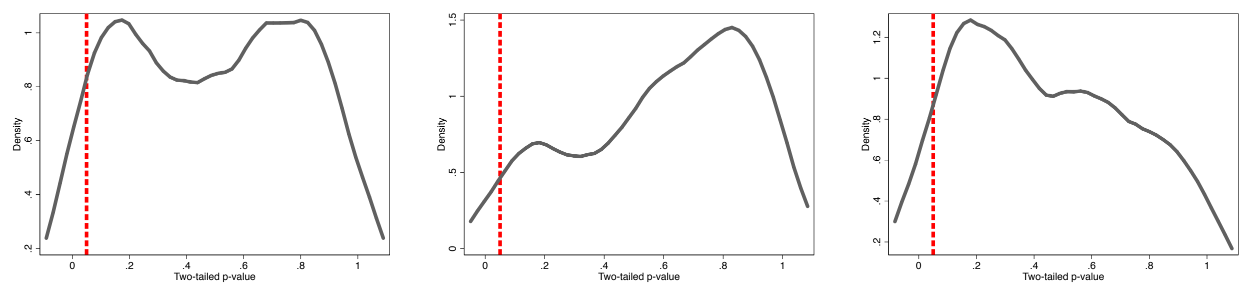


Figure A9 plots the distribution of coefficients from the single chamber switches difference in difference estimates for years t–t+8 downstream.

8.2 Other Estimates for Difference-in-Difference Models with Unified Party Control

Figure A10: Difference in Difference Estimates of Divided Government Compared to Unified Republican Control

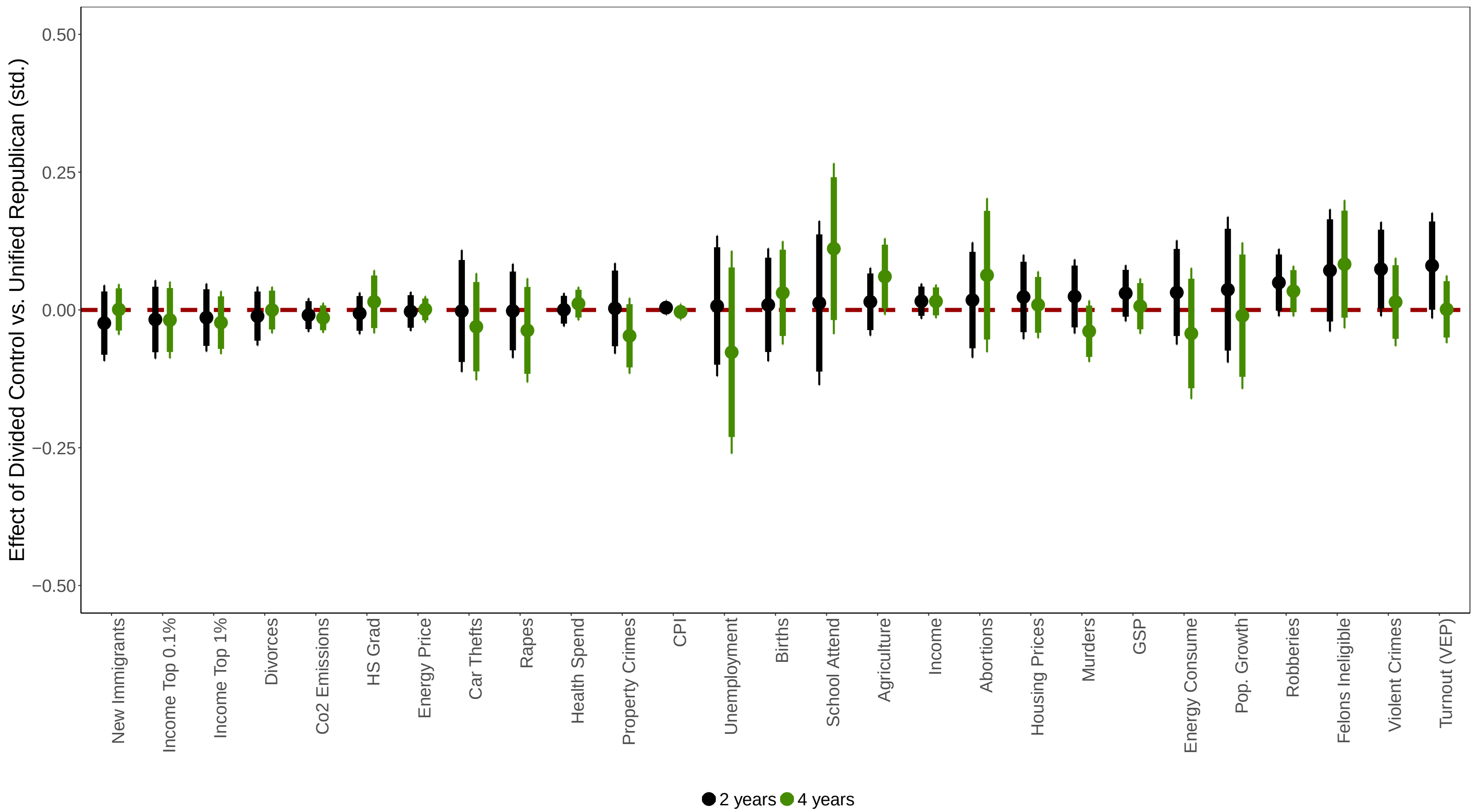


Figure A10 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for divided government to unified republican control. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

Figure A11: Difference in Difference Estimates of Unified Democrat Control Compared to Divided Government

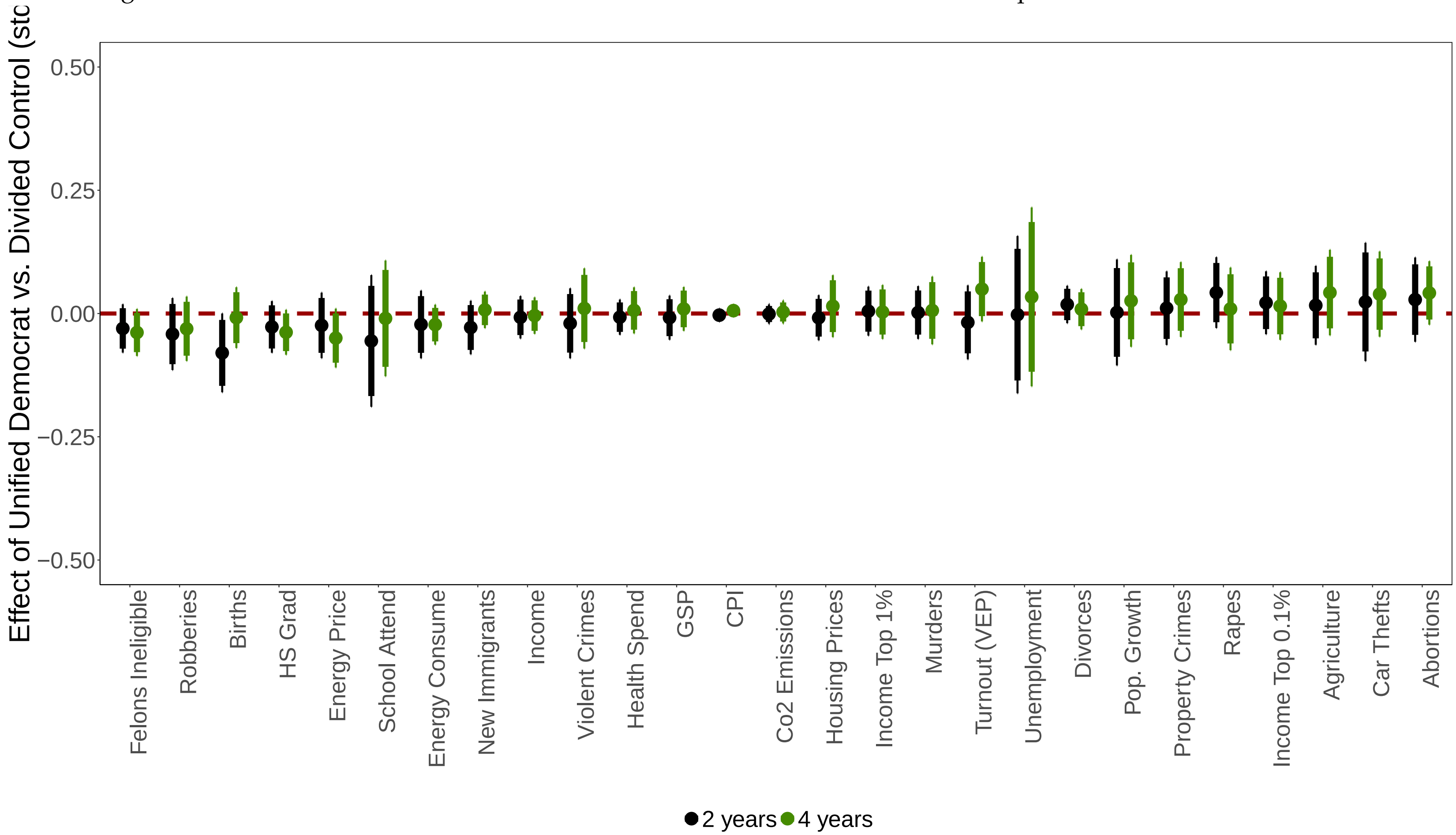


Figure A11 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for unified democrat compared to divided government. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

Figure A12: Summary of Difference in Difference Estimates for Unified Democrat vs. Unified Republican

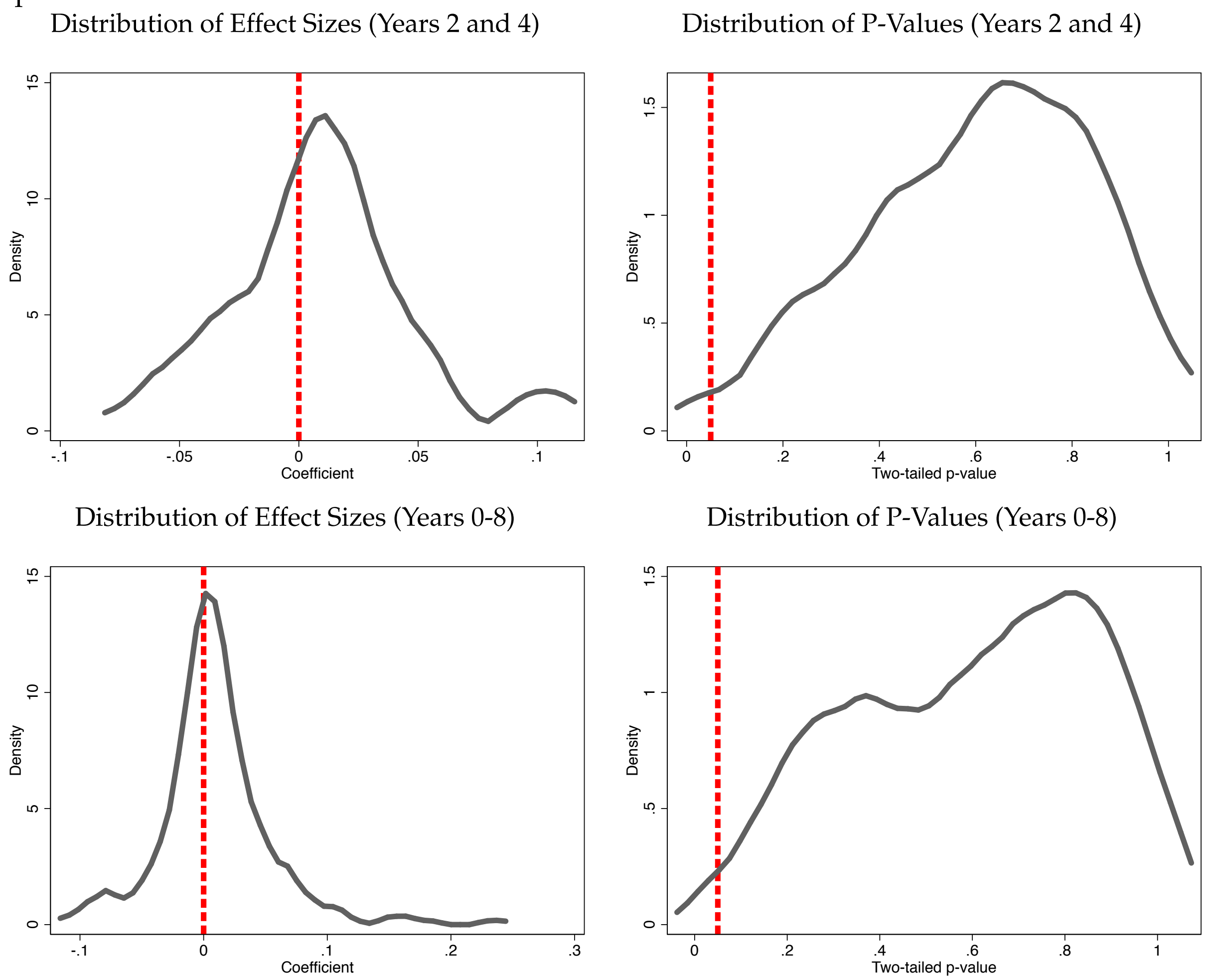


Figure A12 plots the distribution of p-values and coefficients from the unified power difference in difference estimates shown in Figure 4 in the text along with corresponding distributions for years t–t+8 downstream.

Figure A13: Difference in Difference Estimates of Unified Democratic Control Compared to Unified Republican Control, Time Interaction

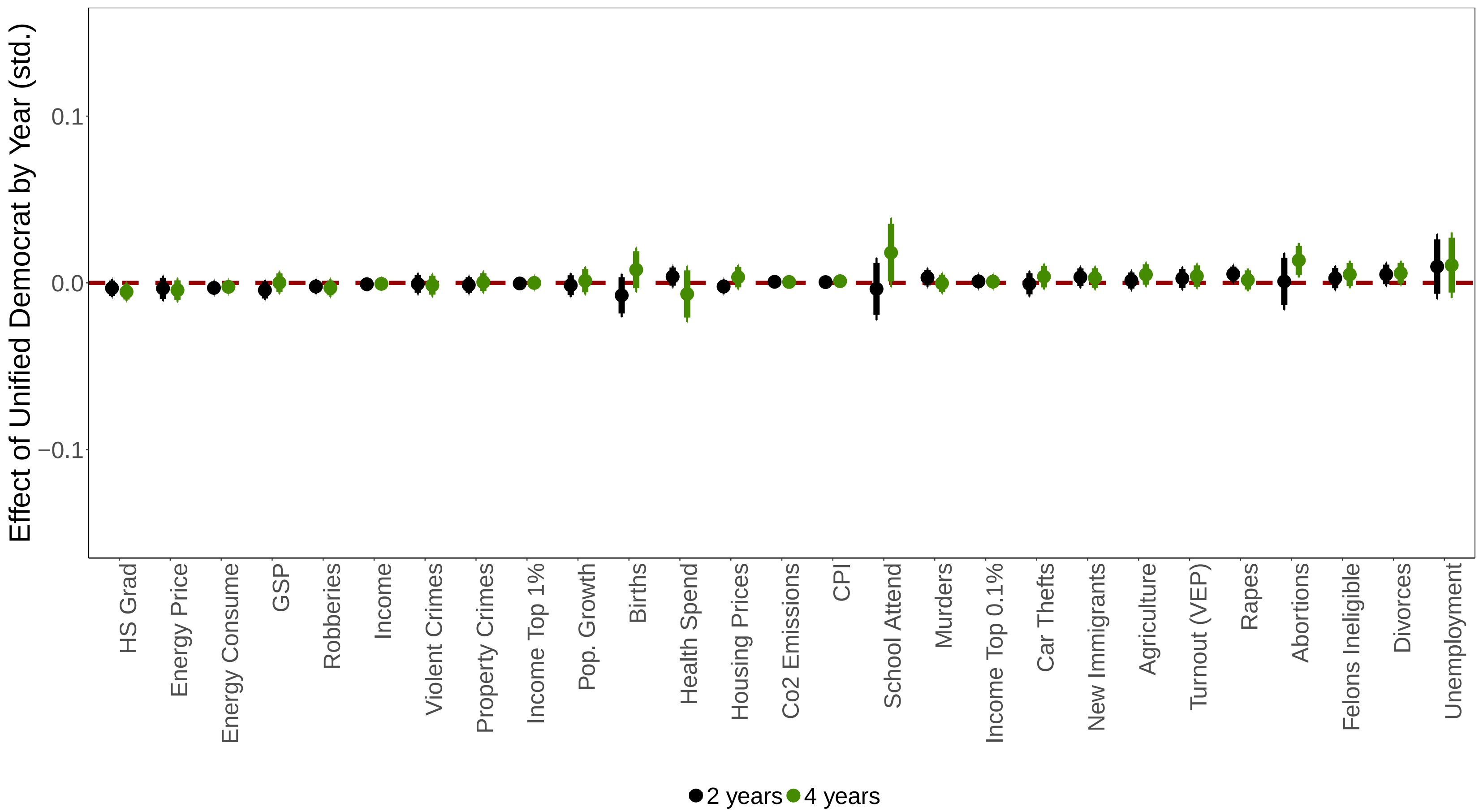


Figure A13 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for unified democratic control compared to unified republican control interacted by a continuous year measure. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

Figure A14: Difference in Difference Estimate of Persistent Unified Democratic Control

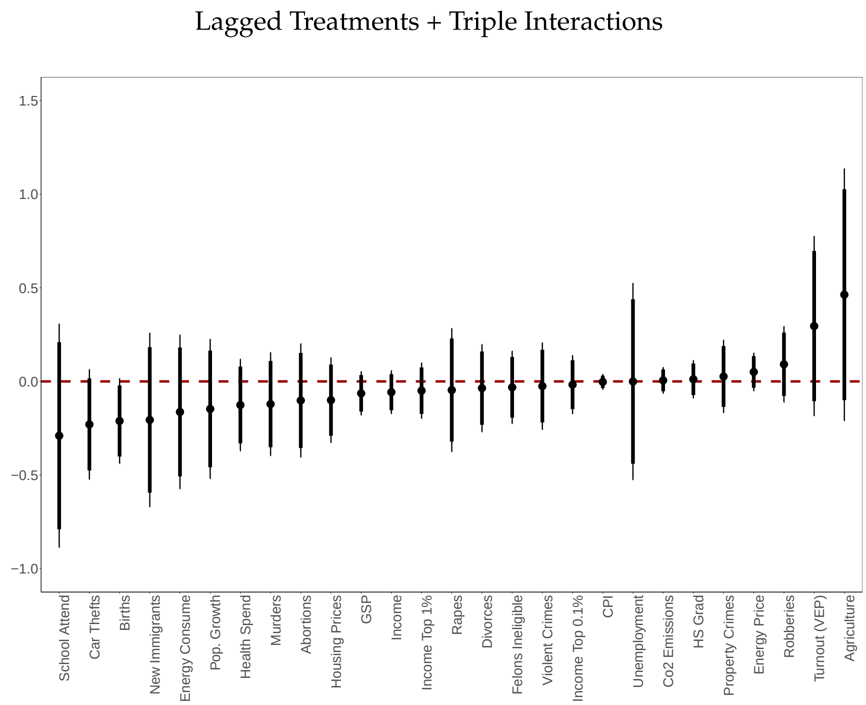


Figure 5 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates for persistent (i.e. control in periods t, $t-3$, and $t-5$) unified democratic control compared to persistent unified republican control. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. This approach uses the triple interaction model described in the paper.

8.3 Full Set of Difference-in-Difference Results for Party Control Elements

In the first year, 11 (5.8%) out of the 189 coefficient estimates (27 outcomes by 7 treatment effects) are significant at the 5% level; only a bit more than we would expect by simple random chance. More importantly, however, none of these clear the Bonferroni or Sidak adjustments for multiple comparisons. In the second year, out of the 189 estimates only 18 (9.5%) are statistically significant at the 5% level. As in the first year, none of the coefficients that appear to show an effect initially actually clear the Bonferroni and Sidak multiple comparisons threshold for statistical significance. This conclusion holds whether we condition on lagged measures of the dependent and treatment variables or not.

When we look at our economic outcomes, it is clear that there is not much of an effect on average income from having a Democratic House or Senate alone or both together. At first glance, there might be a positive effect of having a Democratic governor alone for income. However, this effect does not clear the Bonferroni or Sidak thresholds for multiple-comparison statistical significance. The same holds true for the Democratic Senate and Governor estimates, which is significant (and negative) at the 5% level but not at the multiple comparisons level. The estimate for unified control is positive, but not significant even at the 5% level. Among all the coefficients for income, the median coefficient is 0.07% of a standard deviation. Even if we compare unified Republican (the baseline) to unified Democratic states, the combined effect is 0.14% of a standard deviation.

None of the estimates for the unemployment rate—perhaps the most commonly used single measure of economic well-being—are significant at the 5% level. Moreover, these look to be idiosyncratically dispersed on either side of 0, with the median coefficient being -4.5% of a standard deviation. Simply put, party control appears to have little to no causal effect on unemployment.

The other economic outcomes show a consistent result. Even in cases where there appears to perhaps be a statistically significant effect, the effect sizes are small to modest in size. Based on our 95% confidence intervals, we can rule out large effects with a high degree of certainty. These results are consistent with with little to no effect of Democratic control on the economy.

Our education outcomes show similar results. On the percent graduating high school, two coefficients are significant at the 5% level (Democratic House only and unified control, being both negative). However, neither of these clear the threshold for multiple comparisons. Moreover, these effect sizes tend to be small. Among all the coefficients for the % graduating high school, the median coefficient is 2.7% of a standard deviation. Even if we compare unified Republican (the baseline) to unified Democratic states, the combined effect is only -9.3% of a standard deviation. The same holds true for school attendance rates, with the median effect estimate being 0.9% of a standard deviation and the unified comparison producing a modestly sized coefficient of -18.0% of a standard deviation. Given the lack of statistical significance and generally small to moderate (at best) effect sizes, these results are consistent with no clear effect of Democratic control on the performance of students.

We next turn our attention to two of our family/health outcomes: birth rates and health spending (per capita). Across these, none of the effects clear the 5% level of statistical significance, let alone the threshold for multiple comparisons. Moreover, these effect estimates are all small in size. Among all the coefficients for the birth rate, the median coefficient is 1.1% of a standard deviation. Even if we compare unified Republican (the baseline) to unified Democratic states, the combined effect is only -1.7% of a standard deviation. The same holds true for health spending, with the median effect estimate being 0.6% of a standard deviation and the unified comparison producing a coefficient of 3.0% of a standard deviation. Again, given the lack of statistical significance and small effect sizes, these results are consistent with no clear effect of Democratic control on health/family policy outcomes.

When we look at our difference-in-difference estimates for the effect of party control on two of our environmental outcomes—the level of CO_2 emissions and energy consumed by the commercial sector—the results look similar to our other policy outcomes. Across both of these proxies of environmental well-being, none of the difference-in-difference estimates clear the 5% level of statistical significance, let alone the threshold for multiple comparisons. Among all the coefficients for CO_2 coefficients, the median coefficient is -1.4% of a standard deviation. Even if we compare unified Republican (the baseline) to unified Democratic states, the combined effect is only 0.5% of a standard deviation. The same holds true for commercial energy consumption, with the median effect estimate being -5.6% of a standard deviation and the unified comparison producing a modestly sized coefficient of -11.3% of a standard deviation. Again, given the lack of statistical significance and generally small to moderate (at best) effect sizes, these results are consistent with no clear effect of Democratic control on environmental outcomes.

We next turn our attention to the effect of party control on our two civic outcomes: voter turnout (VEP) and the number of felons ineligible to vote. On the number of felons outcome, one coefficient is negative and significant at the 5% level (Democratic House + Senate). However, this does not clear the multiple comparisons level. The same holds true for the Democratic House alone with voter turnout (and for the Senate alone, which is marginally insignificant that the 5% level). These coefficients appear to aberations, however: most of the coefficients are not close to significant at the 5% level. Moreover, the coefficients are small to modest in size: dispersing themselves idiosyncratically around 0 and averaging about 7.0% of a standard deviation (felons) and 7.6% of a standard deviation (VEP) respectively. These results are consistent with no clear effect of Democratic control on civic outcomes.

Finally, when we look at the effect of party control on two of our criminal justice outcomes—the murder and robbery rates—we find similar null effects. Across both of these measures, none of the difference-in-difference estimates clear the 5% level of statistical significance, let alone the threshold for multiple comparisons. Moreover, the coefficients are small to modest in size: dispersing themselves idiosyncratically around 0 and averaging about 3.7% of a standard deviation (murder) and 8.6% of a standard deviation (robbery) respectively. These results are consistent with no clear effect of Democratic control on the level of crime in the state.

8.4 First Year

Figure A15: Diff/Diff Estimate of Democratic Control on Economic Outcomes (First Year)

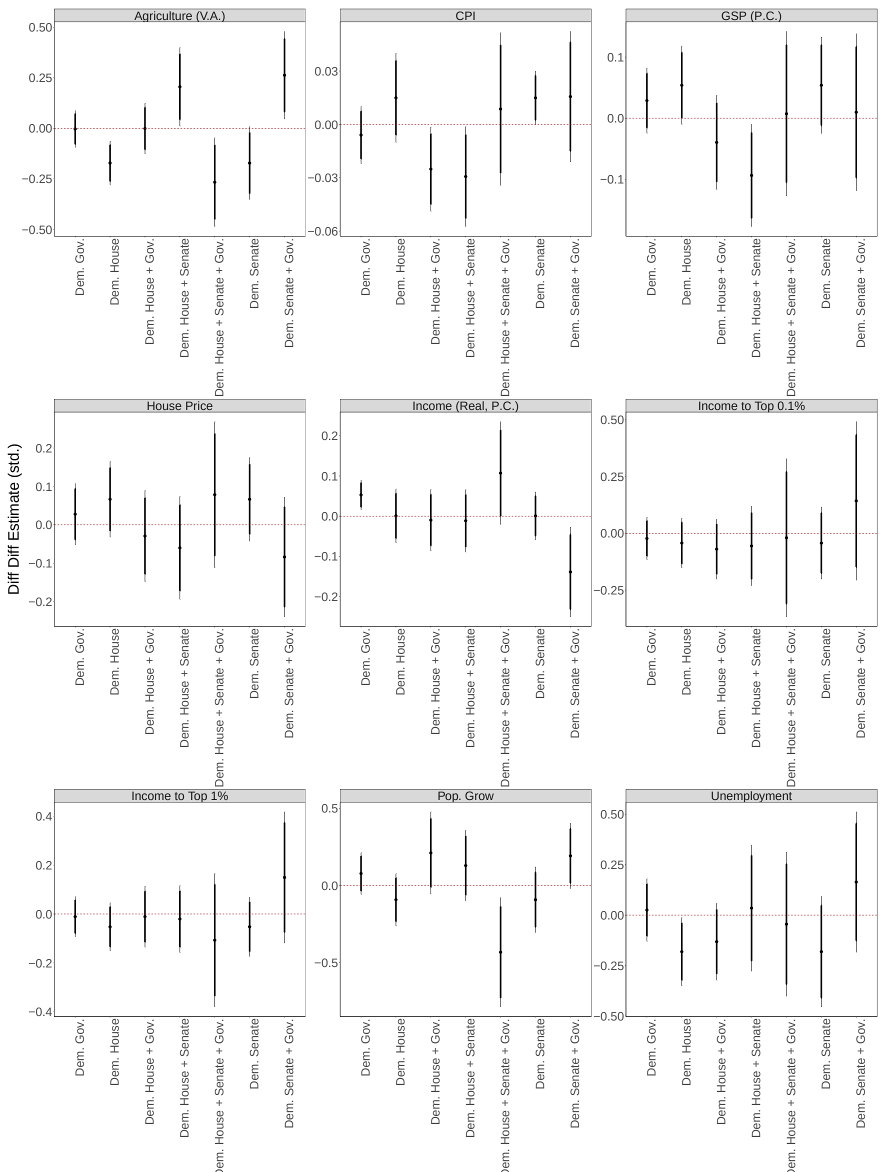


Figure A15 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A16: Diff/Diff Estimate of Democratic Control on Education Outcomes (First Year)

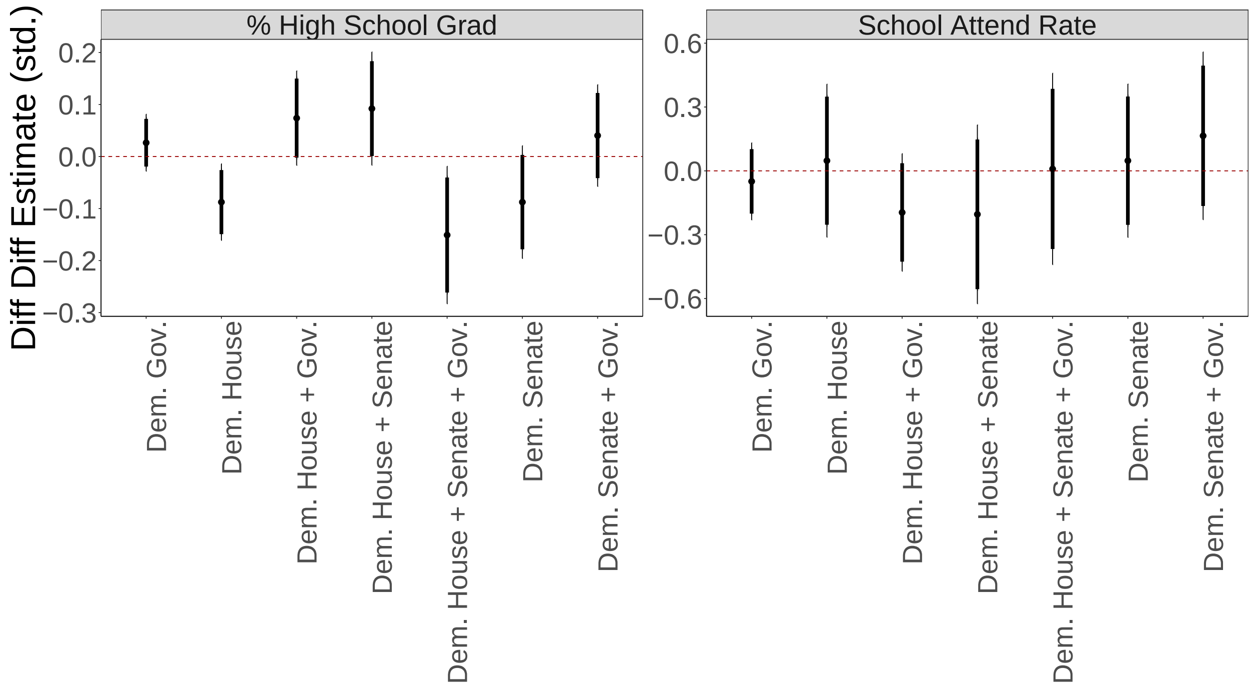


Figure A16 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A17: Diff/Diff Estimate of Democratic Control on Health/Family Outcomes (First Year)

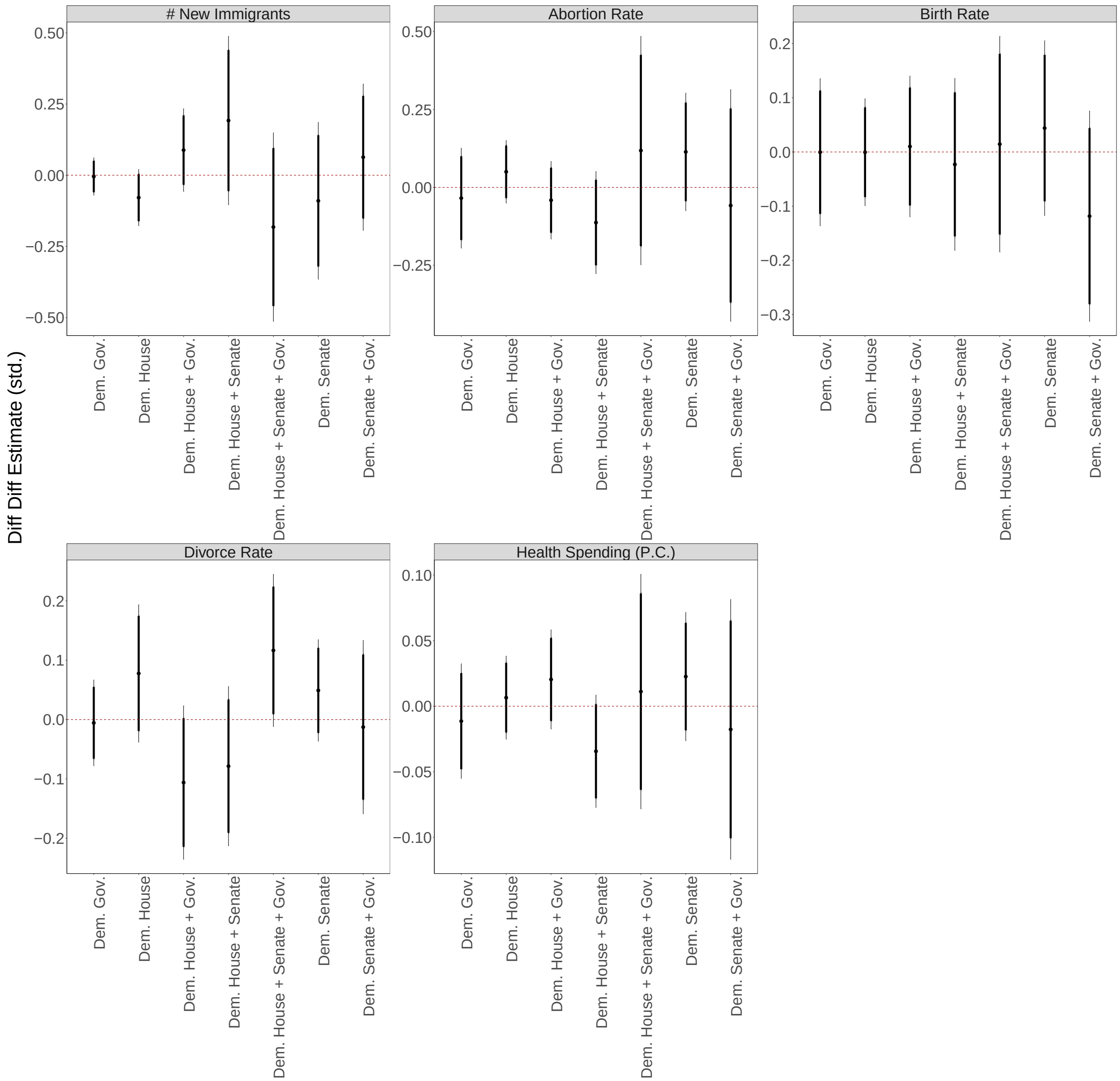


Figure A17 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A18: Diff/Diff Estimate of Democratic Control on Environmental Outcomes (First Year)

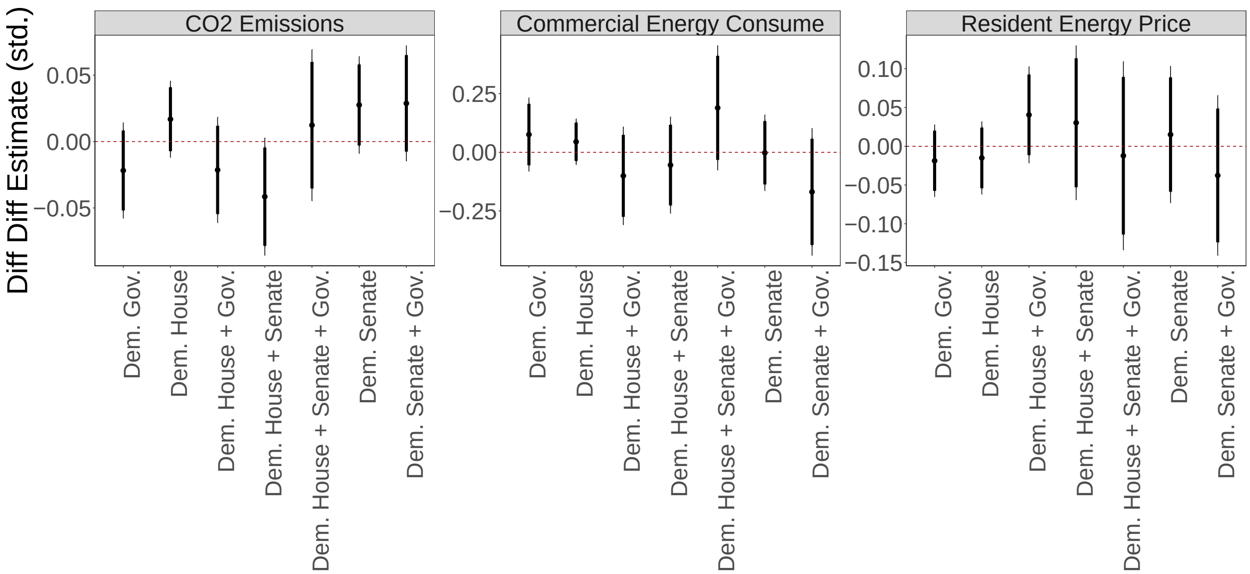


Figure A18 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A19: Diff/Diff Estimate of Democratic Control on Civic Outcomes (First Year)

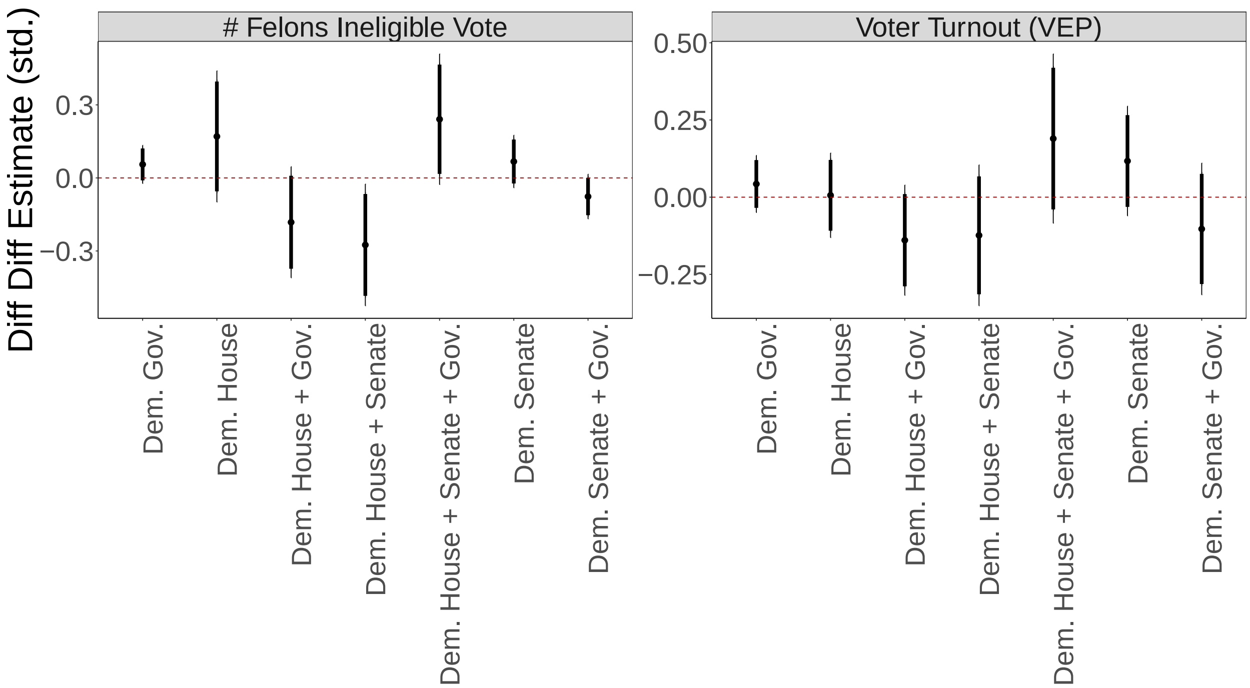


Figure A19 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A20: Diff/Diff Estimate of Democratic Control on Crime Outcomes (First Year)

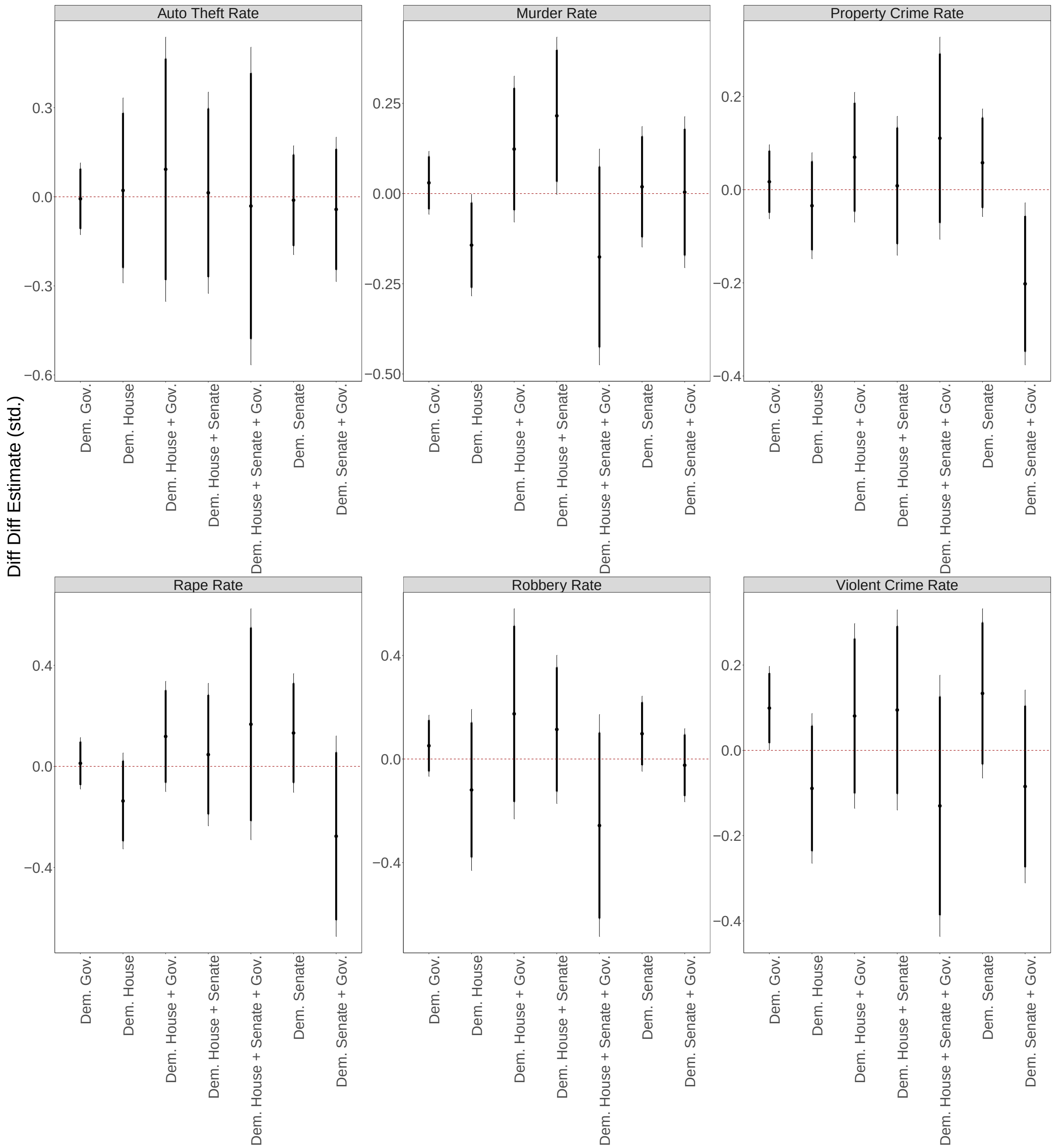


Figure A20 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

8.5 Second Year

Figure A21: Diff/Diff Estimate of Democratic Control on Economic Outcomes (Second Year)

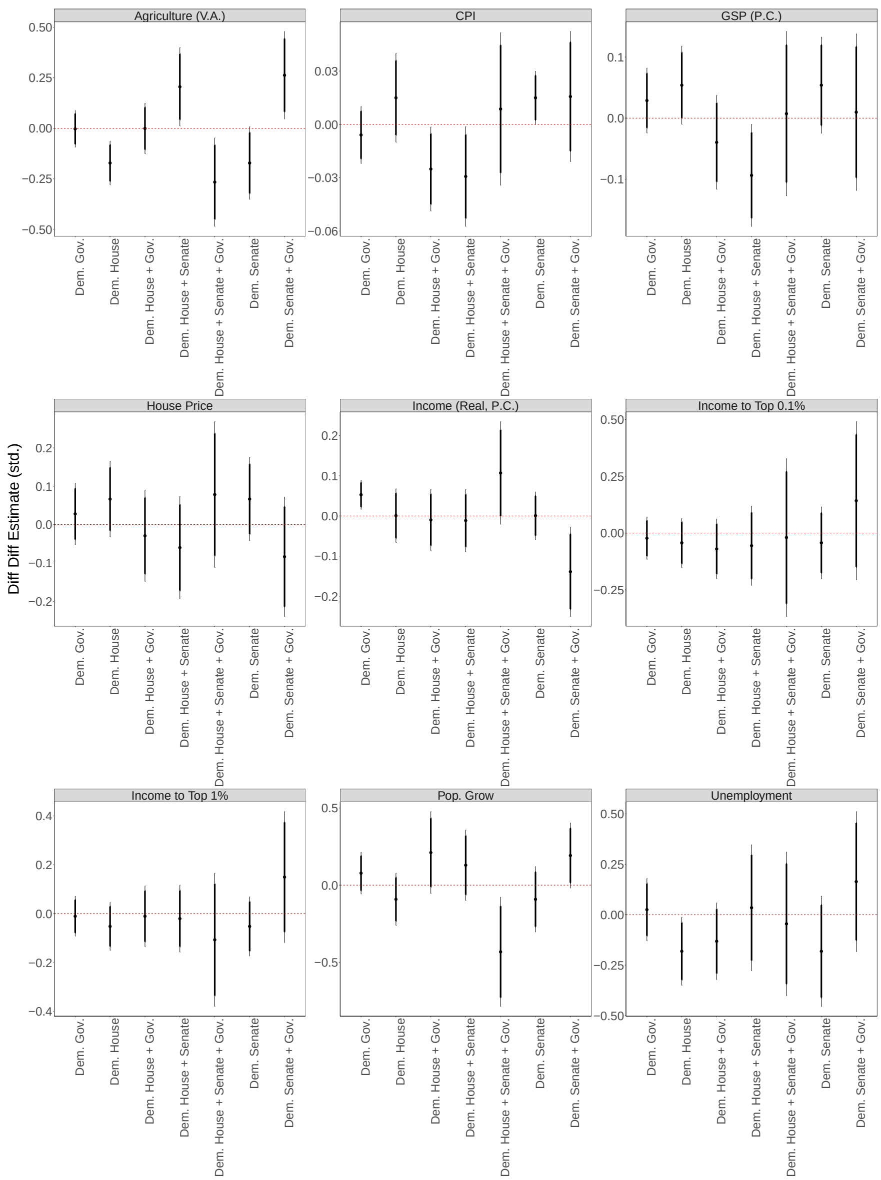


Figure A21 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A22: Diff/Diff Estimate of Democratic Control on Education Outcomes (Second Year)

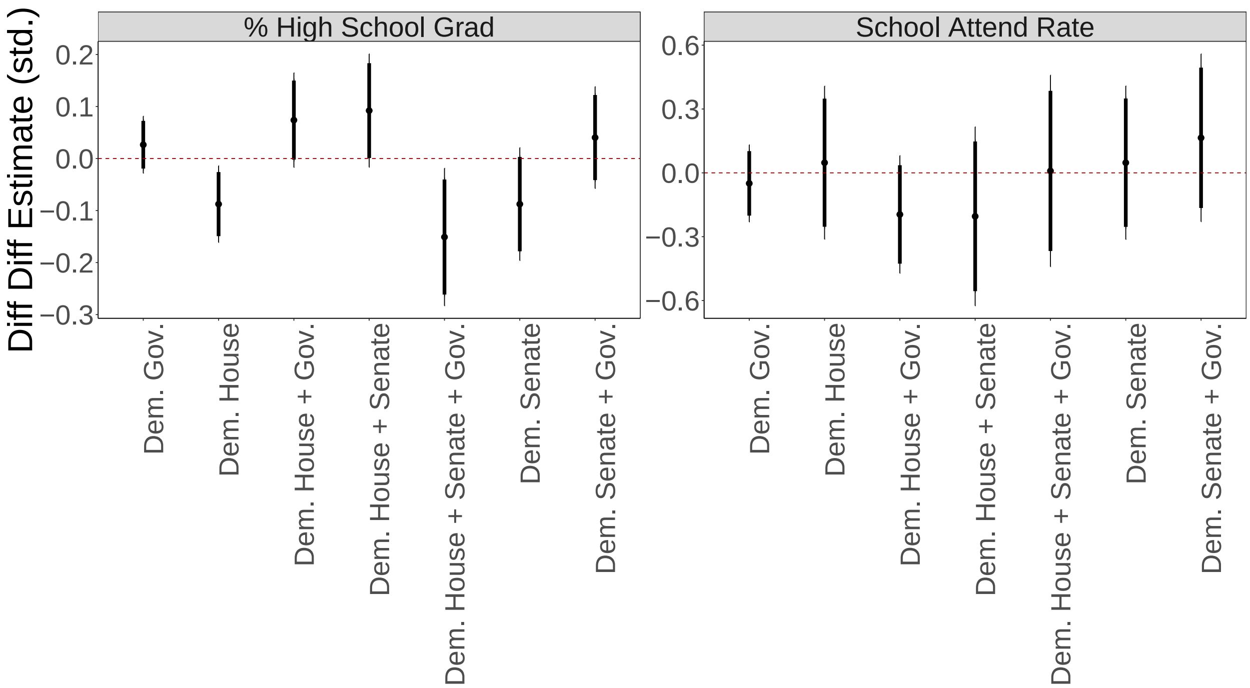


Figure A22 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A23: Diff/Diff Estimate of Democratic Control on Health/Family Outcomes (Second Year)

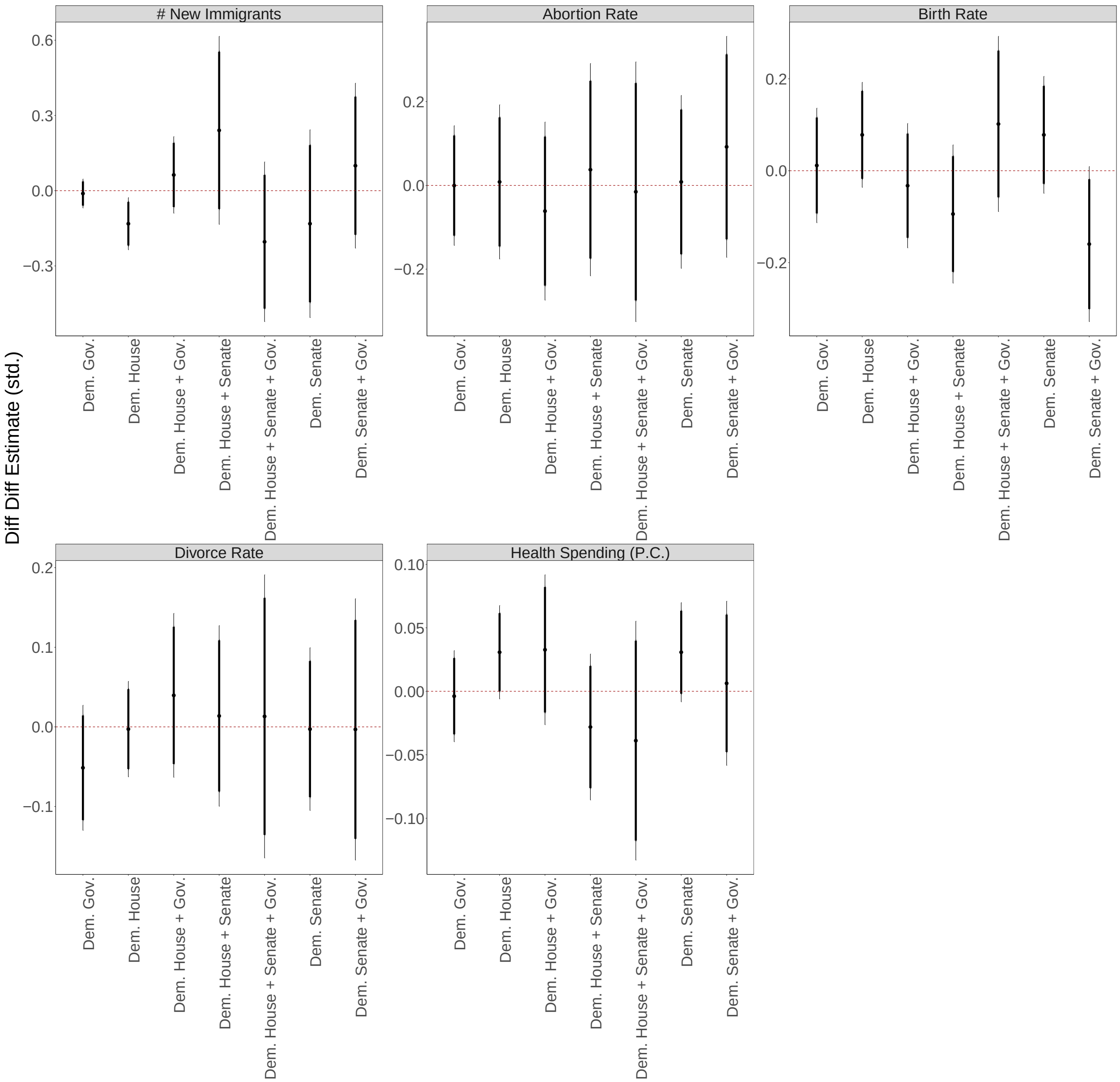


Figure A23 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A24: Diff/Diff Estimate of Democratic Control on Environmental Outcomes (Second Year)

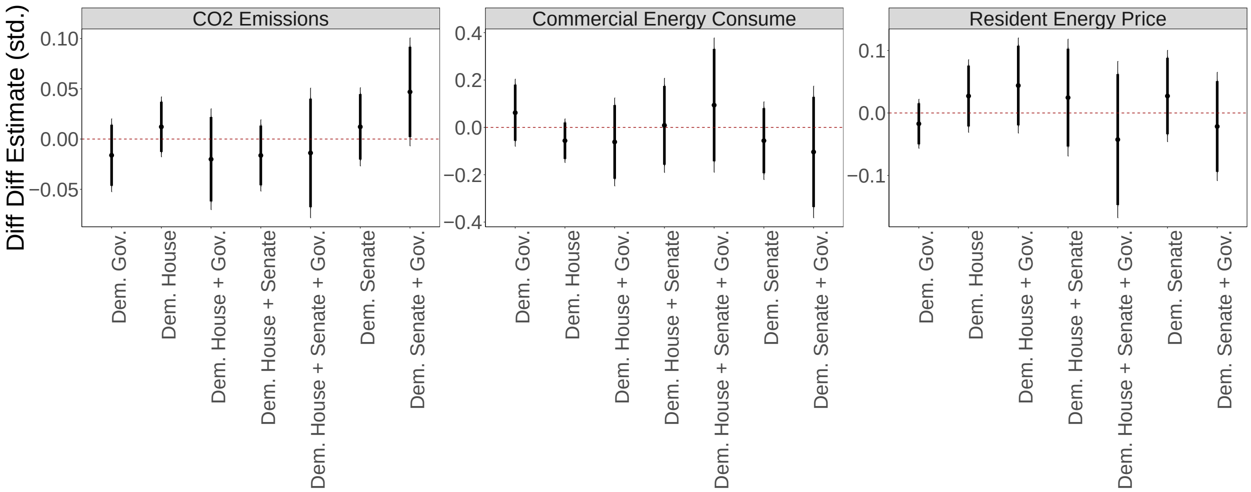


Figure A24 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A25: Diff/Diff Estimate of Democratic Control on Civic Outcomes (Second Year)

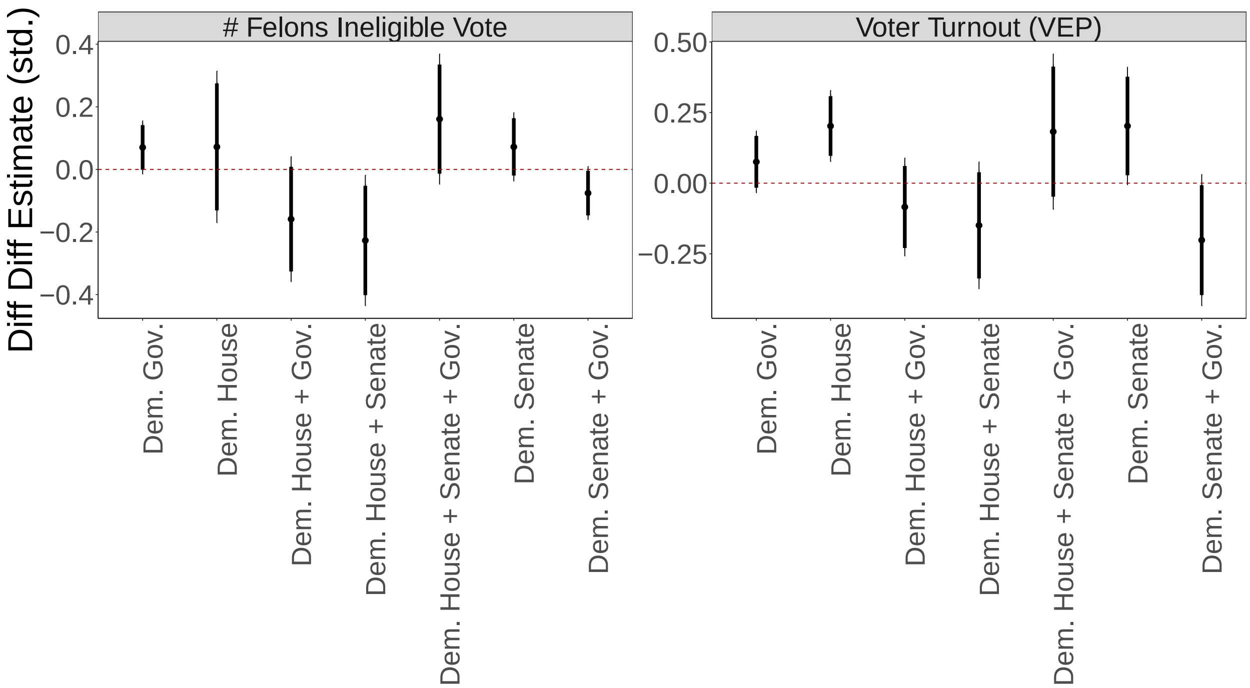


Figure A25 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A26: Diff/Diff Estimate of Democratic Control on Crime Outcomes (Second Year)

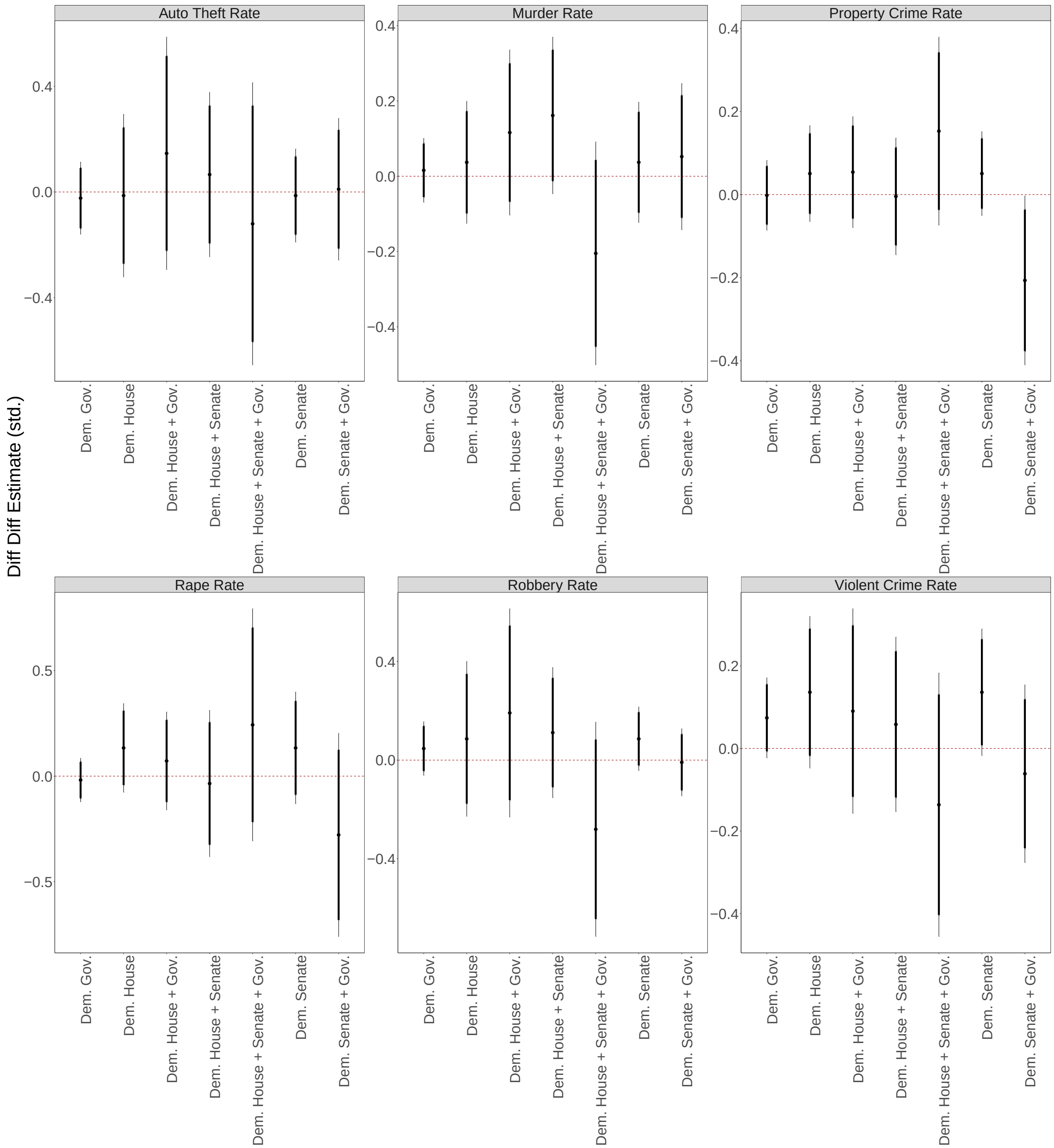


Figure A26 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the difference in difference estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

9 RDD Specification Checks

To explore whether our discontinuity satisfies the conditions necessary to draw causal inferences, this section provides the two standard checks for the validity of a regression discontinuity design. First, table A1 shows tests for covariate balance at the Democrat power discontinuity. Each row displays the results from a separate regression discontinuity model, with lagged versions of these outcomes included as the dependent variable. This test is suggested as a best practice by Eggers et al. (2015) to test for the validity of a discontinuity. The logic is, that if the treatment affects lagged variables, we should be suspicious of the validity of the discontinuity as sorting cases (in this case state legislatures) in an as-good-as random manner. This test is particularly potent for lagged versions of the dependent variables; if lagged versions of these variables are balanced pre-treatment and then the non-lagged versions show effects, we can be even more certain that the discontinuity is estimating an effect that is unbiased from other observed or unobserved factors.

Out of the 81 tests run for lagged measures of our dependent variable (27 outcomes by 3 cutoffs), only 2 (2.5%) are significant at the 5% level: a bit less than what we would expect by chance. Moreover, these two do not clear the significance threshold for multiple hypothesis testing. We also observed balance on lag measures of our key independent variables (party control), with the one exception being an imbalance in gubernatorial control. Our other measures also remain balanced, with factors such as population, citizen ideology and party identification, and various individual demographics remaining balanced at the cutoff. These conclusions hold when we perform the same tests in a multi-cutoff setup; with only 8 (2.7%) out of 294 tests (42 baseline characteristics by 7 treatment estimates) clear the multiple hypothesis testing threshold for statistical significance. This balance across our model specifications helps support the as-good-as random assignment of units to Democratic or Republican control.

We note that our assessment of the validity of the balance around the cutoff with the simple seat count running variable is slightly more optimistic than that provided by Hall, Feigenbaum and Fouirnaies (Forthcoming). As far as we can tell, their main empirical concern about this cutoff is that there is potentially an (not statistically significant) imbalance in lagged treatment status at this cutoff. Here we quote from their paper:

> There are two reasons why we do not use the simple seat share measure as our forcing variable. First, the balance results are not promising. In a parallel exercise to Table 1 (see below), we regress the lag of treatment (majority) status on Democratic seat share in the following election, as well as regressing previous seat share on future seat share. In the first balance test, with majority status as the outcome, we estimate a 10 percentage point increase in the probability of majority status with a standard error of 0.10. In the second balance test, with previous seat share as the outcome, we estimate a treatment effect of 1.46 points with a standard error of 1.34. Though neither of these effects are significant at conventional levels, both appear far from balanced at the discontinuity.
>
> —Hall, Feigenbaum and Fouirnaies (Forthcoming, 281-282)

We see a similar coefficient in our data. However, we think this is not a large concern. When viewed in context with the host of the additional variables we explore at the cutoff, the seat share appears to be much more balanced. In short, we are not as worried about one (potentially idiosyncratic) non-significant potential imbalance. For our purposes, this specification of the running variable appears to be well suited.

That being said, we agree with Hall, Feigenbaum and Fouirnaies (Forthcoming) that the seat share running variable potentially throws away information about the context of electoral races underlying the party control cutoff (their second reason discussed on page 282). It is for that reason, we estimate all of our regression discontinuity models with their running variables (see Figure A56). When we do so, we find similar results—perhaps ones that suggest an even smaller effect on policy outcomes. Hence, we think it important to highlight while their approach for specifying the running variable is highly useful, it may not be universally needed to correctly identify a causal effect with the party control cutoff.

Table A1: Balance in Lagged Outcomes at the Party Control Discontinuity

	Variable	Years Available	House β_{lag}	House p	Senate β_{lag}	Senate p	Governor β_{lag}	Governor p
Ed.	% High School Diploma	1975-2006	1.71	0.48	-0.26	0.91	3.70	0.02
	Average School Attendance Rate	1986-2009	-0.84	0.48	-0.34	0.66	-0.42	0.72
Economy	Real Per Capita Personal Income	1960-2010	1096	0.55	-469.86	0.77	1568.2	0.31
	Population Growth	1930-2012	0.00	0.74	0.00	0.30	0.00	0.80
	Consumer Price Index	1960-2007	0.03	0.72	0.04	0.56	0.05	0.45
	Quarterly Housing Price Index	1975-2012	0.34	0.32	-0.09	0.68	0.18	0.55
	Gross State Product Per Capita	1987-2010	4867.2	0.22	226.15	0.93	6221.3	0.05
	Value Added by Agricultural Sector	1950-2011	-503039	0.53	-293748	0.61	-95627	0.88
	Unemployment rate	1975-2004	0.43	0.52	0.23	0.72	-1.13	0.06
	Fraction Income top 1%	1960-2013	0.29	0.83	0.22	0.84	1.16	0.30
	Fraction Income top 0.1%	1960-2013	0.26	0.75	0.19	0.79	0.62	0.33
Health/Fam.	Health Spending Per Capita	1991-2009	1400	0.14	42.56	0.92	329.92	0.51
	Abortion Rate	1975-1996	-1.03	0.83	-0.12	0.97	-7.69	0.03
	Divorce Rate	1975-2004	0.17	0.79	0.81	0.13	0.39	0.45
	Birth Rate	1991-2008	-0.08	0.37	0.04	0.59	0.10	0.13
	New Immigrant Green Card Holders	1988-2011	376.68	0.97	-21064	0.22	-14829	0.35
Crime	Property Crime Rate	1960-2009	-467.31	0.35	349.94	0.31	-17.46	0.96
	Rape Rate	1960-2009	-3.01	0.57	4.71	0.22	0.28	0.94
	Robbery Rate	1960-2009	-3929.2	0.44	-12114	0.05	-7570.9	0.12
	Violent Crime Rate	1960-2009	-40.26	0.63	-24.52	0.73	-65.83	0.33
	Car Theft Rate	1960-2009	-52.38	0.38	-18.14	0.74	-49.15	0.42
	Murder Rate	1960-2009	0.19	0.85	0.25	0.79	-1.59	0.13
Civic	Voter Turnout (VEP)	1980-2012	0.04	0.16	-0.02	0.70	0.00	0.85
	# Felons Ineligible to Vote	1980-2012	-95967	0.12	-17046	0.54	9857.7	0.57
Enviro.	Commercial Sector Energy Consume	1960-2009	-188.74	0.96	478.92	0.84	-252.02	0.95
	Residential Sector Energy Price	1970-2010	0.46	0.82	-0.26	0.89	3.20	0.16
	CO2 emissions (metric tons)	1960-2001	-33.80	0.46	-94.91	0.07	-35.63	0.23
IVs	Democratic Majority Status (H)	1961-2010	-0.04	0.73	0.00	0.99	0.00	0.99
	Democratic Majority Status (S)	1962-2010	0.10	0.38	0.00	0.98	0.02	0.82
	Democratic Majority Status (G)	1962-2010	0.10	0.53	0.04	0.72	0.40	0.00
Other	Population	1900-2008						
	Citizen Ideology Score	1960-2013	1.12	0.78	-3.50	0.29	-2.24	0.64
	Citizen Party Identification Score	1976-2011	0.05	0.10	0.01	0.79	-0.03	0.42
	Citizen Ideology Score (weighted)	1976-2011	0.00	0.99	0.00	0.94	-0.06	0.13
	% Students Limited English	1999-2013	-0.01	0.85	-0.02	0.34	0.03	0.14
	% Students Disability	1988-2013	0.02	0.40	0.00	0.94	0.02	0.06
	% Students Free/Reduced Lunch	1999-2013	-0.05	0.36	0.06	0.27	0.02	0.66
	% Students Male	1999-2013	0.16	0.18	0.00	0.98	0.03	0.43
	% Students Hispanic	1993-2013	-0.08	0.17	-0.02	0.57	0.03	0.55
	% Students Black	1993-2013	-0.12	0.08	0.03	0.51	-0.06	0.23
	% Students White	1993-2013	0.11	0.23	-0.03	0.67	-0.05	0.39
	% Students Attend	1996-2009	-0.02	0.30	0.00	0.81	0.02	0.22

Estimates come from regression discontinuity models with the optimal bandwidth, local non-parametric specification of the running variable, bias-correction, and robust intervals as specified by the `rdrobust` command in `Stata` created by Calonico, Cattaneo and Titiunik (2014). Columns labeled β_{lag} provide the RDD coefficient estimate for the lagged measures; columns labeled p provide the p-value for the coefficient estimate. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Lagged IVs come from two years previous.

The second specification test recommended is the precise sorting test provided by McCrary (2008).

Precise sorting occurs when observations—in this case state legislative bodies—are able to rampantly manipulate their score on the running variable (Lee and Lemieux 2010). If this were to occur, the discontinuity would lose it's as-good-as random assignment. To test this possibility, McCrary (2008) recommends looking for clusters of observations around the cutoff. The logic is, if observations are able to manipulate what side of the cutoff they fall on, we should be able to see this by a discontinuity in the number of observations at the cutoff.

Figure A27 plots the distribution of legislatures at the party power discontinuity for the three cutoffs. As can be seen, the distribution of legislatures is relatively smooth at the cutoff for the Governor and House; in these, neither party appears to dominate scenarios close to the cutoff. In the Senate, however, there is some evidence of a discontinuity at the cutoff. We note two things about this imbalance. First, the McCrary Density Check is inherently limited. As McCrary himself notes, "a running variable with a continuous density is neither necessary nor sufficient for identification" except for strong auxiliary assumptions (2008, 701). In addition, the McCrary Density Check has not been generalized to situations where multiple running variables are used. Put differently, with multiple cutoffs, the expectations for balance across all of these are less straightforward. For these reasons, we take the position that the covariate balance checks just shown offer a more informative check for precise sorting. Given overwhelming balance, we deem precise sorting to be unlikely. Second, in attempt to address any potential for precise sorting, we run the recommended so called "donut RD" check (Barreca, Lindo and Waddell 2016). When we do so, the conclusions presented in the paper do not change. Taken with the results from the covariate balance test, this check is assuring that the state legislature party power discontinuity sorts states in an as-good-as random manner. This allows for these cutoffs to be used to estimate the causal impact of party control on policy outcomes.

Figure A27: McCrary Density Check for Precise Sorting

(a) Governor (b) House (c) Senate

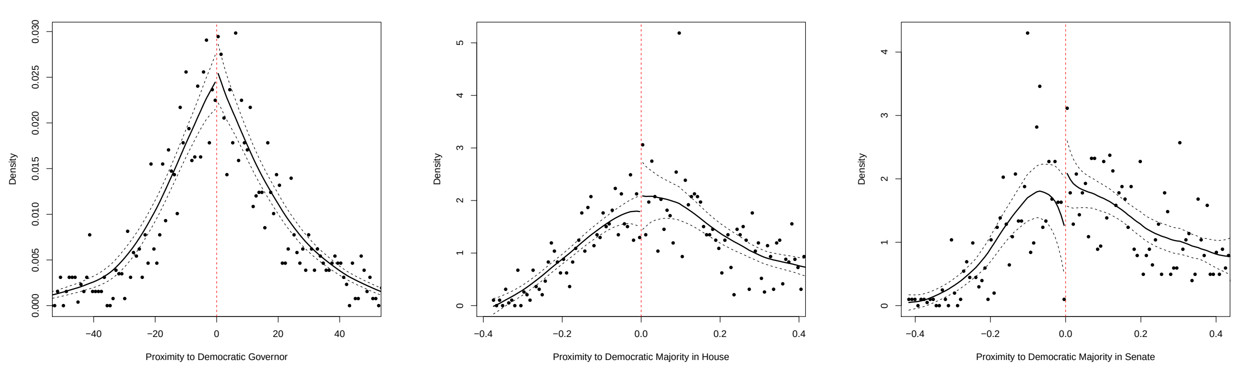


Figure A27 displays the McCrary Density Test for precise sorting (McCrary 2008). The x-axis displays the running variable for these three individual cutoffs. Corresponding p-values for H_0 = continuity at the cutoff: Governor = 0.59, House = 0.15, Senate = 0.00.

10 Supplemental Single-Cutoff RDD Results

Figure A28: Single Cutoff RDD Effect of Democratic Control on Policy Outcomes (Second Year)

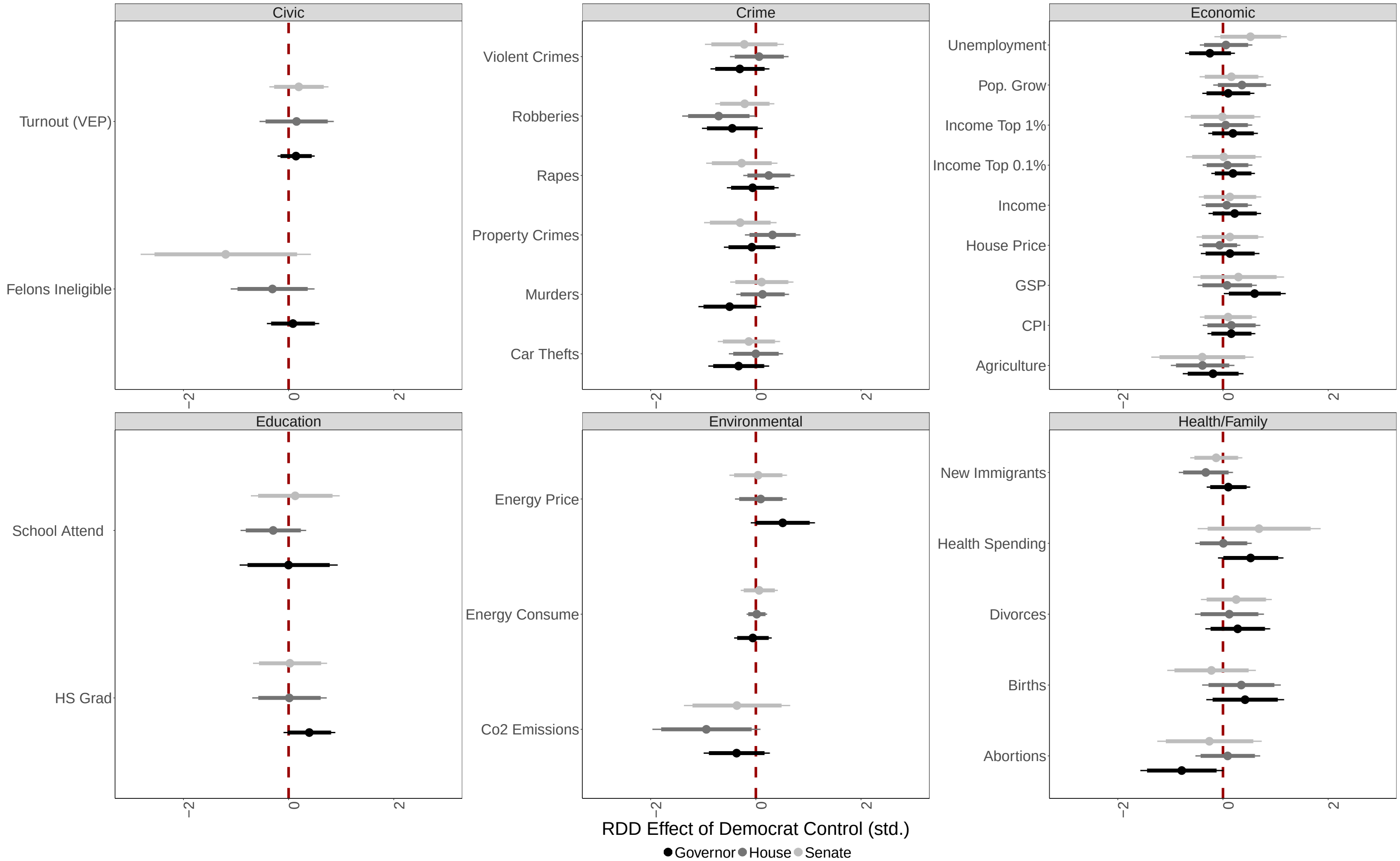


Figure A28 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the regression discontinuity estimates. The estimates are broken by the chamber that switches power. The running variable is modeled with a local kernel smoothed function. The estimates use the optimal bandwidth as specified by the `rdrobust` command in `STATA` created by Calonico, Cattaneo and Titiunik (2014). Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

11 RDD for Unified Control

Figure A29: RDD Estimates of Unified Democratic Control Compared to Unified Republican Control/Divided Gov't

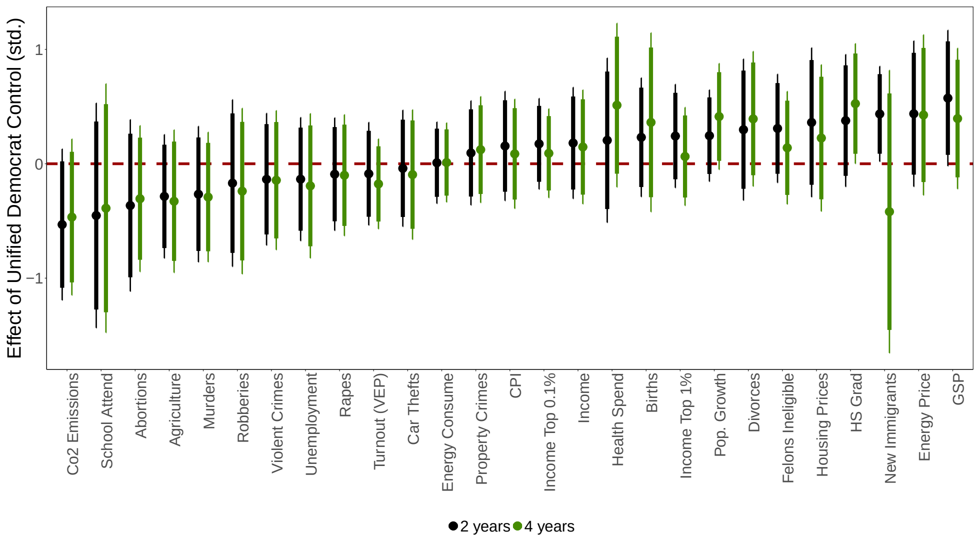


Figure A29 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the regression discontinuity estimates for unified democratic control compared to unified republican control and divided government. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

Figure A30: RDD Estimates of Unified Democratic Control Compared to Unified Republican Control/Divided Gov't

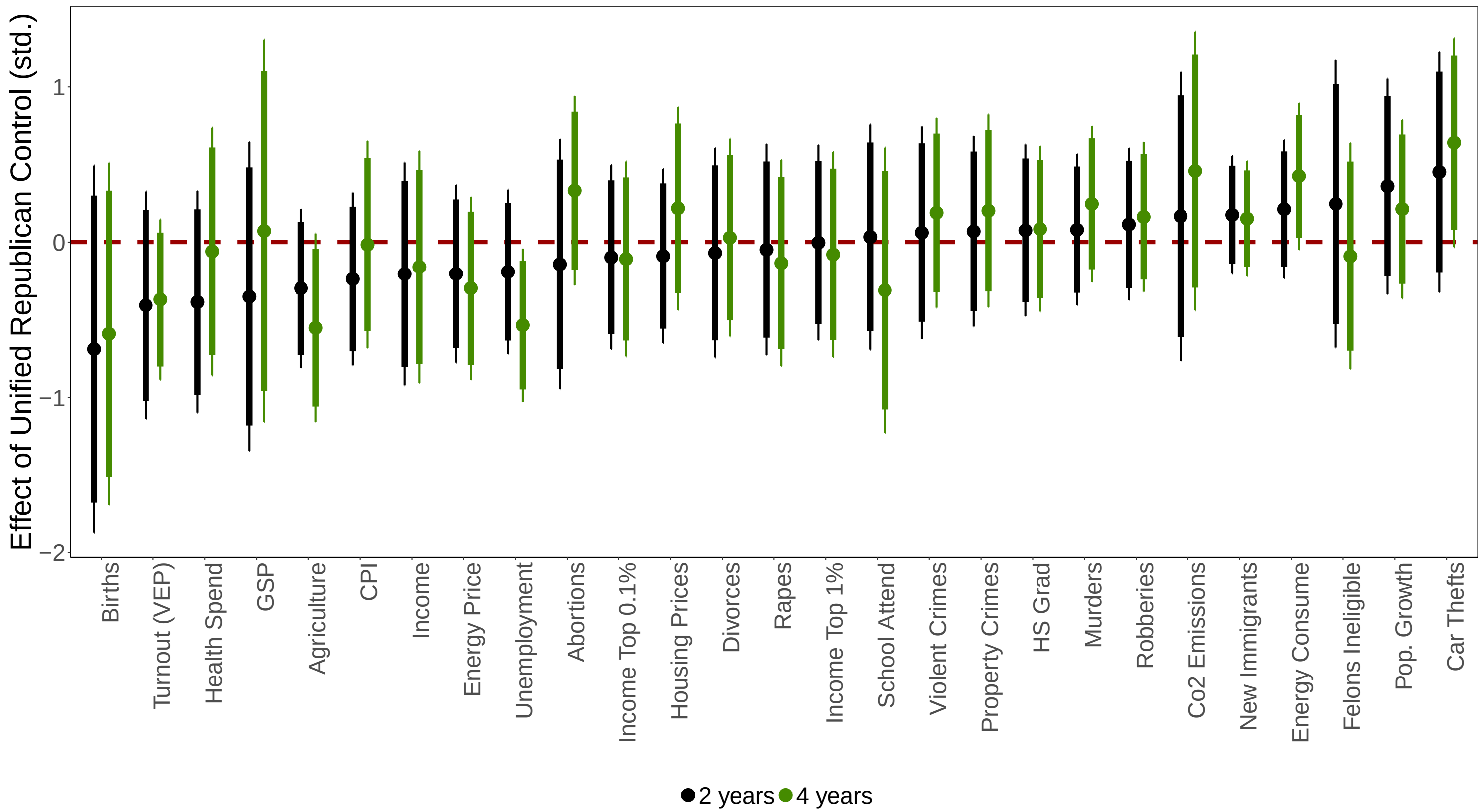


Figure A30 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the regression discontinuity estimates for unified republican control compared to unified democratic control and divided government. Coefficients are sorted from smallest to largest for year 2 effects. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level.

12 Full Set of Multi-Cutoff RDD Results

12.1 First Year

Figure A31: MRDD Estimate of Democratic Control on Economic Outcomes (First Year)

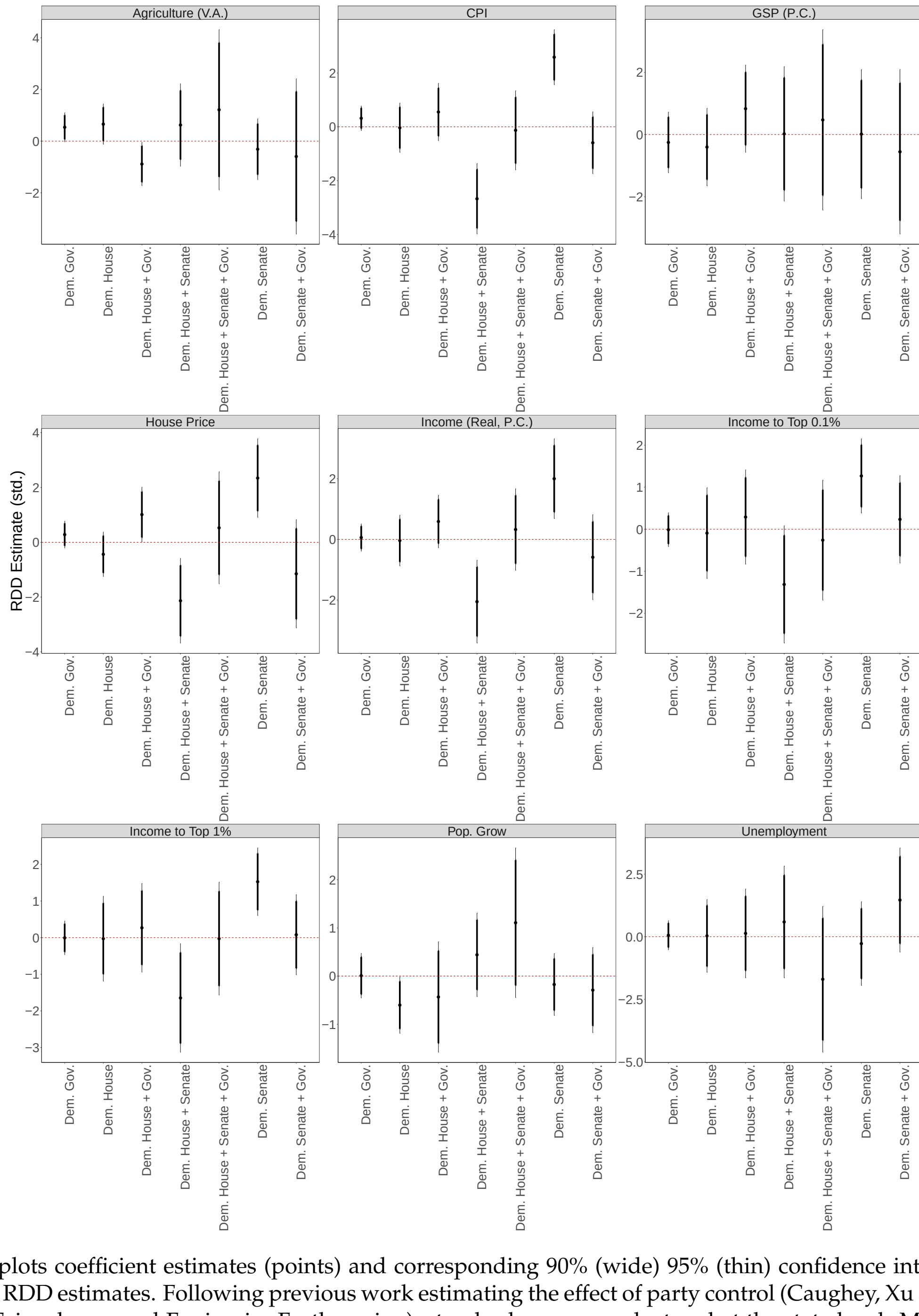


Figure A31 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A32: MRDD Estimate of Democratic Control on Education Outcomes (First Year)

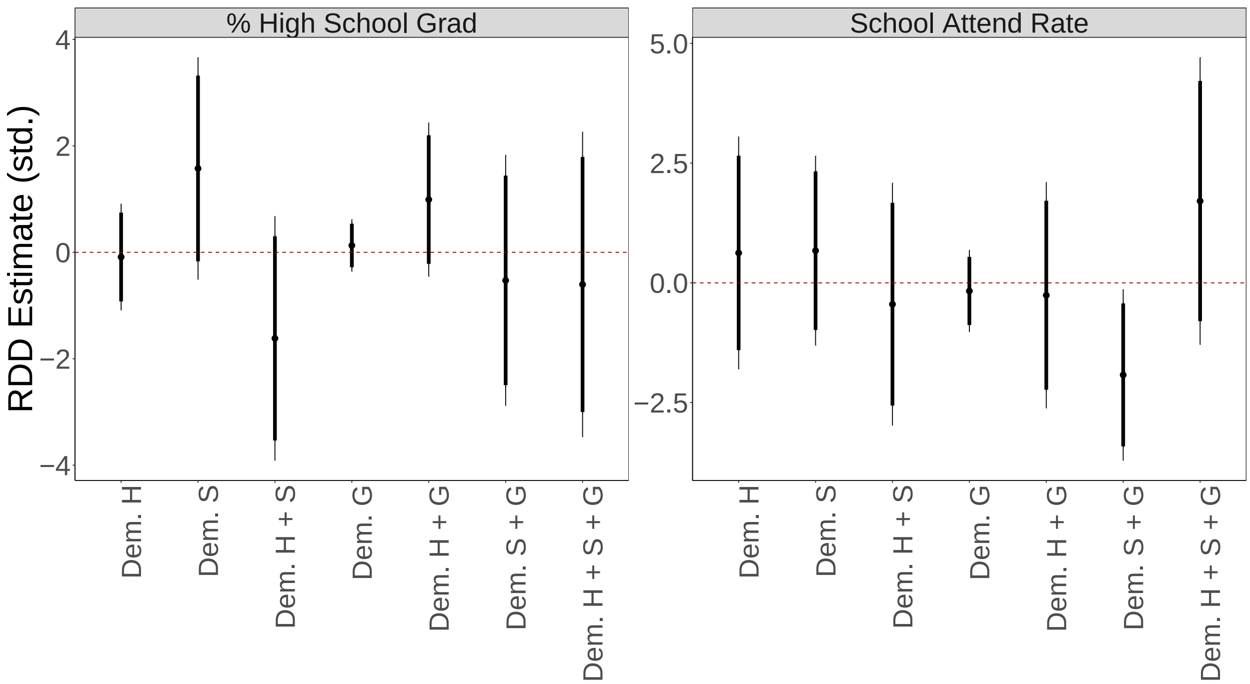


Figure A32 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A33: MRDD Estimate of Democratic Control on Health/Family Outcomes (First Year)

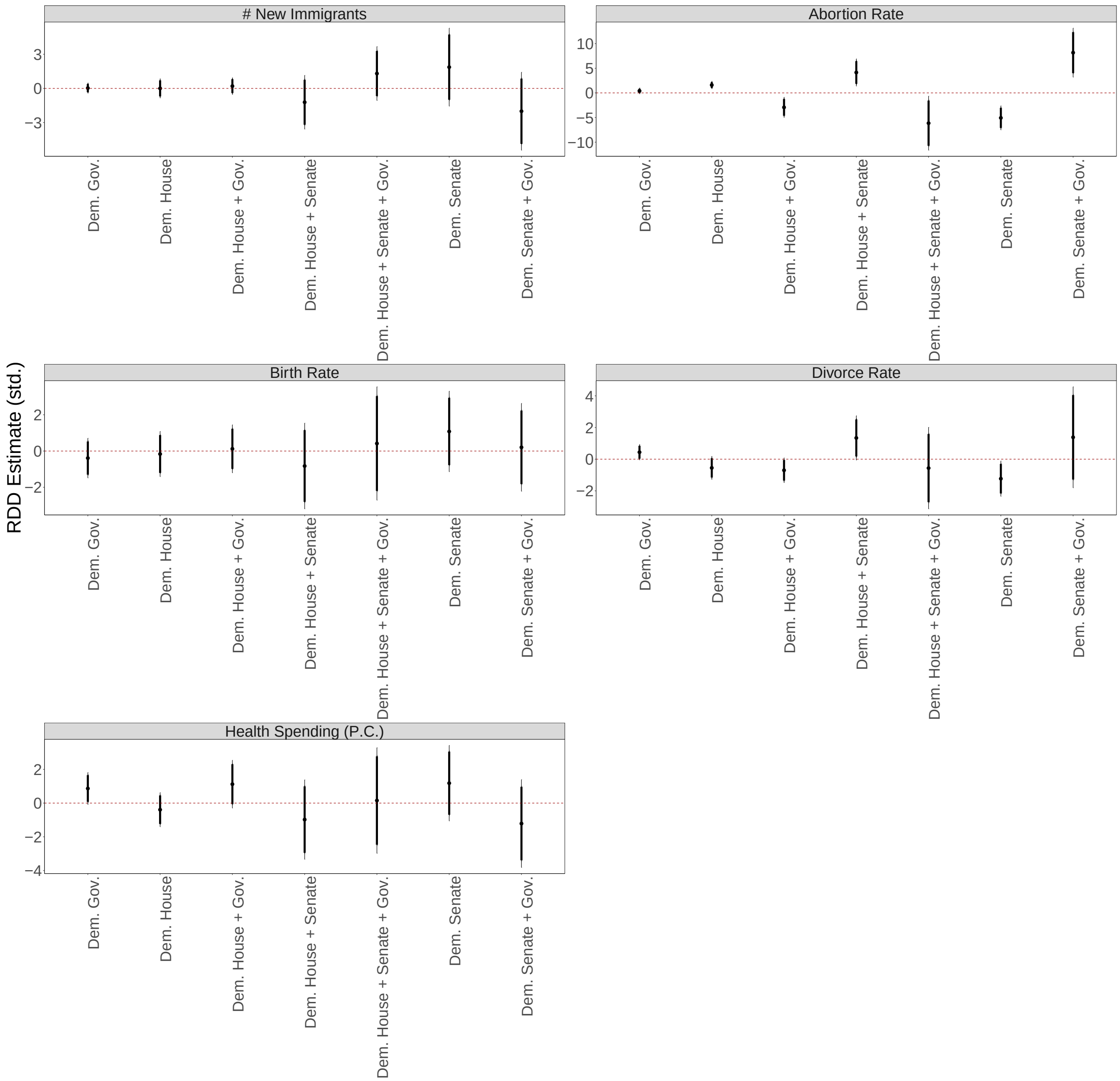


Figure A33 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A34: MRDD Estimate of Democratic Control on Environmental Outcomes (First Year)

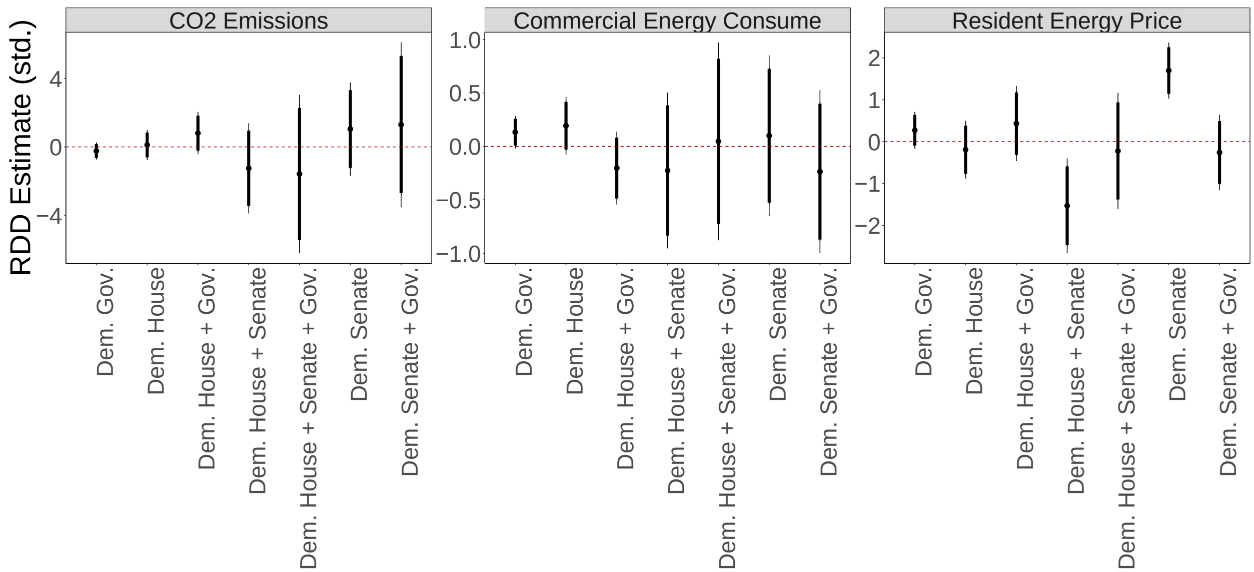


Figure A34 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A35: MRDD Estimate of Democratic Control on Civic Outcomes (First Year)

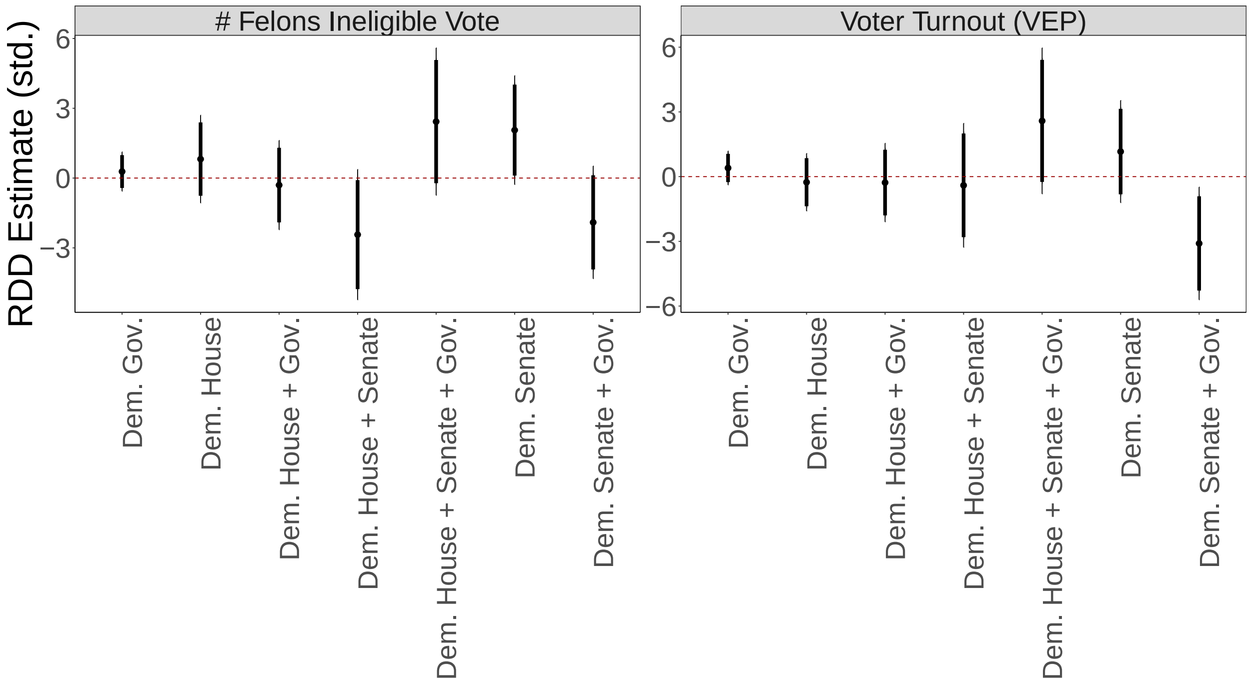


Figure A35 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A36: MRDD Estimate of Democratic Control on Crime Outcomes (First Year)

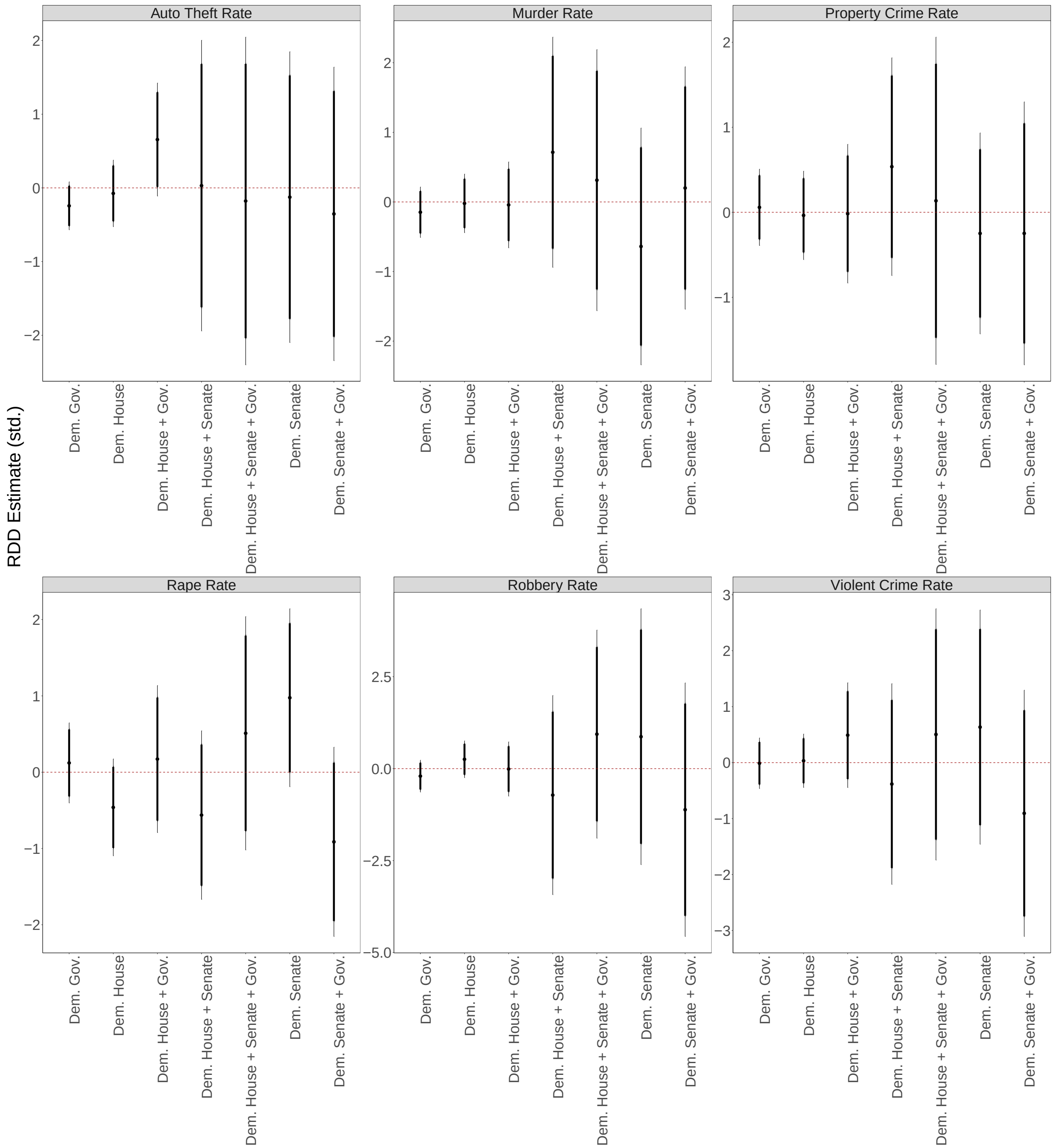


Figure A36 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

12.2 Second Year

Figure A37: MRDD Estimate of Democratic Control on Economic Outcomes (Second Year)

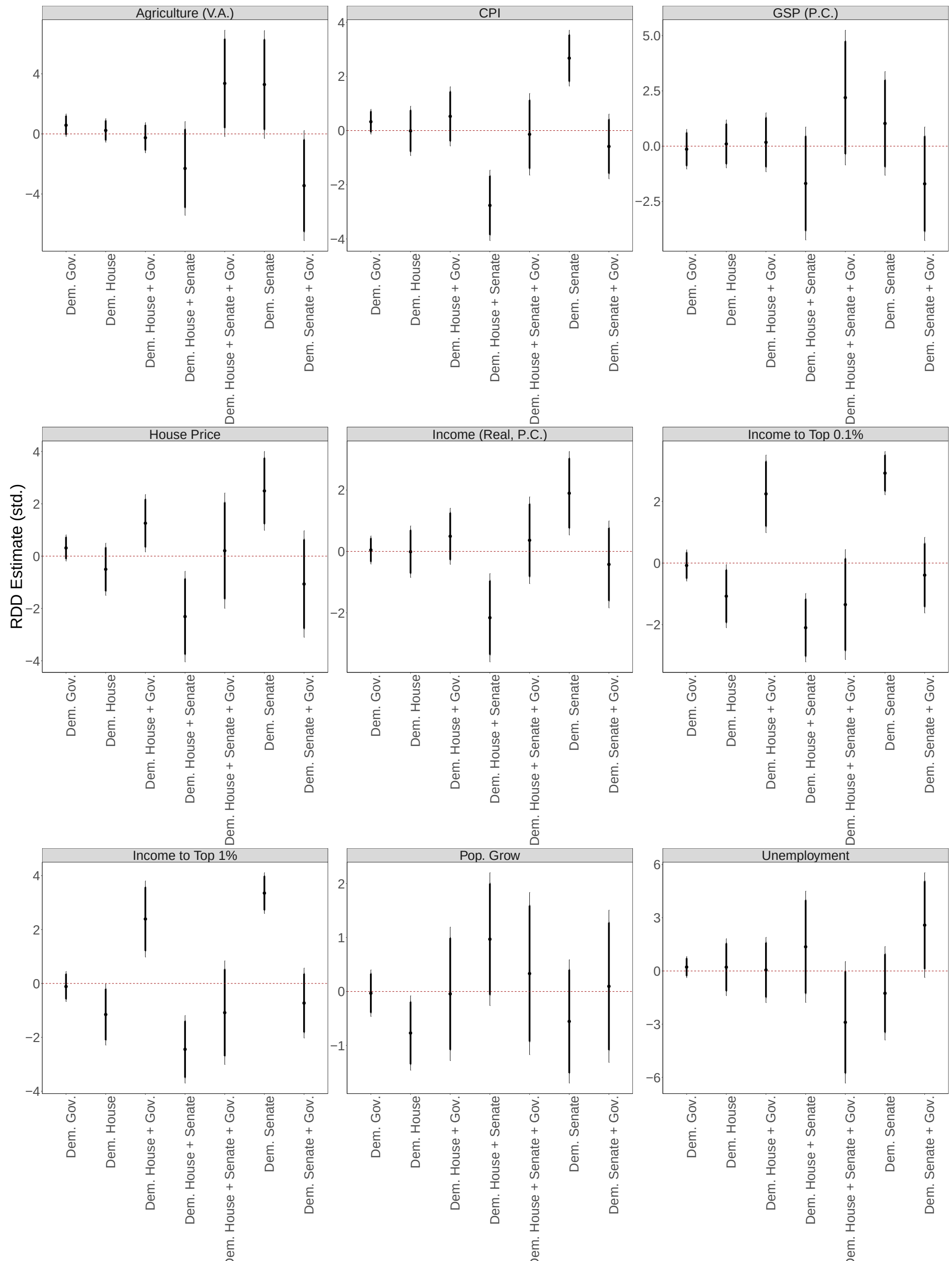


Figure A37 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A38: MRDD Estimate of Democratic Control on Education Outcomes (Second Year)

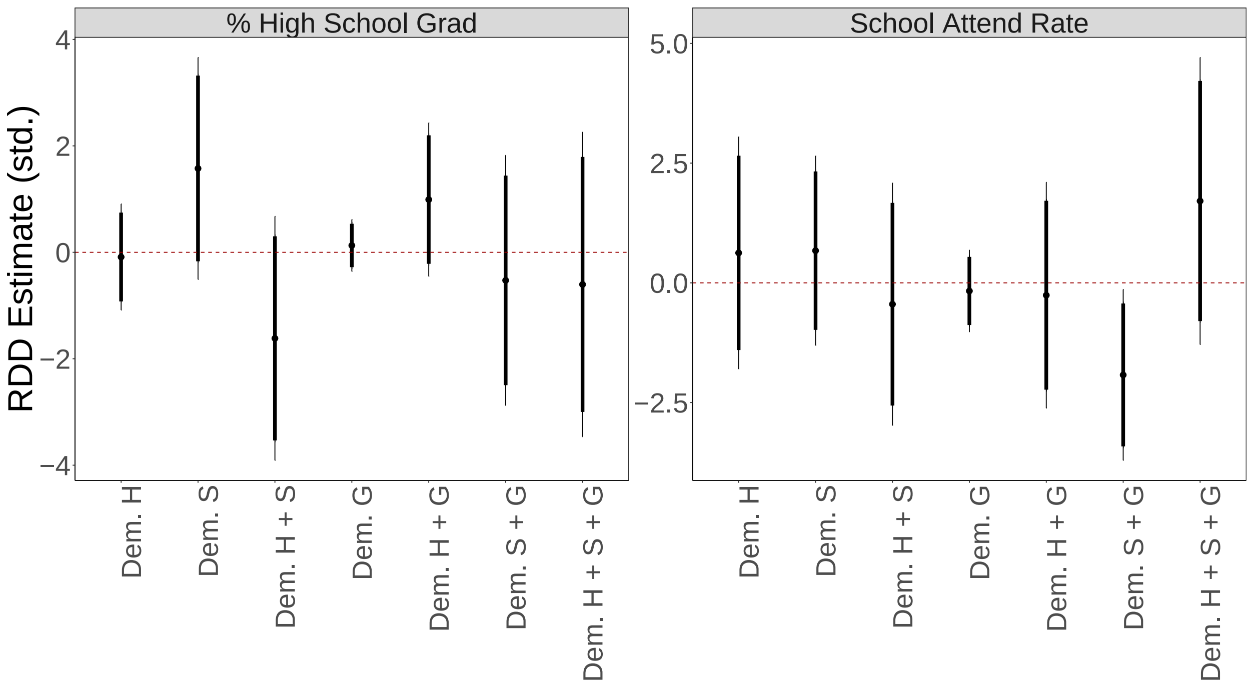


Figure A38 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A39: MRDD Estimate of Democratic Control on Health/Family Outcomes (Second Year)

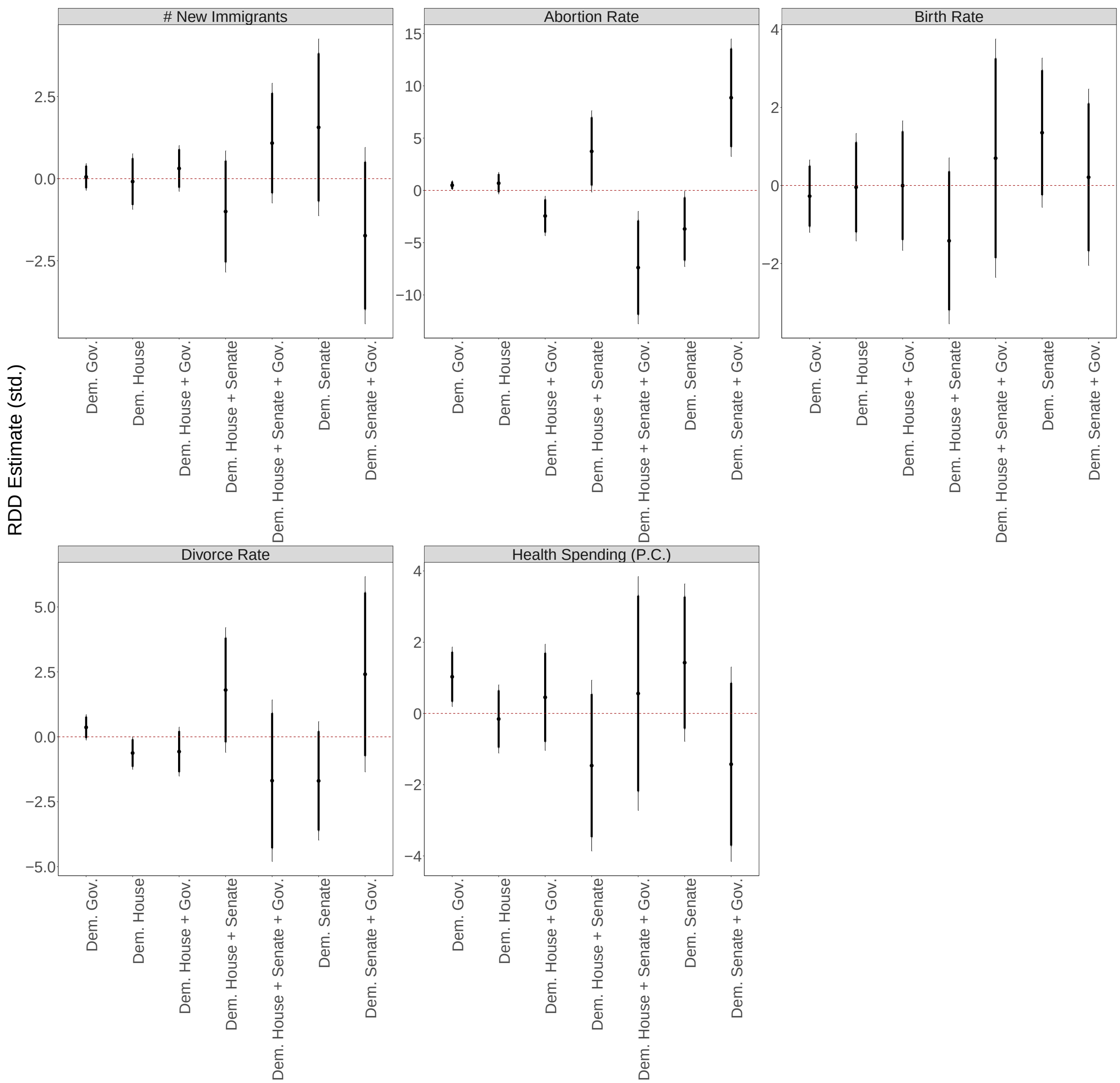


Figure A39 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A40: MRDD Estimate of Democratic Control on Environmental Outcomes (Second Year)

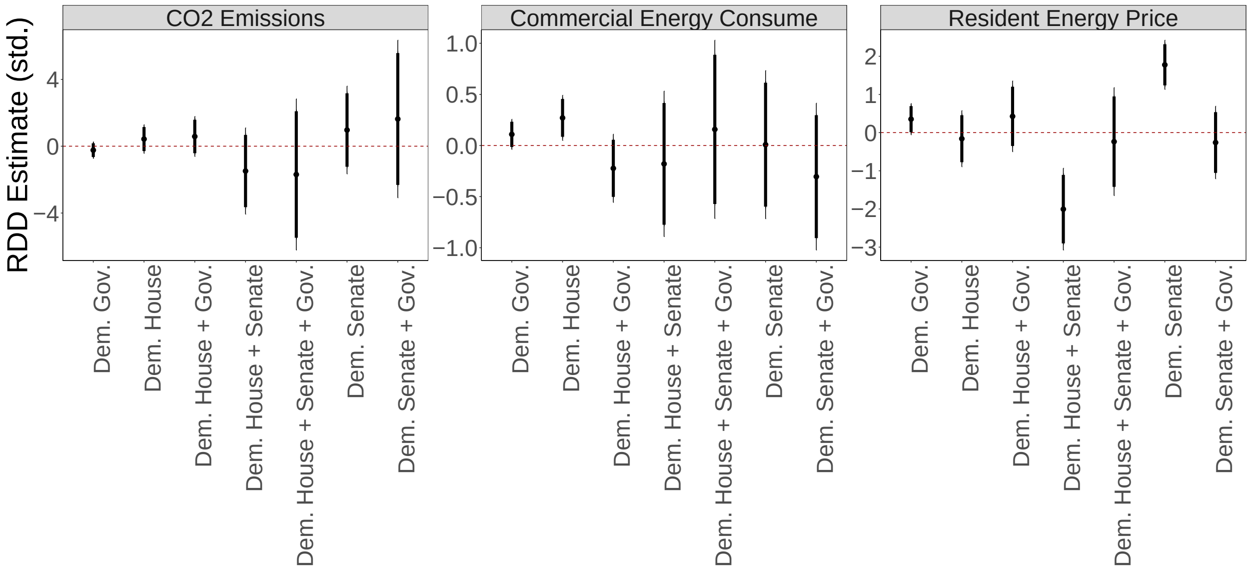


Figure A40 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A41: MRDD Estimate of Democratic Control on Civic Outcomes (Second Year)

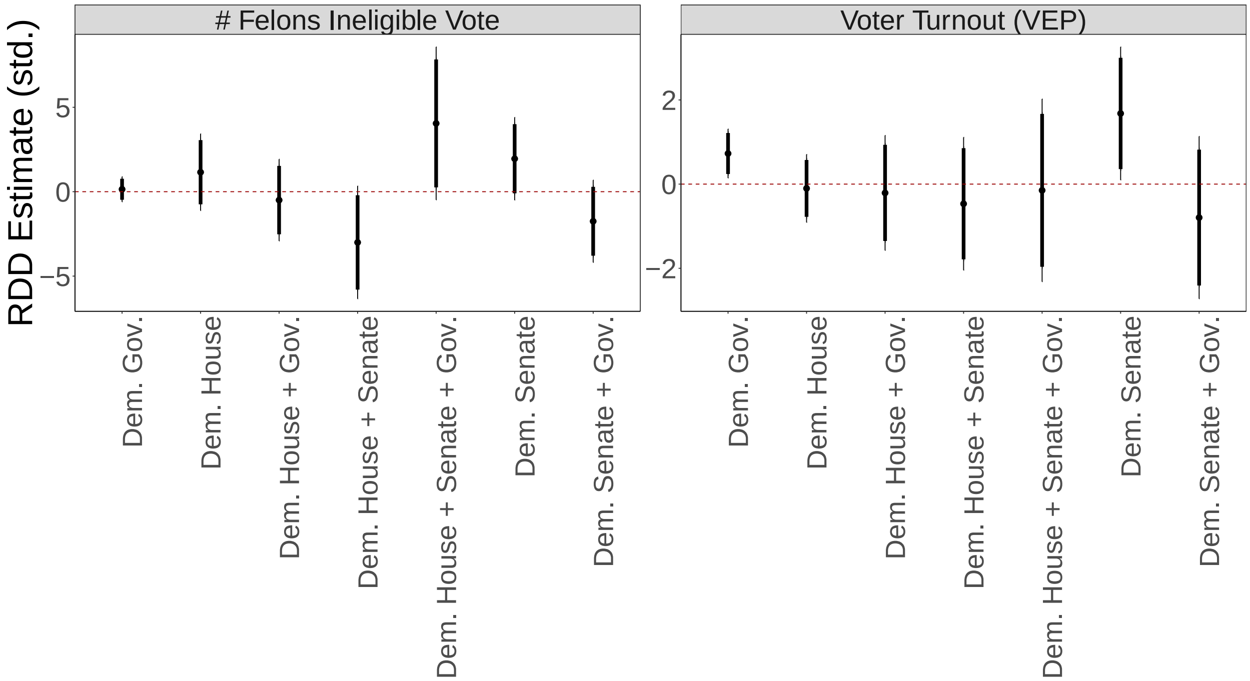


Figure A41 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A42: MRDD Estimate of Democratic Control on Crime Outcomes (Second Year)

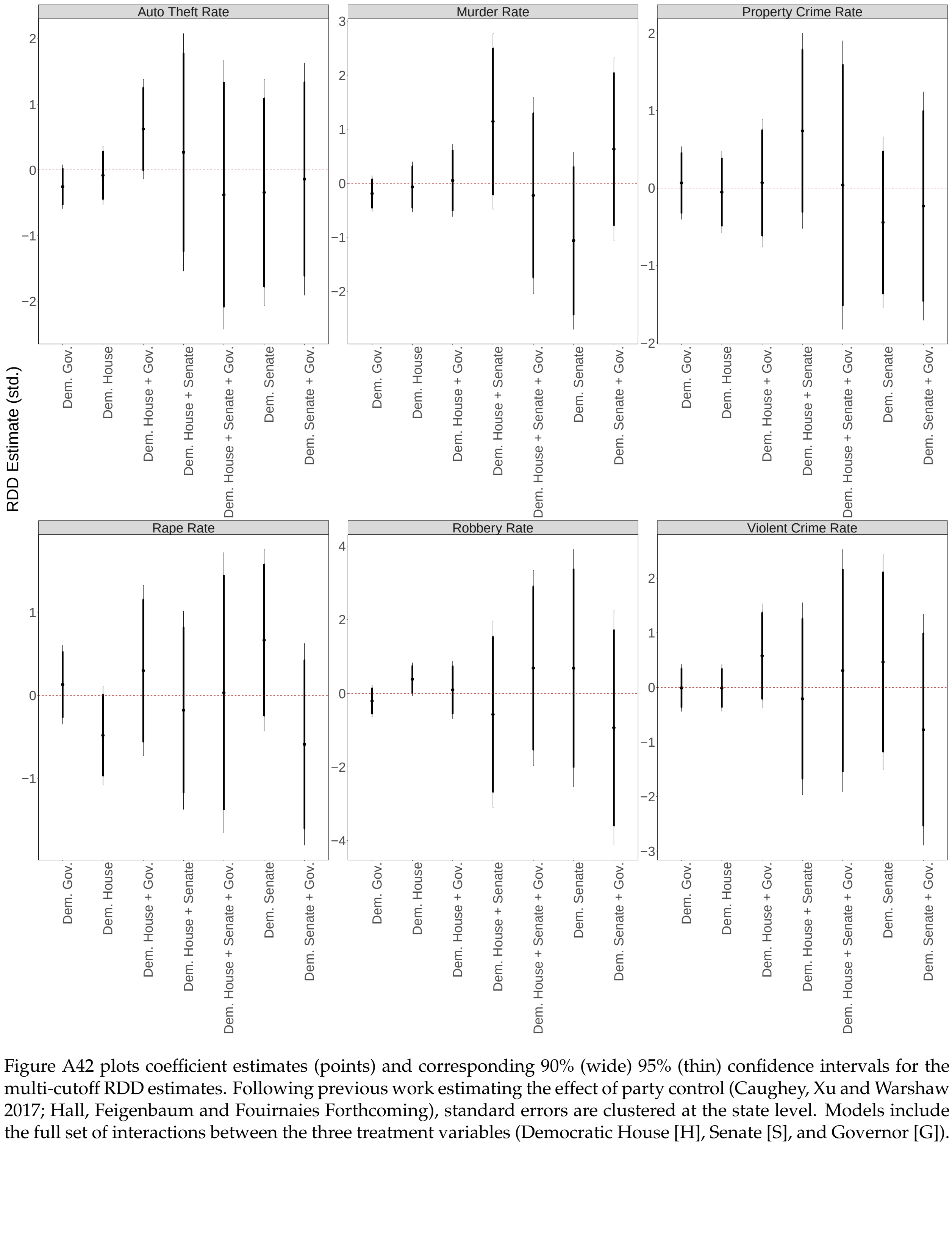


Figure A42 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

13 MRDD + Diff/Diff Results

13.1 First Year

Figure A43: MRDD + Diff/Diff Estimate for Economic Outcomes (First Year)

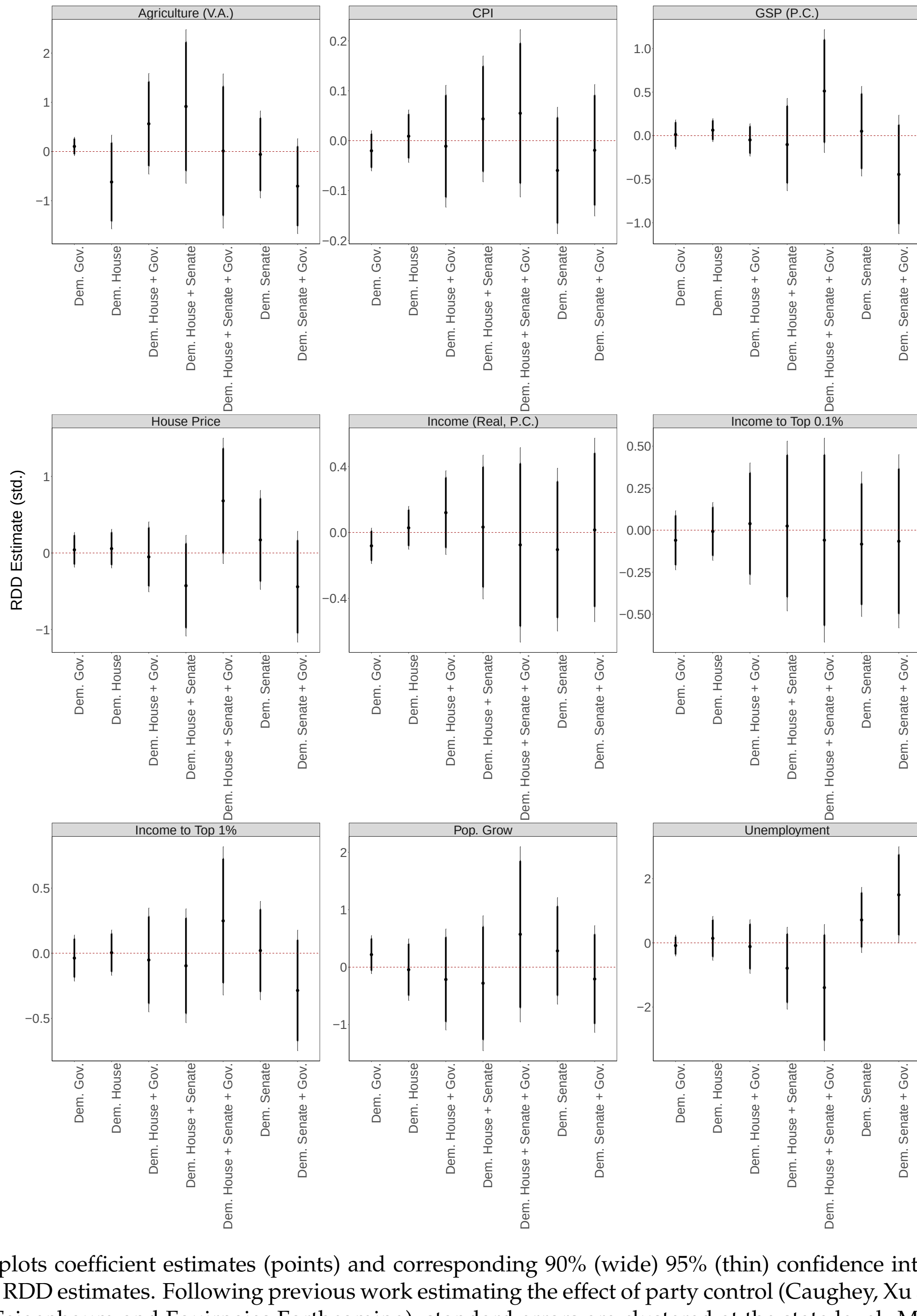


Figure A43 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A44: MRDD + Diff/Diff Estimate for Education Outcomes (First Year)

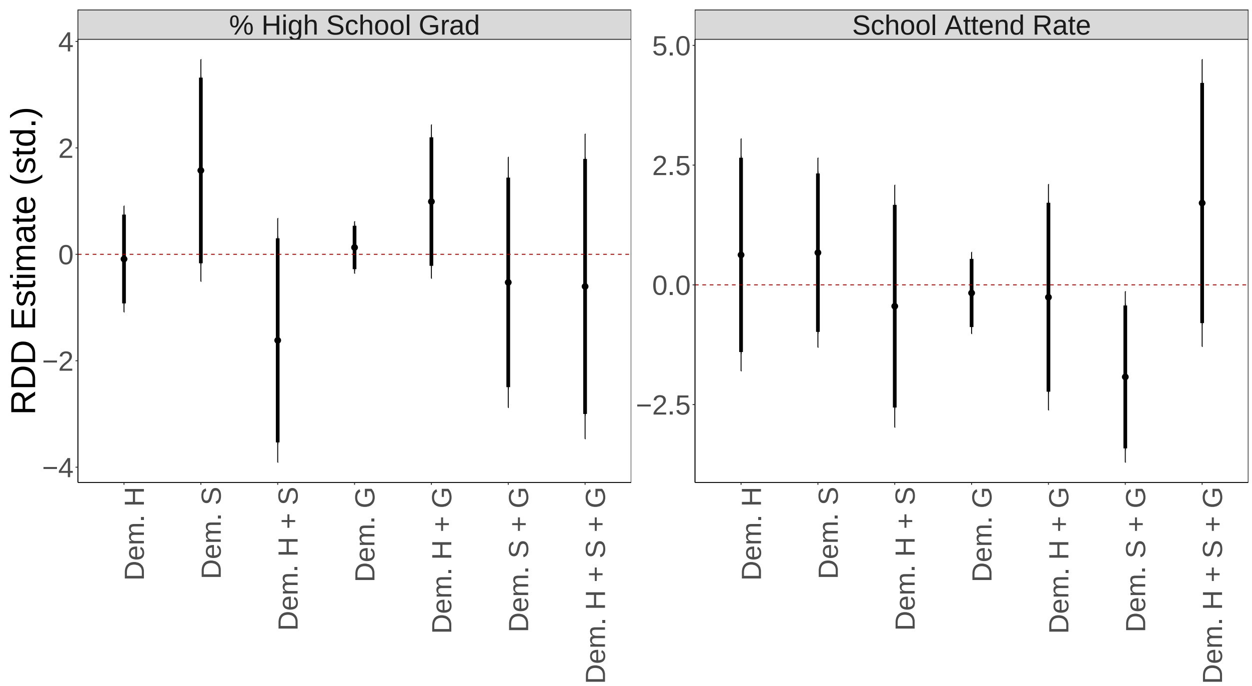


Figure A44 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A45: MRDD + Diff/Diff Estimate for Health/Family Outcomes (First Year)

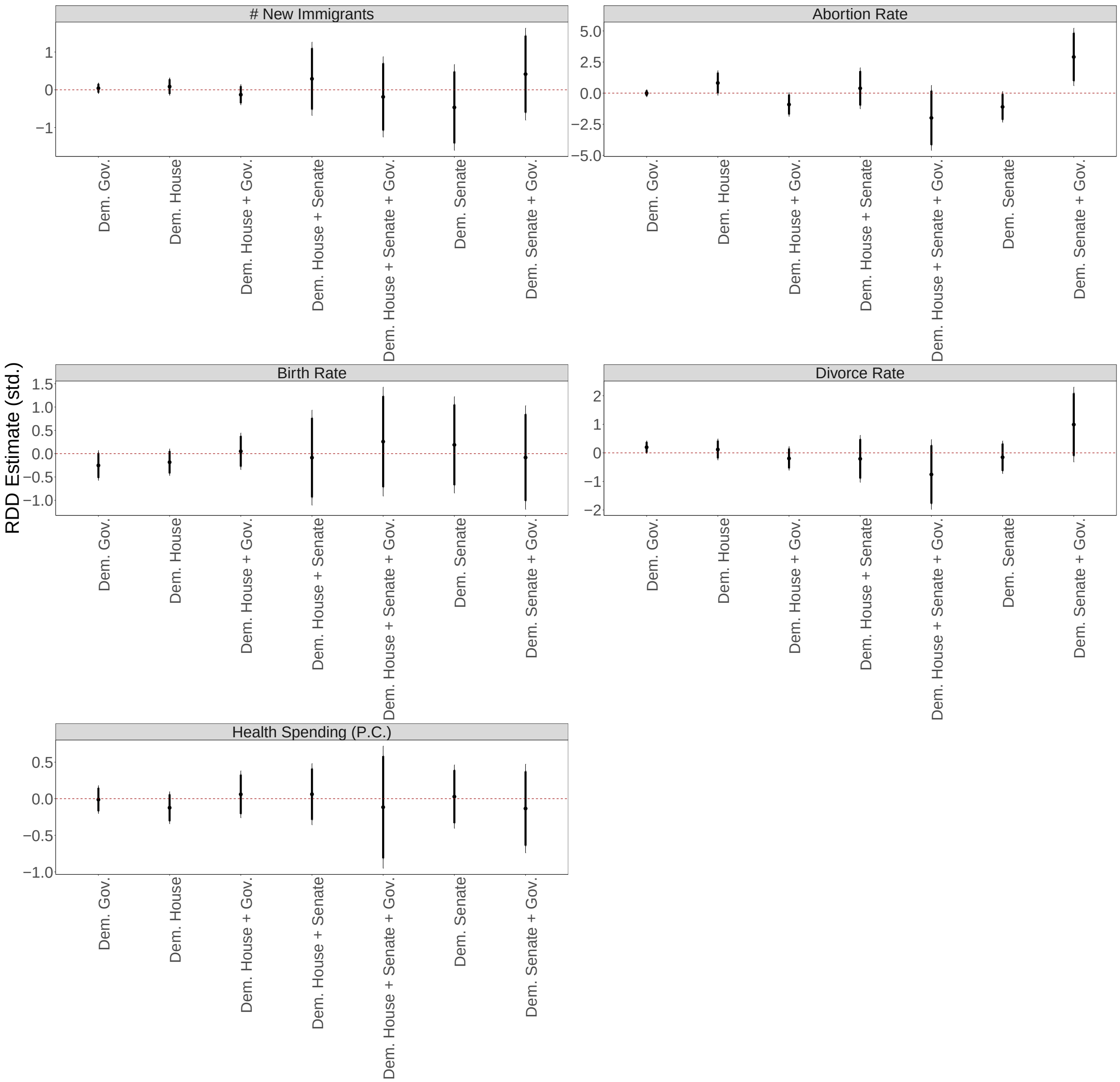


Figure A45 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A46: MRDD + Diff/Diff Estimate for Environmental Outcomes (First Year)

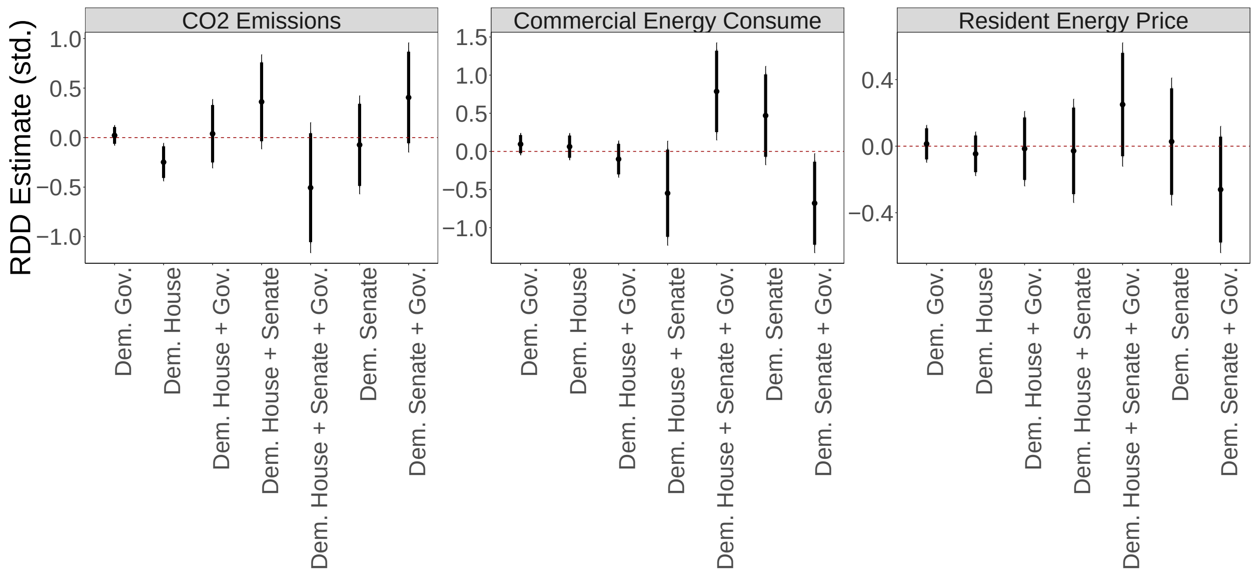


Figure A46 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A47: MRDD + Diff/Diff Estimate for Civic Outcomes (First Year)

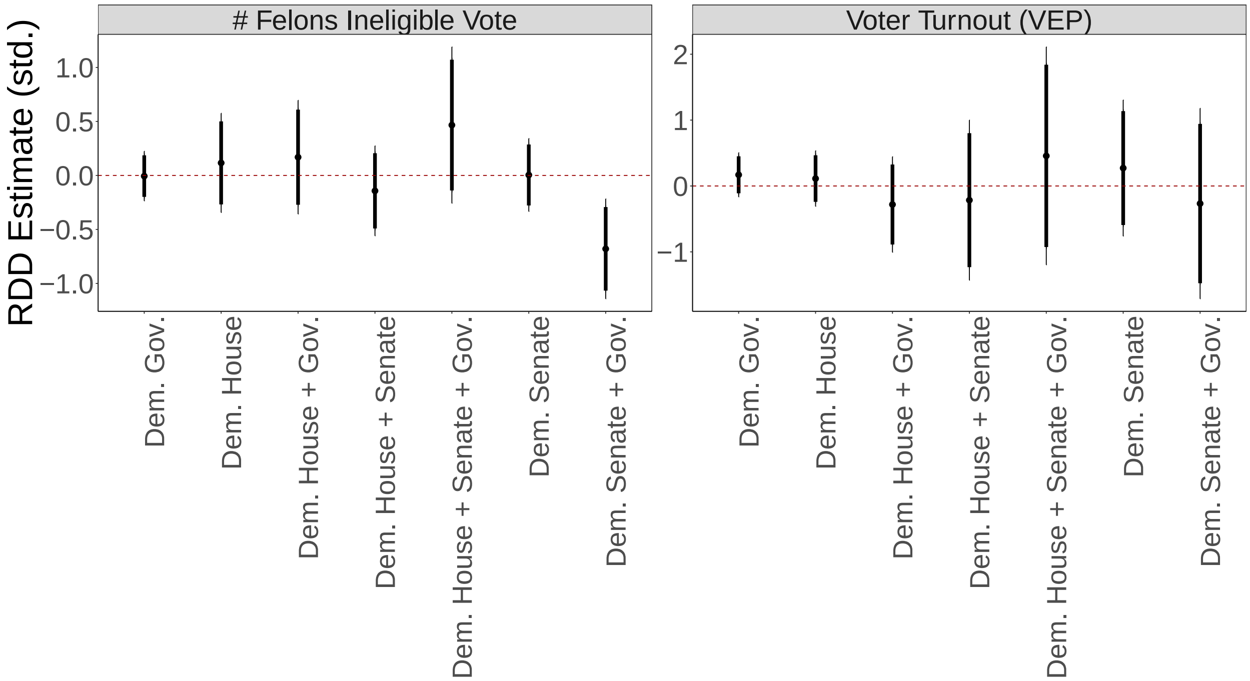


Figure A47 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A48: MRDD + Diff/Diff Estimate for Crime Outcomes (First Year)

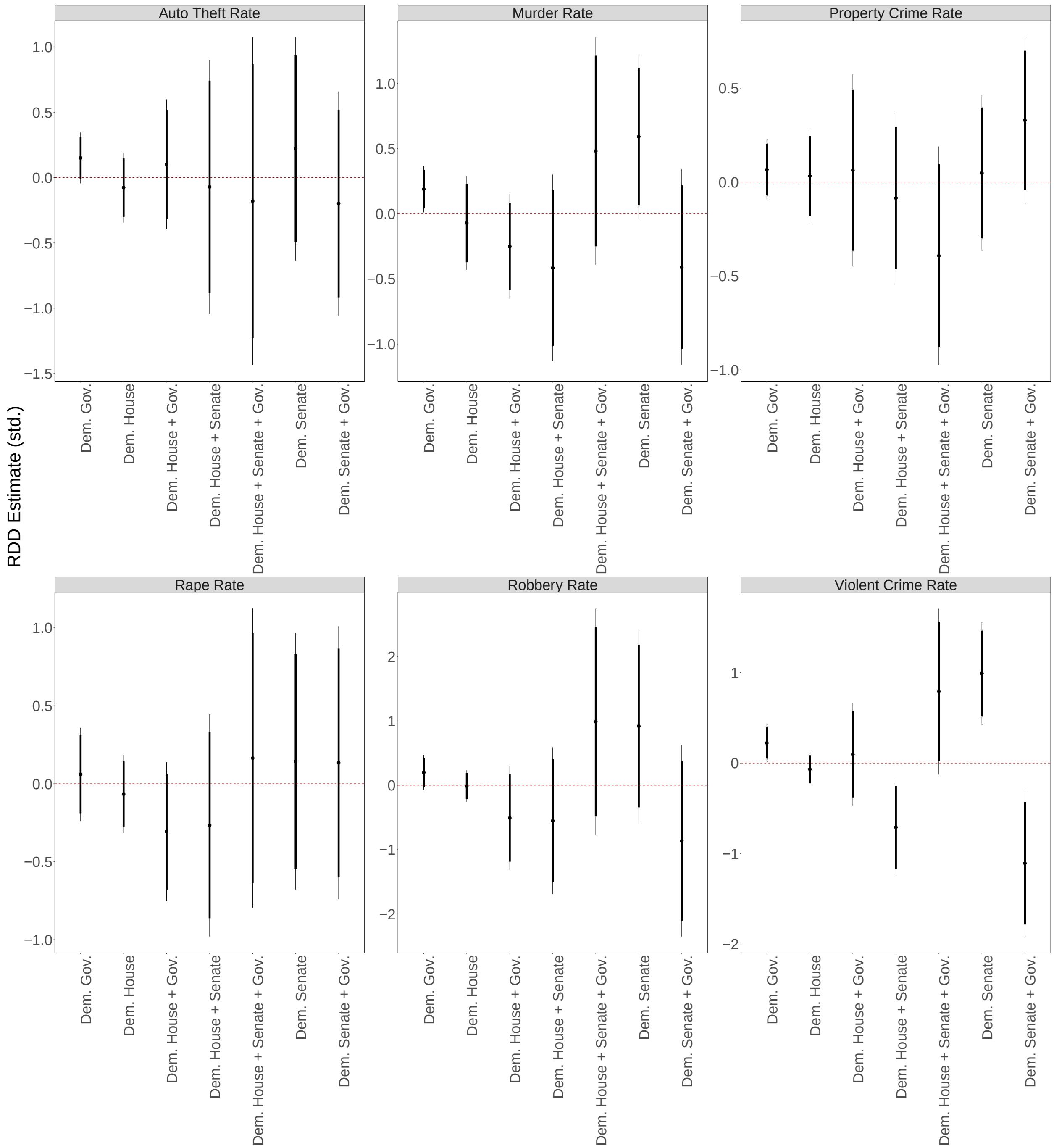


Figure A48 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

13.2 Second Year

Figure A49: MRDD + Diff/Diff Estimate for Economic Outcomes (Second Year)

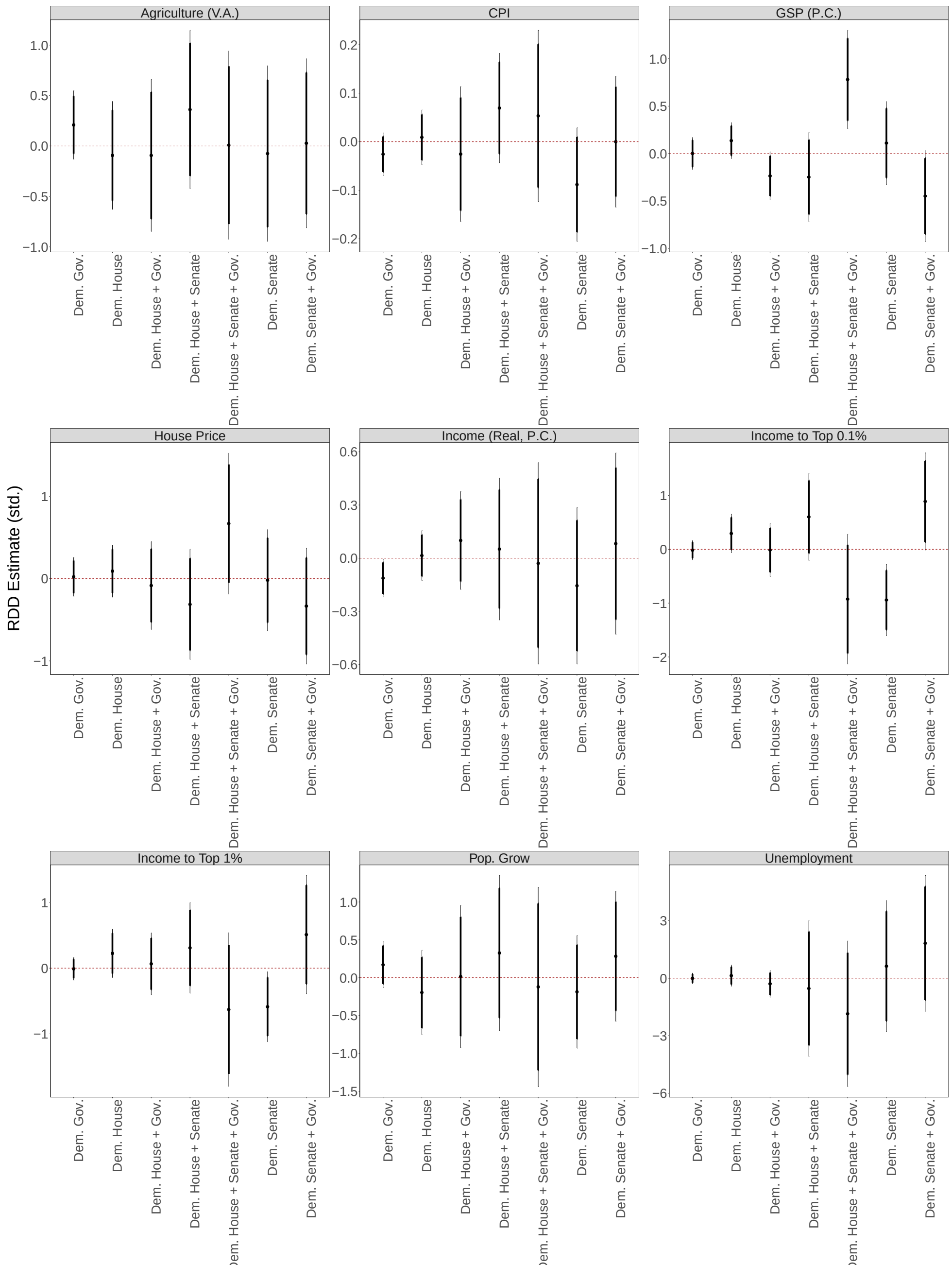


Figure A49 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A50: MRDD + Diff/Diff Estimate for Education Outcomes (Second Year)

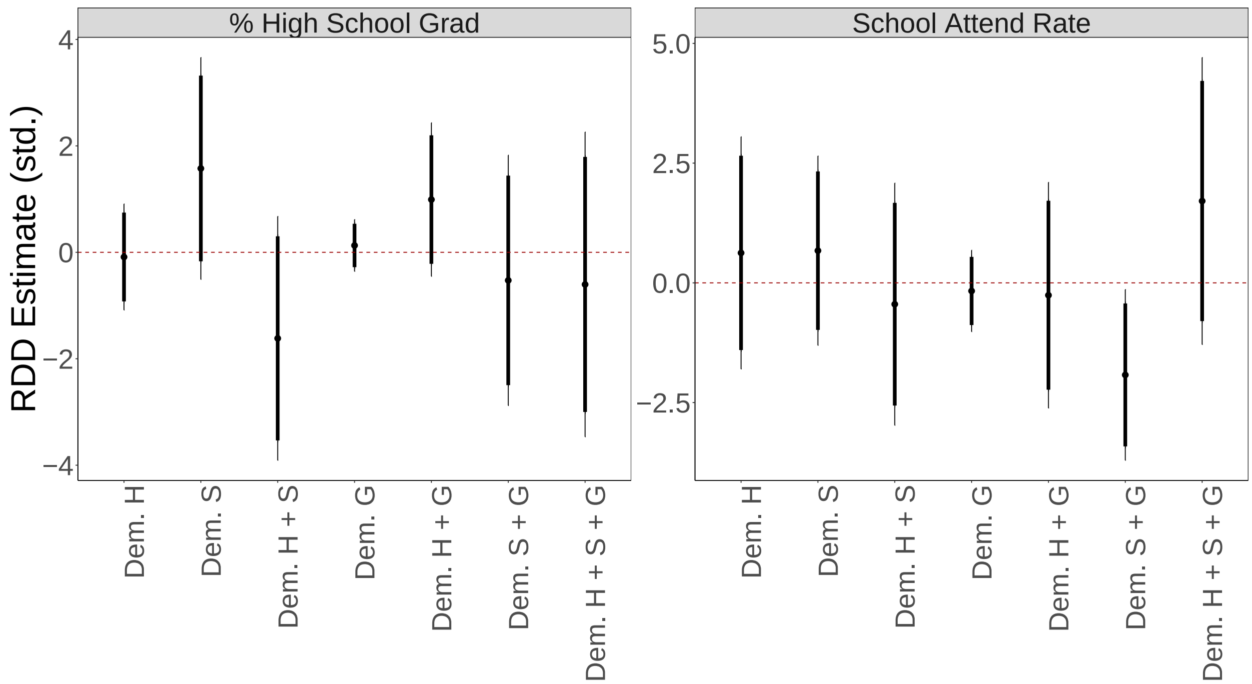


Figure A50 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A51: MRDD + Diff/Diff Estimate for Health/Family Outcomes (Second Year)

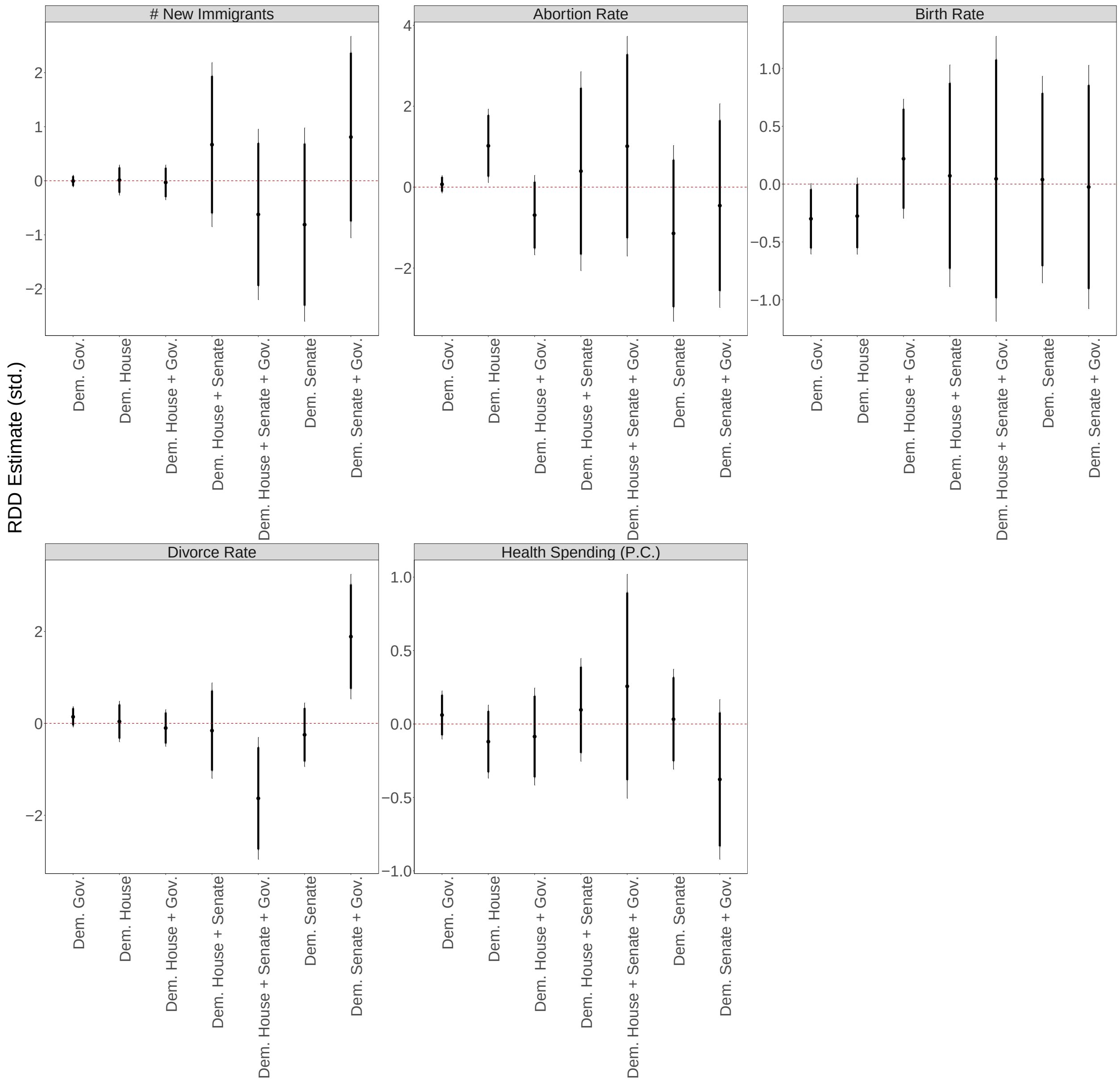


Figure A51 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A52: MRDD + Diff/Diff Estimate for Environmental Outcomes (Second Year)

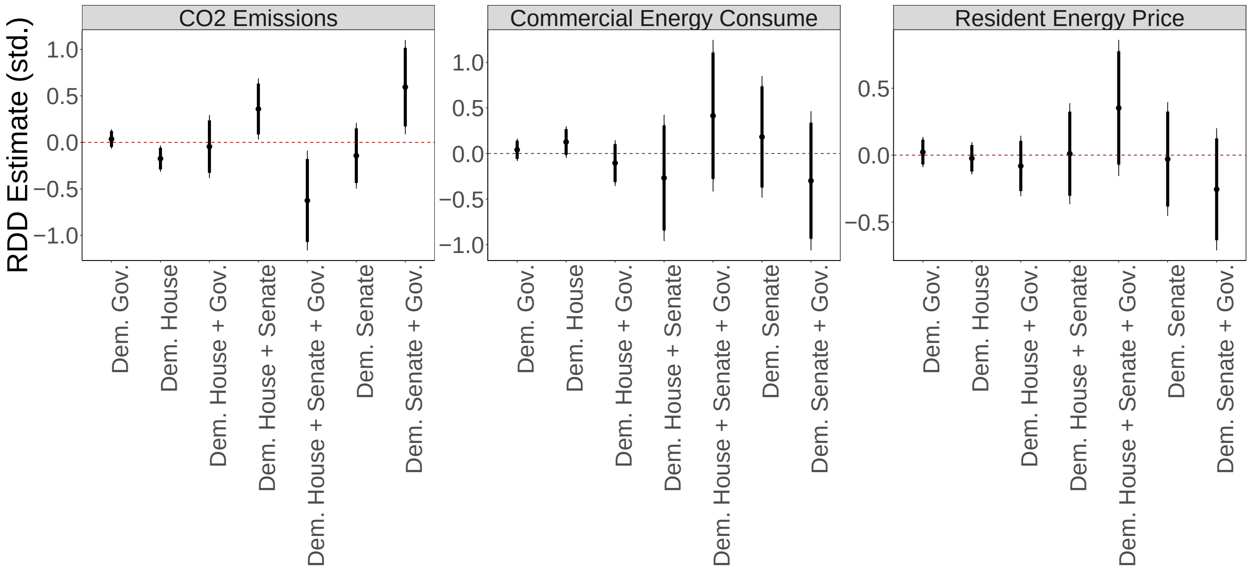


Figure A52 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A53: MRDD + Diff/Diff Estimate for Civic Outcomes (Second Year)

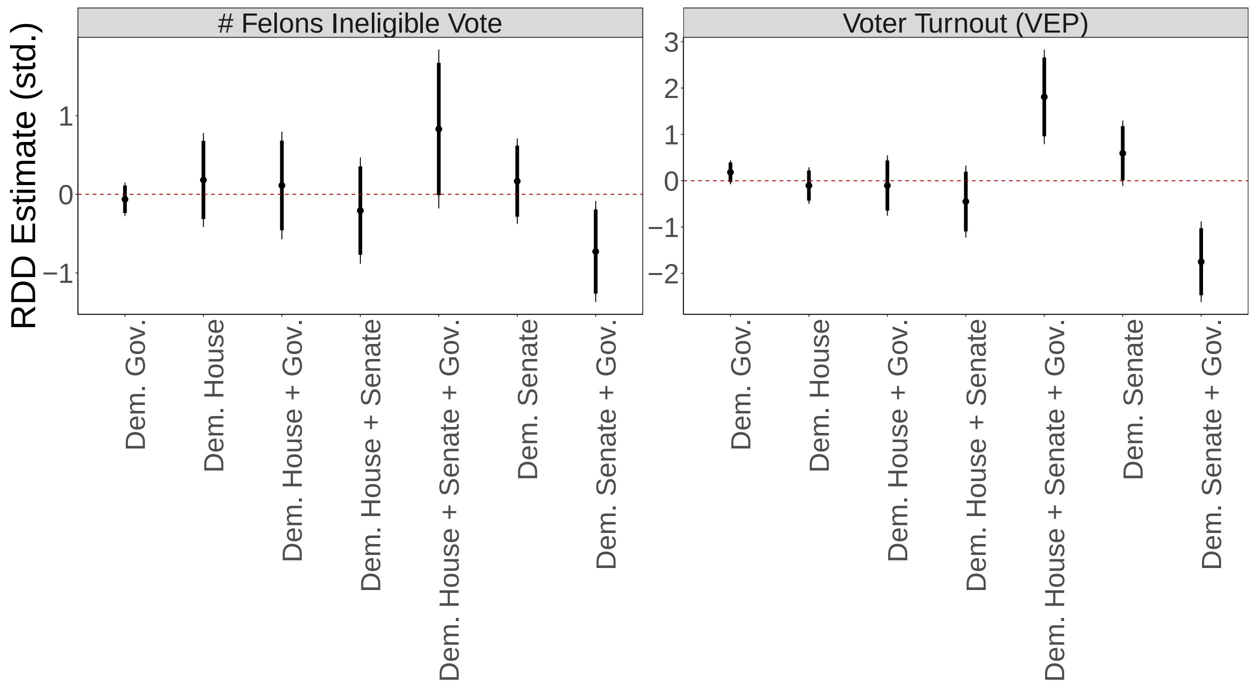


Figure A53 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

Figure A54: MRDD + Diff/Diff Estimate for Crime Outcomes (Second Year)

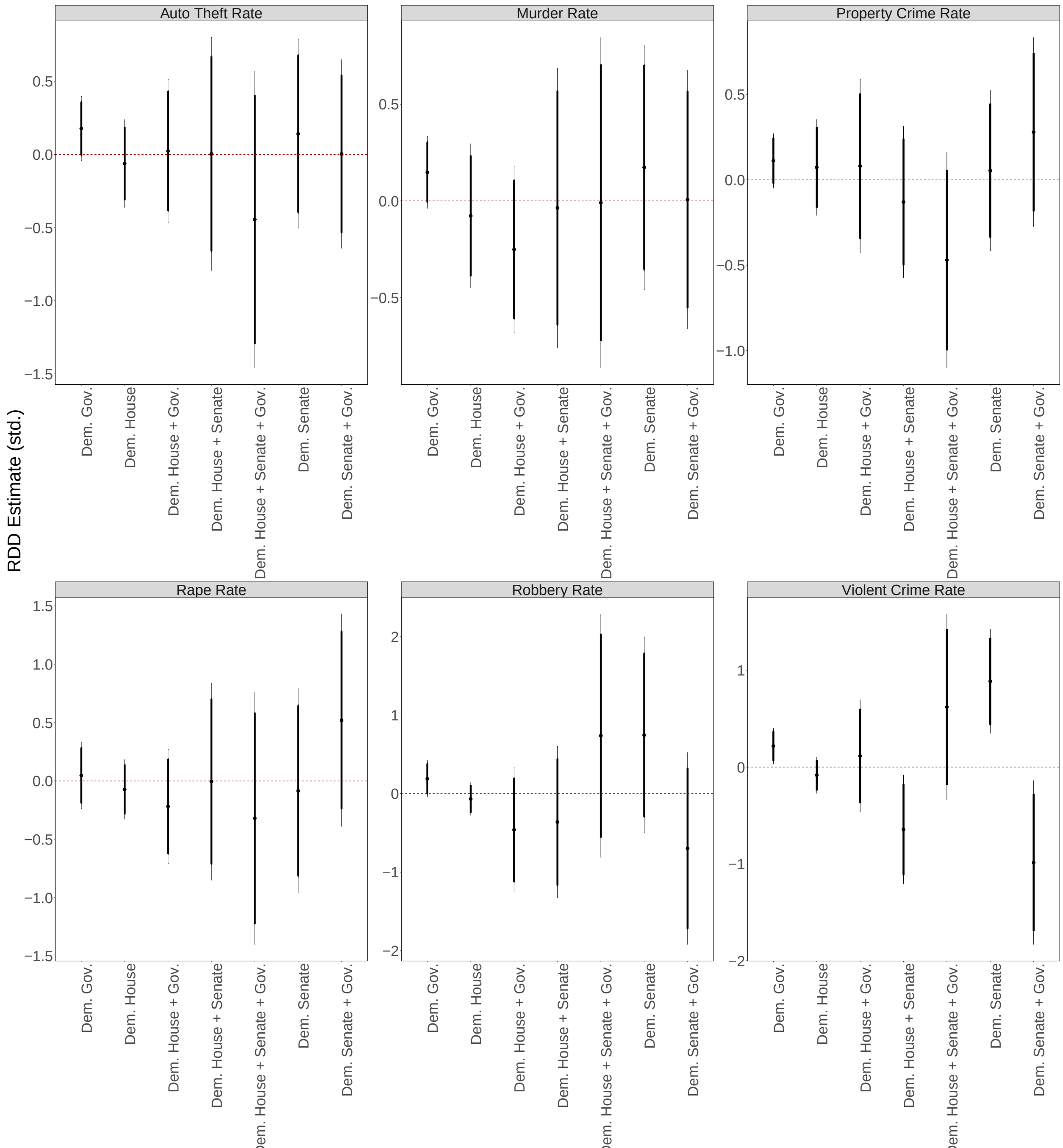


Figure A54 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the multi-cutoff RDD estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Models include the full set of interactions between the three treatment variables (Democratic House [H], Senate [S], and Governor [G]).

14 Alternate Ways to Define the Running Variable

Hall, Feigenbaum and Fouirnaies (Forthcoming) propose three alternate ways of creating the running variable for legislative party control, all of which rely on the closeness of individual state legislative races. The first—what they call the "Euclidean Distance" approach—measures the "distance between the vector of running variables and the treatment boundary" Hall, Feigenbaum and Fouirnaies (Forthcoming, 13). While having a nice geometrical procedure, Hall, Feigenbaum and Fouirnaies (Forthcoming) note that this specification is less interpretable. The second approach—what they call the "Manhattan distance" method—measures the cumulative total of "how many additional percentage points the party would have to be given to flip majority status" Hall, Feigenbaum and Fouirnaies (Forthcoming, 13). For example, if a party needed to win three seats in order to secure a majority, the "Manhattan Distance" would be the sum of the three closest seats distance below their individual race cutoffs.[37] The third approach—what they call the "Uniform Swing" method—uses the individual race score for the candidate that would push the legislature over the cutoff. That is, if a state were to be three seats away from the majority, the "Uniform Swing" method would use the third lowest race below the cutoff Hall, Feigenbaum and Fouirnaies (Forthcoming). The rationale here is that you are only as close to achieving control as your lowest race. Hall, Feigenbaum and Fouirnaies (Forthcoming, 13) note that this measure "assumes perfect correlation across elections."[38]

Figure A55 shows the McCrary density check across these three variables. As can be seen, there is a slight imbalance in the Manhattan distance, but balance across the other two. This combined with the covariate balance reported by Hall, Feigenbaum and Fouirnaies (Forthcoming) suggests that this is a valid way for specifying proximity to treatment.

[37] Conversely, if a legislature were in the majority by three seats, they would only be as close to falling into the minority as their three seats above the cutoff

[38] Specifying the cutoff in these ways preserves the balance that we show in Table A1. Consistent with work by Hall, Feigenbaum and Fouirnaies (Forthcoming) and Caughey, Xu and Warshaw (2017), there is, perhaps, even more balance with these alternate running variable scores.

Figure A55: McCrary Density Check for Precise Sorting (Alternate RVs)

(a) Euclidean (b) Uniform (c) Manhattan

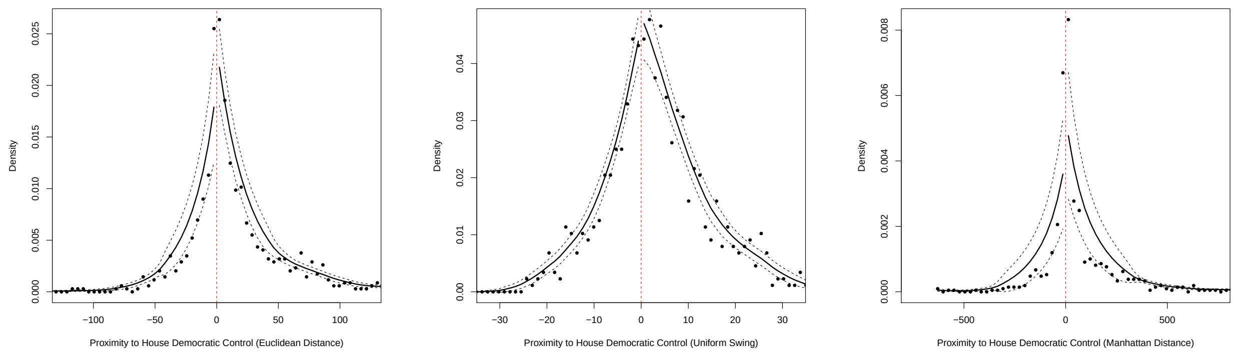


Figure A55 displays the McCrary Density Test for precise sorting (McCrary 2008). The x-axis displays the running variable for these three individual cutoffs. Corresponding p-values for H_0 = continuity at the cutoff: Euclidean = 0.11, Uniform = 0.82, Manhattan = 0.003.

Following the lead of Hall, Feigenbaum and Fouirnaies (Forthcoming), we estimate these models for the lower chamber as this is a cleaner comparison given non-overlapping election windows. However, the results do not change if we do our own calculation of their running variable scores for the upper chamber (available upon request). Figure A56 shows our RDD results using these alternate specifications of the running variable. The results are very consistent with those that we have outlined in the paper. There is little evidence of systematic effects on policy outcomes. And when there is divergence, the Hall, Feigenbaum and Fouirnaies (Forthcoming) running variables show estimates that are close to zero, with less evidence of significant effects. These results suggest that our conclusions are not an artifact of the construction of the running variable.

Figure A56: RDD Effect of Democratic Control: Alternate Running Variables (House; All Variables)

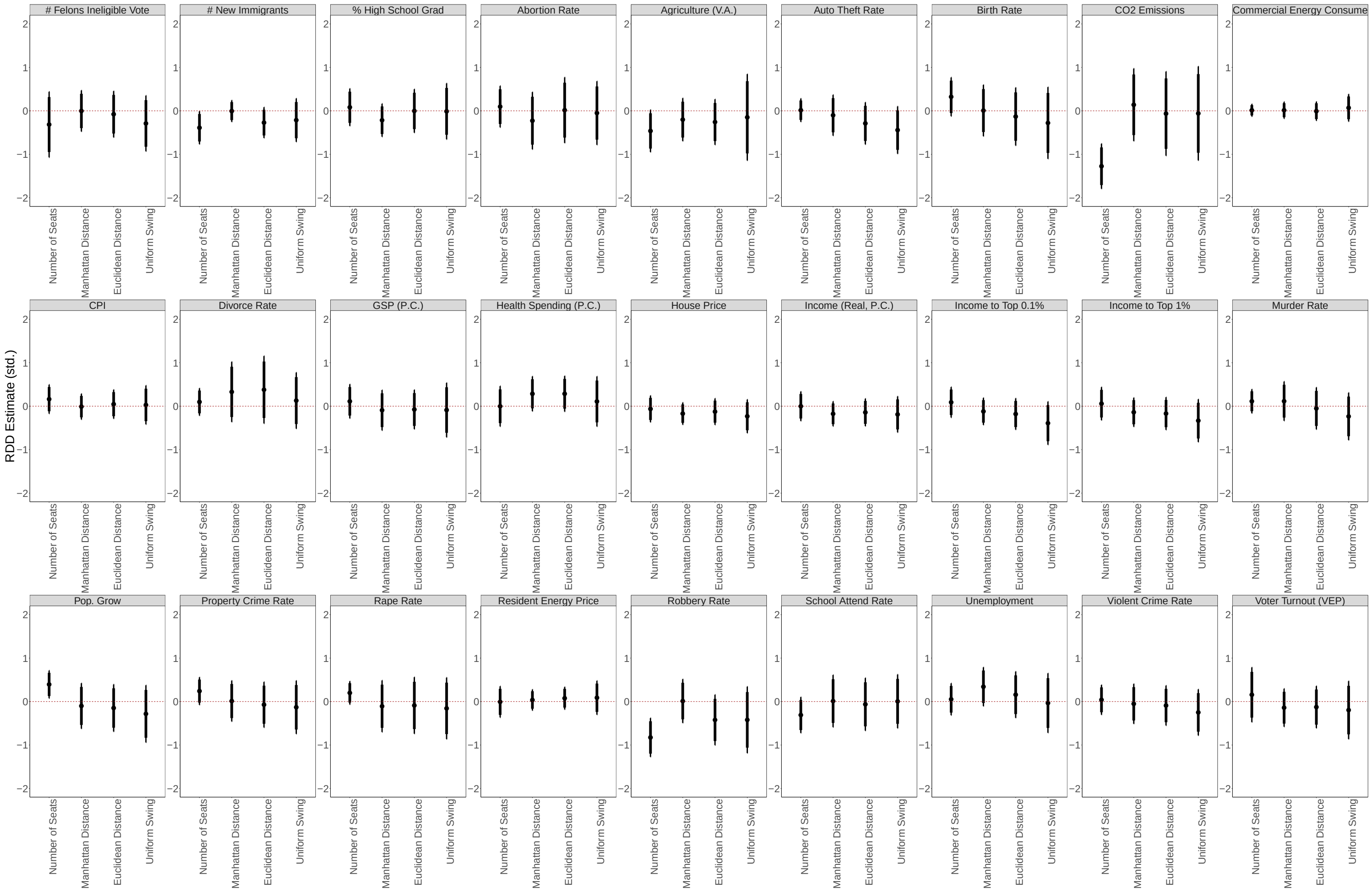


Figure A56 plots coefficient estimates (points) and corresponding 90% (wide) 95% (thin) confidence intervals for the regression discontinuity estimates. Following previous work estimating the effect of party control (Caughey, Xu and Warshaw 2017; Hall, Feigenbaum and Fouirnaies Forthcoming), standard errors are clustered at the state level. Following Hall, Feigenbaum and Fouirnaies (Forthcoming) we focus our attention on the lower chamber. The estimates use the optimal bandwidth as specified by the `rdrobust` command in `STATA` created by Calonico, Cattaneo and Titiunik (2014).