How Do Voters Matter?
Evidence from US Congressional Redistricting

Daniel B. Jones
University of South Carolina[a]

Randall Walsh
University of Pittsburgh[b]

This version: August 2016

Abstract. How does the partisan composition of an electorate impact the policies adopted by an elected representative? We take advantage of variation in the partisan composition of Congressional districts stemming from Census-initiated redistricting in the 1990's, 2000's, and 2010's. Using this variation, we examine how an increase in Democrat share within a district impacts the district representative's roll call voting. We find that an increase in Democrat share within a district causes more leftist roll call voting. This occurs because a Democrat is more likely to hold the seat, but also because – in contrast to existing empirical work – partisan composition has a direct effect on the roll call voting of individual representatives. This is true of both Democrats and Republicans. It is also true regardless of the nature of the redistricting (e.g., whether the redistricting was generated by a partisan or non-partisan process).

[a] University of South Carolina, Darla Moore School of Business, Department of Economics, 1014 Greene St., Columbia, SC 29208.
[b] University of Pittsburgh, Department of Economics, 230 S. Bouquet St., Pittsburgh, PA 15213.

1. Introduction

What is the relationship between voters' preferences and the policies supported and enacted by their representatives? Broadly speaking, voters influence policy on both an extensive margin and an intensive margin. On the extensive margin: voters choose between candidates through elections. If different candidates are expected to support different policies once elected, voters are essentially choosing which policy bundle they prefer when they vote for a given candidate. On the intensive margin, shifts in voter preferences may directly lead an already-elected representative to support different policies.

Theoretical models of electoral competition differ on which of these margins matter. The Downsian model of electoral competition (Downs, 1957) and related models suggest that, in order to achieve and maintain electoral support, politicians adopt policies that please the median voter. Thus, shifts in preferences of voters may lead to shifts in policymaking by their representatives – an intensive margin response. These models also, therefore, imply that who is elected (the extensive margin) is of less consequence: all candidates propose policies close to the median voter's ideal. Other models (e.g., "citizen-candidate" models[1]) assume that politicians adopt their personally preferred policies if elected, so elections only serve to select the candidate whose policy proposals are most preferred by voters. That is, under these models, only the extensive margin is operative.

In our paper, we take advantage of variation in the partisan composition of Congressional districts stemming from post-Census redistricting in the 1990's, 2000's, and 2010's in order to empirically assess the importance of these two margins. To do so, we construct a new measure of

[1] For Citizen-candidate models, see: Osborne & Slivinski (1996), Besley & Coate (1997). The idea that politicians simply enact their personally preferred policy is also consistent with Alesina's (1988) model with limited concerns about future election outcomes.

"predicted Democrat share" within each district, which allows us to observe the share of Democrats within every Congressional district just before and after each wave of redistricting.

Using a difference-in-differences strategy (with continuous treatment), we ask: "Does a larger share of Democrats within a district lead to more leftist representation in Congress?" If so, does this happen only because a Democrat is more likely to be elected (an extensive margin response)? Or, does a leftward shift in representation occur even if the incumbent party or candidate remains in office both before and after the shift in the electorate (an intensive margin response)? The difference-in-differences strategy allows us to answer these questions while stripping away the influence of (1) general time trends in ideological positions in Congress and (2) unobservable differences between Congressional districts and representatives (that are constant across redistricting).

The extant empirical literature on this question has led to mixed results. Several researchers have documented a relationship between voters' preferences and the ideological position of their elected representatives in the legislature (e.g., Levitt, 1996; Gerber and Lewis, 2004), but disentangling whether such a relationship occurs through the intensive or extensive margin is not the main goal of those papers. Most relevant is the work of Lee, Moretti, and Butler (2004) who also study the US House of Representatives. They use a regression discontinuity strategy to isolate quasi-random variation in the "electoral strength" of a party, building on the notion that a narrow (quasi-randomly assigned) Democrat victory in the previous contest generates strength for the Democratic candidate in the next election due to the incumbency advantage. Ultimately, they find that increased electoral strength only impacts the roll call voting behavior of a district's representative through the extensive margin, with no intensive margin response. Indeed, they conclude: "Voters merely *elect* policies," and that once a candidate has been elected: "the degree

of electoral strength has no effect on a legislator's voting behavior." In work concurrent to ours, Fedaseyeu et al. (2015) document that voters become significantly more likely to vote for Republicans in areas where hydraulic fracturing (or "fracking") has driven fossil fuel extraction booms. This move toward Republican representation then in turn leads to more conservative representation in the House. As in Lee, Moretti, and Butler (2004), they find that this result comes entirely through the extensive margin. Conditional on being elected, representatives from areas with shale booms, on average, vote no differently than do other members of their party.

However, other research using the same empirical approach as Lee, Moretti, and Butler (2004) has found conflicting evidence in the context of the US Senate (Albouy, 2011). Moreover, recent research suggests that the assumptions necessary for a valid regression discontinuity design are not satisfied in US Congressional elections (Caughey and Sekhon, 2011).[2] Given these mixed results and the concerns raised around using a regression discontinuity strategy in the US House, as do Fedaseyeu et al. (2015), we aim to contribute to this literature by providing new evidence from a different empirical approach.

We find clear evidence that both margins matter. First, not surprisingly, an increase in Democrats within a district leads to more leftist representation overall. Part of this result stems from the extensive margin: a positive shock to the number of Democrats in the district increases the likelihood that a Democrat is elected, and Democrats are more likely to hold a leftist ideological position in their roll call voting. However, this simple extensive margin effect does not entirely explain the shift to the left. We find that an increase in the number of Democrats within a district leads to more leftist representation even when controlling for party affiliation. Indeed, only about 63% of the overall shift to the left in response to a higher share of Democrats appears to be driven

[2] Caughey and Sekhon (2011) find that narrowly elected Democrats are different in a number of ways (other than just the fact that they won): incumbency status, financial resources, political experience, and other observables.

by increased likelihood of electing a Democrat. This is in contrast to Lee, Moretti, and Butler's (2004) result; in their paper, a change in Democrats' electoral strength within a district led to a shift to the left in roll call voting, but roughly 100% of this change was explained by increased likelihood of electing a Democrat.

The main threat to identification in our analysis is the fact that Congressional districts are not randomly drawn and therefore our treatment is not randomly assigned. As in any difference-in-differences approach, this fact only threatens the validity of our research design if the factors that determine treatment are also related to the anticipated trend of the outcome variable. We would therefore be concerned if districts experiencing the largest changes in partisan composition were markedly different in their pre-existing partisan composition or if the pattern of redistricting varied substantially by the circumstances surrounding redistricting (e.g., party of incumbent, cause of redistricting, partisanship of redistricting authority). We address these concerns in a number of ways.

We begin by directly assessing the relationship between pre-existing Democrat share and redistricting-prompted changes in Democrat share – both in isolation and in relation to various types of redistricting processes. In the aggregate, we find no meaningful difference in the post redistricting change in Democratic share between districts with a low baseline democrat share and those with a high baseline democrat share.[3] Although this finding may seem surprising given frequent discussion of heavy manipulation of redistricting for political purposes, recent research in fact suggests that the conventional wisdom on redistricting is not borne out in data. McCarty et al. (2009) provide evidence to suggest that there is in fact very little relationship between redistricting and an increase in polarization in Congress, which would be expected if redistricting

[3] A one percentage point increase in baseline Democrat share is, on average, associated with a 0.034 percentage point decrease in Democrat share following redistricting.

was used by state governments to minimize the competitiveness of districts.[4] Further, and perhaps more importantly for our analysis, the relationship between baseline Democrat share and the redistricting outcome does not appear to vary with the nature of the redistricting. To explore this, we split our sample along several dimensions that proxy for the likelihood that redistricting was associated with political motivations (e.g., non-partisan vs. partisan processes); we find no evidence that this relationship varies based on the likelihood that states were engaged in politically motivated redistricting.

While the descriptive evidence suggests that selective redistricting may not pose a threat to our analysis, in our empirical work we demonstrate that our results are robust to three different strategies for addressing the issue. First, we include a rich set of time trends (interacted at both the district and congress person level). Second, we replicate our analysis on different subsamples of the data, focusing on states whose redistricting processes were less likely to have been politically motivated. Finally, we evaluate the impact of district composition on a second demographic dimension, percent black. Utilizing the Leadership Conference on Civil Right's (LCCR's) Congressional ratings as our outcome variable, we demonstrate the presence of an intensive margin effect of percent black on voting behavior relative to the LCCR's agenda. This result is robust to controlling for Democrat vote share.

Our finding of both an intensive margin and an extensive margin effect contributes to a literature in political science exploring the impacts of redistricting on legislators' behavior (Boatright, 2004; Crespin, 2010; Bertelli & Carson, 2011). These authors focus on representatives present before and after a single wave of redistricting and study their response to a change in

[4] Friedman and Holden (2009) challenge the notion that redistricting is aimed to provide incumbents an advantage; they in fact provide causal evidence that the incumbent reelection rate is lower after each wave of post-Census redistricting during the time period we study, perhaps due to a tightening in the legal constraints (and enforcement of constraints) on redistricting in recent decades.

partisan composition (as measured by presidential vote share in the most recent presidential election). Ultimately, results from those papers are mixed: Boatright (2004) and Crespin (2010) provide some evidence that representatives do change their voting behavior after redistricting; Bertelli & Carson (2011) do not. While related, the focus of our paper is different. We are especially motivated to quantify the relative impact of electorate shifts on the extensive and intensive margins. In doing so, we provide a new measure of partisan composition that is not tied to any particular candidate. We measure the effects of several waves of redistricting, and demonstrate the robustness of our analysis to potential non-randomness of redistricting.

Our results also speak to a more general literature testing implications of models of electoral competition. One prediction of the Downsian model has received substantial empirical scrutiny: policy convergence. Under the Downsian model, rival candidates both aim to please the median voter and therefore offer (and, if elected, enact) identical policies; thus it ultimately does not matter who is elected. Competing models predict policy *divergence* – that is, that different parties adopt different policies if elected.[5] It is worth noting that the "extensive margin" effect we discuss is only relevant if there is some degree of policy divergence. A number of papers have empirically tested whether there is more evidence of policy convergence or divergence, often using a regression discontinuity design to randomly assign which party wins the election. Results are very mixed; some papers provide clear evidence of policy divergence, while others document convergence.[6]

[5] For instance, if candidates are motivated to run to enact their personally preferred policy (Osborne and Slivinski, 2006; Besley and Coate, 2007), policy platforms are not credible commitments (Alesina, 1988), or voters rationally abstain from voting if they are indifferent between candidates (Llavador, 2006), then we may expect divergence: candidates from different parties adopt different policies if elected.

[6] Lee, Moretti, & Butler (2004), Albouy (2013), Beland (2015), and Hill & Jones (2016) all find clear evidence of partisan differences in enacted policy and therefore policy *di*vergence. Other studies find little or no partisan difference in policy (e.g., Reed, 2006; Leigh, 2008; Ferreira and Gyourko, 2009; Gerber and Hopkins, 2011).

Despite the substantial amount of empirical work on policy convergence/divergence, convergence is of course just one prediction of the Downsian model. A more general prediction of the model (and various extensions of the model) is that candidates move *towards* the preferences of the median voter, or at least shift their policies in reaction to a shift in the preferences of the median voter. This can occur even if policies of competing candidates do not fully converge. That is, we may observe "partial policy convergence" (Alesina, 1988). Testing whether candidates move towards the median voter in their district in general (even if we do not observe full policy convergence) is the main goal of our paper; indeed, a model which predicts some response – but not complete convergence – to the median voter best describes our results. Put differently, while our results do not support all of the predictions of the Median Voter Theorem, they do support the importance of the median voter's preferences.

2. Empirical approach

How does the partisan composition of her electorate impact an elected official's policy decisions? To answer this question, we employ a continuous-treatment difference-in-differences approach. The "treatment" is variation in partisan composition of a Congressional district. To measure this variation, we construct a novel measure of predicted partisan composition, which we describe in the next section. Because the context we focus on is the US House of Representatives, the policy decisions of interest are roll call votes. As we discuss in more detail in the next section, our main outcome variable is the first dimension of Poole & Rosenthal's "DW-Nominate" score.[7] This measure collapses all of a representative's votes from a particular session of Congress into a single measure capturing their ideological position along a continuum; increasingly negative

[7] There is a second dimension of the DW-Nominate score which accounts for regional differences within parties, which we do not incorporate into our analysis.

numbers indicate increasingly leftist voting, while increasingly positive represent increasingly conservative voting. Thus, our analysis makes use of a panel of representatives, with one observation per Congress.

Our source of variation in partisan composition (measured as share of Democrats in a district) stems from redistricting in the early 1990's, 2000's, and 2010's. For each wave of redistricting, the "pre" period consists of the two (two-year) Congresses before redistricting; the "post" period consists of the two (two-year) Congresses after redistricting. In practice, this means that we treat our data (described in more detail in the next section) as three pooled panels, rather than a single long panel. For instance, Alabama's 1st Congressional District is coded as an entirely different district for each of the three waves of redistricting. (For that reason, although we discuss the use of "district fixed effects" in this section, in practice, our empirics use "District X Redistricting wave fixed effects". This allows a district that is named "Alabama - 1st District" to be treated differently each decade.) We elaborate on the reasons for this approach in the data section.

With this setup, we estimate variations on the following equation:

$$DW_{idt} = \alpha + \beta_1 DemShare_{dt} + \gamma_t + \delta_d + \epsilon_{idt},$$

where "DW_{idt}" is the first dimension of the DW-Nominate score for representative *i* in district *d* at time (or Congress) *t*. On the right hand side, "$DemShare_{dt}$" measures the partisan composition of district *d* at time *t* and is of primary interest. We exploit variation in the partisan composition of district *d* stemming from redistricting to identify the impact of this variable. We also include time (Congress) fixed effects and, at a minimum, congressional district fixed effects. (Extensions of the specification, discussed below, include individual representative fixed effects.)

Just as in a difference-in-differences approach with a binary treatment variable, identification is based on *changes* in the treatment variable before and after redistricting (in this case, "DemShare"). If partisan composition is identical before and after redistricting (i.e., the district was not redistricted), this will be captured by the district fixed effect.

Having established the basic specification, we now consider the hypotheses we test. Note that the simple specification described thus far identifies the combined effects of the intensive and extensive margins discussed in the introduction. That is, an increase in *DemShare* may be expected to impact ideological positioning (*DW*) in two ways:

(1) Intensive margin effect: *DemShare* may have a direct effect on *DW*. Under the Downsian model of electoral competition, politicians respond to a leftist shift in the preferences of their electorate by proposing and adopting more leftist policies.

(2) Extensive margin effect: An increase in *DemShare* increases the likelihood that a Democrat is elected. If, conditional on being elected, Democrats enact more leftist policies (or engage in more leftist roll call voting), then the increased likelihood of Democratic victory implies more leftist representation.

Given these two channels, we expect β_1 to be negative: an increase in *DemShare* should lead to more leftist roll-call voting (reflected by a more negative DW-Nominate score). The magnitude of the coefficient is of interest as it provides a baseline measure of the *overall* impact of a shift in partisan composition. We will compare this to the magnitude of coefficients in other specifications to disentangle the intensive and extensive margin effects.

We adopt several strategies to disentangle these different drivers. First, we can of course simply include a dummy indicating the partisan affiliation of district *d*'s representative in Congress *t*. That is, we estimate:

$$DW_{idt} = \alpha + \beta_1 DemShare_{dt} + \beta_2 DemRep_{dt} + \gamma_t + \delta_d + \epsilon_{idt},$$

where $DemRep_{dt}$ is an indicator that a Democrat is elected in district d at time t. In other specifications, we restrict our focus to districts where the partisan affiliation of a district's representation in Congress is the same before and after redistricting. In either case, the resulting estimate of β_1 captures the *direct* impact of DemShare on roll call voting separated from the indirect effect that comes through an increased likelihood of a Democrat being elected.

Of course, there is a second potential *extensive* margin impact that is not dealt with through either of the two approaches we just noted. In particular, just as an increase in the share of Democrats increases the likelihood of electing a Democrat, there may also be an increased likelihood that the candidate put forth by a given party is relatively leftist. That is, there may be a higher likelihood of electing a centrist Republican in an otherwise mostly Republican district, or a higher likelihood of electing a far-left Democrat in an otherwise mostly Democratic district. This would make it appear as though higher DemShare leads to more leftist voting even controlling for partisan affiliation (or restricting to districts where partisan affiliation of the representative remains constant across districts), and indeed it would, but not necessarily because politicians are *responding* to their electorate. Instead, this would represent a more nuanced extensive margin effect, wherein voters choose a more leftist politician (conditional on party) and this politician then proceeds to vote according to his or her own preferences.

In our most robust specifications, we include individual representative fixed effects. In those specifications, identification implicitly stems from treatment-induced changes in roll call voting of representatives who were present both before and after redistricting. Within-party ideological variation is captured by the representative fixed effect, so this approach removes the influence of even the nuanced within-party extensive margin effects from the estimate of β_1. The

resulting β_1 captures the intensive margin impact of *DemShare* on a given representative's roll call voting separated from any effect driven by electing a different type of candidate (different party or otherwise). This specification is also, in some sense, the most conservative approach, as individual candidate fixed effects absorb a large amount of the variation in ideological positioning.

What are the main threats to our empirical approach? The main concern is that the partisan composition of Congressional districts is not randomly determined. State governments are responsible for redrawing Congressional districts within their state; the potential for states to draw districts favorable a particular party or incumbents has been much discussed in academia and in popular media. Given our identification strategy, an important threat is that pre-existing trends in ideological positioning of representatives cause large changes in partisan composition, either to secure the seat for the incumbent whose political capital has grown or to remove a party's growing stronghold on a particular district. To deal with this, we: a) check for patterns consistent with this form of non-random redistricting behavior in the descriptive data; b) include specifications that control for either district- or representative-specific time trends; c) split our sample into states that one might expect would be more or less impacted by non-random redistricting (e.g., states forced to redistrict because of reapportionment vs. states that redistricted without gaining or losing seats); and, d) demonstrate that similar findings hold when we analyze the impact of share black on voting relative to the priorities of the LCCR's legislative priorities – results that are robust to additional controls for share Democrat.

3. Data

3.1 Measure of Congressional district partisan composition

Our identification hinges on changes in partisan composition of Congressional districts stemming from redistricting; thus, we require data on district composition (share of Democrats within a district) that is (or can be) measured immediately before and after redistricting. Thirty-one states track partisan affiliation of registered voters and make aggregate statistics publicly available; however, these statistics are usually reported at the county-level. In urban areas, Congressional districts often make up a small subset of a county; elsewhere, Congressional districts often cut across county borders. Thus, the official voter registration statistics are not immediately usable, as they do not capture the composition of a representative's electorate.

An alternative approach might be to use actual election results immediately before and after redistricting. One could take vote share received by a Democratic candidate as a measure of partisan composition within a district. Obviously, taking election results from US House elections as our measure of partisan composition would be problematic as it would introduce substantial endogeneity into our estimation approach. Vote share received by a particular candidate is, itself, an outcome driven by a variety of factors: personal characteristics of the candidate (and the opposing candidate), incumbency status, etc.

Given the lack of immediately available data that suits our needs, we construct our own measure capturing the *predicted* Democrat share for each Congressional district before and after each wave of redistricting in our sample period (1990's, 2000's, and 2010's). Broadly, we use demographic characteristics of Congressional districts to predict districts' partisan compositions. To do so, we draw on three sources of data: (1) county-level voter registration statistics from the year 2010, (2) Census data at the block group level for the years 1990, 2000, and 2010, and at the county level for 2010, and (3) Congressional district geographic boundary data, which we use to map Census block-groups into Congressional districts before and after redistricting.

The 2010 county-level registration statistics are drawn from Dave Leip's US Election Atlas (Leip, 2013). From these data, we can calculate *actual* Democrat shares at the county-level. We measure Democrat share in these data as the number of voters registered as Democrats in a district divided by the number of voters registered as Democrats or Republicans. Thus, when we discuss Democrat share throughout the paper, we are capturing Democrats' electoral strength relative to Republicans (ignoring independents or third parties). We do so to more closely map into the hypotheses we are testing (discussed in the previous section), which focus on competition between two parties. The Congressional district boundary data is taken from Lewis et al.'s (2013) compilation of Congressional district shapefiles for every Congress in US history; they have made these files available online.[8]

Broadly, our construction of predicted Democrat shares requires four steps. (We describe the construction of this variable very generally in this paragraph; full details are in an appendix.) First, we use the 2010 county-level voter registration data and the 2010 county-level demographic data to estimate coefficients that we will eventually use to predict Democrat share at other geographic levels. Second, we map Census block groups into pre- and post-redistricting Congressional districts. Third, we use those mappings to aggregate Census block group demographic data up to pre- and post-redistricting Congressional district-level demographics for each wave of redistricting. Finally, we use the coefficients from the estimation in the first step to construct *predicted* Democrat share at the Congressional district level, for every Congressional district immediately before and after every wave of redistricting. This leaves us with the "*DemShare*" variable that is used in the empirical approach described in the previous section. We

[8] http://cdmaps.polisci.ucla.edu/

summarize the distribution of the resulting measure in Figure 1, using a kernel density estimate plot.

Although we discuss the full details of the construction of Democrat share in the appendix, two details are worth emphasizing: (1) Census data and block group boundaries are held constant within each redistricting wave and (2) we do not rely on actual district names in employing district fixed effects; instead, we account for the fact that a given post-redistricting district may be essentially the same as some pre-redistricting district, but may have a different name. We elaborate on both of these points below (with additional detail in the appendix).

First, when we map block group data into Congressional districts before and after redistricting, we hold the source of the Census data constant. Consider, for instance, the 1990's wave of redistricting. The first elections impacted by redistricting were in November 1992. Thus, the 1989/1990 and 1991/1992 Congresses are considered "pre-1990's redistricting", while the 1993/1994 and 1995/1996 Congresses are considered "post-1990's redistricting" in our analysis. We use 1990 Census data to construct demographics and Democrat share for both the pre-1990's redistricting districts *and* post-1990's redistricting districts. This means that the shift in partisan composition of districts that we measure in our data stems *only* from redistricting and not from sorting in or out of Congressional districts. If we had used intercensal demographic estimates for the post-1990's redistricting demographics, we would be capturing a joint effect of redistricting and within-decade demographic shifts in the district. These latter shifts may be an endogenous response to redistricting, so we view the fact that we isolate variation stemming only from redistricting as an advantage of our data.

This fact, it should be noted, is the reason that we treat our data as three pooled panels (with four periods each), rather than one long panel (with twelve periods). If we treated the data

as one long panel, we would implicitly be comparing the 1st Alabama District *after* redistricting in the 1990's to the same district in the early 2000's, before 2000's redistricting. Indeed, this district has the same boundaries throughout this period, but – despite this – we would observe variation in the district's demographics because we shift from using 1990 Census data in the first case to 2000 Census data in the second case. This therefore would capture variation stemming *only* from sorting in or out of the district, which is obviously counter to our goal of observing the impact of changes in partisan composition resulting *only* from redistricting. For that reason, we silo each of the three redistricting waves and make only within-wave comparisons.

Table 1 summarizes the construction of our data with regards to usage of Census data and district mapping. As noted, our data is essentially set up as three pooled panels; each "redistricting wave" is a separate panel with a separate pre- and post-redistricting period. Each pre- and post-period consists of two Congresses. (This means we ultimately omit data from Congresses exactly half way in between two redistricting waves: e.g., the 105th.) The table illustrates the point we just made: within each redistricting wave, the Census data used is held constant (see the second to last column). The observed demographics of a given Congressional district change within a redistricting wave only because of redistricting, which is reflected by the fact that we use a different mapping of blocks to districts in the pre- and post- phases of each redistricting wave (see the final column).

The second detail worth noting is that we do not rely on the name of a Congressional district to match a district just before redistricting to the same district just after redistricting. In many cases, especially when seats are gained or lost, a district that is called "District 1" before redistricting may overlap very little with the district called "District 1" afterwards; conversely, a district that *most* overlaps with "District 1" after redistricting could very well be assigned an entirely different

district number. Thus, in implementing district fixed effects, we construct our own district naming based on our own matching of pre- and post-redistricting districts.

We match pre- and post- districts using our block-group level population data. Given our mapping of block-groups to districts, we can identify the districts that block-groups were assigned to under every drawing of districts and can therefore calculate the population-based overlap in two districts before and after redistricting. For instance, if District 7 is dissolved and split between Districts 5 and 6, we can identify what fraction of the population that was in District 7 (based on the nearest Decennial Census) is now in each of the two new districts (based on the same Decennial Census). We then say that some pre-redistricting district "X_{pre}" and some post-redistricting district "Y_{post}" are matched if: (1) of all the post-redistricting districts that have any population overlap with District X_{pre}, District Y_{post} has the most overlap, and (2) of all the pre-redistricting districts that have any population overlap with District Y_{post}, District X_{pre} has the most overlap. If two districts satisfy these conditions, we rename them in our data accordingly: that is, we would code both "X_{pre}" and "Y_{post}" as "X_{match}" in our data, and use the matched name (X_{match}) when we employ district fixed effects.

3.2 Measure of ideological position of Congressional representatives

The main outcome variable used throughout our paper is the first dimension of the "DW-Nominate" score (Poole & Rosenthal, 2007). Specifically, because we are interested (in some analyses) in how an individual Congressperson's ideological position changes from one Congress

to the next, we use the period-specific version of the score as introduced in Nokken & Poole (2004).[9]

The first dimension of the DW-Nominate score locates a politician along a left-right continuum based on their roll call voting patterns in a particular Congress. Increasingly positive DW-Nominate scores indicate a politician whose roll call voting is generally farther to the right, while increasingly negative scores indicate a politician who is farther left. Thus, in our sample, Republicans tend to have positive scores and Democrats tend to have negative scores. This is documented in the top panel of Table 2, which summarizes the DW-Nominate scores in our sample. The top row summarizes the scores in the full sample, while the two following rows split the sample by Democrats and Republicans.

In additional analyses, we test how a shift in racial composition of a district impacts voting on issues specifically related to race and civil rights. To do so, we draw on the scores given to members of Congress by the *Leadership Conference on Civil Rights* (LCCR). They choose a small number of bills from each Congress that are important to race and civil rights and rate members of Congress based on the number of times they vote in line with LCCR's positions. LCCR scores have been used widely in the political science literature as a measure of the quality of black representation in Congress (e.g., Avery & Fine, 2012; Grose, 2005; Whitby & Krause, 2001). We use a version of these scores that have been adjusted by Groseclose, Levitt, and Snyder (1999) to allow for intertemporal comparisons. The adjusted scores are only available through the 110th Congress; as a result, our analyses taking LCCR scores are restricted to the 1990s and 2000s

[9] The standard DW-Nominate score varies over time, but is forced to evolve linearly. Essentially, each Congressperson is assigned a DW-Nominate score at the beginning and end of his or her career; for periods in between the beginning and end, the score is assigned through linear interpolation. This is not true of the Nokken & Poole (2004) version.

redistricting waves. The LCCR scores (overall and split by party) are summarized in the bottom panel of Table 2.

4. Documenting the impact of redistricting on predicted Democrat share

Before proceeding to our main results, we first graphically document the impact of redistricting on changes in partisan composition. To do so, in Figure 2 we plot a two-dimensional kernel density estimate. Each observation in the data is a single Congressional district during a particular wave of redistricting (matched across redistricting as described in the data section). For each district-by-redistricting wave pair, we observe the predicted Democrat share prior to redistricting and the change in Democrat share resulting from redistricting. In Figure 2, pre-redistricting Democrat share is on the x-axis, while change in Democrat share as a result of redistricting is on the y-axis. Thus, if no districts changed as a result of redistricting, all points would fall along the line y=0; positive (negative) values on the y-axis indicate gains (lossed) in the predicted share of Democrats within a district.

Our goal in reporting this figure is twofold: First, our construction of "predicted Democrat share" is novel, so we generate this plot to further summarize the data. Moreover, most immediately available measures of partisan composition of districts do not allow for such a clear picture of the relationship between pre-existing partisan composition and redistricting-prompted changes in composition (while holding demographic data constant), so the figure is of interest in its own right to illustrate the impacts of redistricting. Second, we are interested in assessing whether "heavily treated" districts are clearly different than "less treated" districts along observable dimensions in ways that may threaten the validity of the research design.

Figure 2 reveals that beyond a very small regression to the mean (a 1 percentage point increase in baseline Democrat share is, on average, associated with a .034 percentage point decrease in Democrat share following redistricting), there is no clear relationship between pre-existing Democrat share and the conditional distribution of change in Democrat share. Importantly, while most districts change very little, there is still substantial variation across the entire support pre-existing Democrat share.

Still, the basic concern is that there is some fact that explains why some areas gain (or lose) Democrats, while others do not, and that this fact also explains changes in roll-call voting behavior. To probe this possibility, in Appendix 2, we provide a series of plots and regression coefficients associated with splitting our sample along a variety of dimensions that may be associated with the level/type of selection involved in the redistricting process. If non-randomness in redistricting is driving the redistricting process, one might reasonably expect the form that redistricting takes to vary across these different subsamples. We split the sample by:

(1) states that must redistrict because they gained/lost seats vs. those that did not gain/lose seats,

(2) states where the redistricting authority (e.g., state legislature) is dominated by Democrats vs. dominated by Republicans,

(3) districts with Republican incumbents vs. Democrat incumbents,

(4) districts where DW-Nominate scores of elected representatives has trended left in the preceding decade vs. trended right,

(5) redistricting that occurred in the 1990s vs. 2000s vs. 2010s, and

(6) redistricting by executive and legislative branches vs. courts vs. redistricting commissions.

We find no evidence of meaningful differences in the redistricting process across any of these comparisons.[10]

5. Results

Our primary goal is to assess whether a shift in partisan composition impacts the ideological position of a district's representative. More importantly, we aim to understand *why* this might happen: Does an increase in the number of Democrats in a district lead to more leftist representation only because a leftist is more likely to be elected, or is there a *direct* effect of partisan composition on representatives' behavior (as predicted by the Downsian model)?

5.1 Does "predicted Democrat share" predict success for Democratic candidates?

Prior to addressing these questions, it is important to first confirm that our predicted Democrat share measure – and the redistricting-driven variation in this measure – meaningfully captures an increase in electoral strength for Democratic candidates. To do so, we estimate the basic specification described in Section 2 (a continuous difference-in-differences model with district-by-redistricting wave fixed effects), but instead of taking "DW-Nominate" as the outcome variable we take two measures of Democratic electoral success.

Results are reported in Table 3. Column 1 reports the result of a linear probability model taking as the outcome a dummy variable equal to one if a Democrat represents the district. In Column 2, we take the vote share received by the Democratic candidate in the election for the relevant Congress as the outcome (noting that we observe electoral vote shares for only a subset

[10] As noted in the appendix, within each sample splitting category, we conduct pairwise tests of differences between lines fit to the data to assess differences in patterns of redistricting. None of these tests reveal a statistically significant difference.

of our observations). In either case, our constructed "predicted Democrat share" measure is clearly related to an increase in Democrat electoral strength. Keep in mind that within the set of observations for a given district (within a particular redistricting wave), variation in predicted Democrat share stems *only* from redistricting. Thus, the interpretation of the coefficient in Column 1 is: if the predicted share of Democrats within a district is increased by 10 percentage points as a result of redistricting, then the likelihood of a Democrat winning increases by 14 percentage points. Similarly, a 10 percentage point increase in predicted Democrat share within a district leads to an additional 8 percentage points in vote share (Column 2).

As further evidence, Figure 3 presents a binned scatterplot, relating predicted Democrat share (x-axis) to frequency of Democratic representation within bins of the x-axis variable (y-axis). Here we see that, as we would expect, there is nonlinear relationship between the probability that a Democrat represents a district and the predicted Democrat share within that district. When districts are competitive (close to 50% Democrat share), a small increase in the Democrat share of the district dramatically increases the likelihood that a Democrat holds the seat. When districts are uncompetitive (far from 50% Democrat share), an increase in Democrat share has much less impact on the likelihood of a Democrat being elected.

The pattern of results in Table 3 and Figure 3 is important for two reasons: first, they document the validity to our constructed measure. Second, one of our goals in this paper is to understand how a change in Democrat share impacts representative's behavior even when the shift in electorate composition is not large enough to change *who* represents the district. Figure 3 suggests that there is substantial scope to do so.

5.2 How does partisan composition of an electorate impact a representative's ideological position?

Having documented that our constructed measure of Democrat share performs well in predicting the election of a Democrat, we turn to the main question of the paper. Here, the outcome of interest is the period specific DW-Nominate score. As a reminder, increasingly leftist voting corresponds to a score that is increasingly negative, while conservative roll call voting yields a positive DW-Nominate score.

Our first results on this issue are reported in Table 4. Column 1 reports results from the most basic form of the estimating equation described in the methodology section. We regress DW-Nominate scores on predicted Democrat share, district (by redistricting wave) fixed effects, and Congress fixed effects. Thus, this specification captures the combined, within district, effect of the extensive and intensive margins. On the extensive margin, a higher predicted Democrat share is associated with a higher likelihood that a Democrat is elected (as documented in Table 3). In the absence of complete policy convergence, we expect this to generate more leftist roll call voting. On the intensive margin, if individual politicians directly respond to changing preferences in their electorate (rather than strictly implementing their personally preferred policies), this too could push roll call voting further left. As we are combining these two effects and we have already documented a substantially higher likelihood of electing a Democrat when Democrat share increases, it is perhaps not surprising that Column 1 indicates a sizable negative (leftward) effect of Democrat share on the DW-Nominate score. This result establishes the baseline overall effect of a shift in partisan composition. The question of interest now turns to what drives this leftward push. What fraction of this overall effect is driven by the intensive rather than extensive margin effect?

We adopt several strategies to decompose the extensive and intensive margin effects. We first note that simply controlling for partisan affiliation of the elected representative (as we do in Column 2 of Table 4) leads to a smaller but still highly significantly negative coefficient on predicted Democrat share.

We adopt more direct means of isolating the intensive margin effect in Table 5. In Column 1 of Table 5, we restrict the sample to districts where the partisan affiliation of the representative does not change before and after redistricting. That is, partisan affiliation remains constant across the four Congress that make up the pre and post periods of each redistricting wave. Note that there are many districts for which this is true (as previewed in Figure 3); we lose only 21% of our sample with this restriction. Thus, far more often than not, a shift in Democrat share does not have an effect on partisan affiliation of representatives.

If the increase in leftist voting found in Table 4 was entirely driven by a change in the party representing a district (the main extensive margin effect), then predicted Democrat share would be unrelated to DW-Nominate scores after restricting our sample. This is not the case. Instead, in Column 1 of Table 5, we again find a clear negative (leftist) impact of Democrat share on the DW-Nominate score. Comparing the magnitude of this coefficient to Column 1 of Table 4, the results suggest that only 60% of the leftward shift is driven by increased likelihood of electing a Democrat. The fact that it is 60% rather than 0% confirms that partisan affiliation is important for policymaking and that policy divergence appears to be present in the US House (as others have documented); however, while some degree of policy divergence is present, it is not complete: 40% of the leftward shift in roll call voting in response to a shift in the composition of the electorate is *not* explained by this partisan affiliation effect. Thus far, our results therefore suggest that both the extensive and intensive margins are important.

Of course, as noted in a previous section, there is a second type of extensive margin effect which may still explain the result in Column 1 of Table 5. In particular, even if a change in Democrat share is not large enough to change *which party* is elected, it may still impact the candidate that is put forth *within* a party. That is, an increase in the number of Democrats within an otherwise right-leaning district may not lead the district to elect a Democrat, but it may increase the likelihood of electing a centrist Republican. Note, of course, that this story is more nuanced: this is only an extensive margin effect if the centrist Republican in question genuinely holds centrist preferences and would adopt them if elected independent of the composition of the district.

To eliminate the influence of even this more nuanced extensive margin effect, our most robust specification includes individual representative fixed effects.[11] In this specification, we identify changes in individual representatives' DW-Nominate scores in response to redistricting-generated changes in the partisan composition of the district they represent. Identification, therefore, is implicitly based on individuals who were present in Congress immediately before and after redistricting. This specification strips away the influence of any form of extensive margin effect.

Column 2 of Table 5 reports the result. The negative effect of an increase in Democrats within a district survives even in this specification. The coefficient is smaller, but is still of substantial magnitude. As a point of comparison, the within party standard deviation in DW-Nominate score is 0.21 for Republicans and 0.15 for Democrats. With this result, we can confidently say that – while there is an extensive margin effect – there is also clearly an intensive margin effect. An increase in the number of Democrats within a district leads individual representatives to change the way the vote. This confirms that, even if complete policy

[11] To be more specific, like our district fixed effects, they are representative-by-redistricting wave fixed waves. Thus, for each Congressperson, we identify the impacts of each redistricting wave separately.

convergence is not observed, at least one basic prediction of Downsian-type models is observed in the US House: individual politicians' policy positions move with the preferences of their electorates.

In the remaining columns of Table 5, we assess whether the result we have documented is restricted to just one party. Do both Democrats *and* Republicans shift to the left when there are more Democrats in their district? In Columns 3 and 4, respectively, we repeat the specifications of Column 1 (restricting the sample to districts with no party change) and Column 2 (individual level fixed effects), but we now allow for differential effects for Republican and Democratic representatives. The basic pattern of results – a clear negative effect of Democrat share on roll call voting – is true whether representative is a Democrat or Republican. The magnitudes are slightly different across parties, but not significantly so.

5.3 Robustness checks – DW Nominate

Our remaining results probe the validity of our empirical approach. While our analysis of the raw redistricting data showed no evidence of systematic non-randomness. Selection remains a potential concern. Empirically, we adopt three approaches to rule out that the factors that drive redistricting do not drive our results: 1) we include district- and individual-level time trends; 2) we partition the sample into states where redistricting is more or less likely to have partisan motivations; and, 3) we evaluate the impact of racial composition on voting relative to the LCCR's voting priorities. We begin by probing the robustness of the intensive margin result as that is our main contribution.

Table 6 reports the results of specifications that are similar to the first two columns of Table 5, except that we include unit-specific time trends. Column 1 of Table 6 repeats the specification

wherein we restrict the sample to districts where party does not change after redistricting, but with district-specific time trends included on the right hand side. Column 2 of Table 6 uses the full sample but includes individual representative fixed effects; there, we add individual representative time trends. In both cases the goal is to eliminate the influence of pre-existing trends in ideological position that may have caused redistricting authorities to target particular districts. This concern does not appear to drive our results; the results in Table 6 are nearly identical to the results in Table 5, even with the time trends.

Next, we test whether our results are different in states where there might be of particular concern that redistricting is endogenous to pre-existing ideology and Democrat share as compared to states where the reassignment of census block groups to different congressional districts is more plausibly exogenous. Specifically, we split the sample in three ways: first, we split the sample by states where redistricting was or was not conducted by a unified partisan state government (or subset of state government). For instance, many states require that redistricting is decided upon by the state legislature with approval by the governor. In these cases, we code a state as "partisan" in their redistricting if the majority party of the legislature and the governor are of the same party when redistricting occurs. If the separate branches are not all of the same party, we code the redistricting as "not partisan". Other states make use of an independent nonpartisan redistricting commission. These states are coded as nonpartisan regardless of the partisan balance in the state legislative and executive branches. We interact our main treatment variable (predicted Democrat share) with a dummy indicating whether a state's redistricting process was partisan or not partisan. Results are reported in the first panel of Table 7. The first column includes district fixed effects and restricts attention to districts where the partisan affiliation of the representative did not change; the second column uses the full sample but includes individual fixed effects. We find that whether

redistricting was conducted by a partisan or nonpartisan subset of government, the basic result holds: an increase in Democrat share within a district leads to more leftist voting.

Following the same idea, we split the sample in a second way. Specifically, if we are concerned that non-random drawing of Congressional districts impacts our results, this is likely to be more of a problem in states that did not gain or lose any seats as a result of the post-Censal reapportionment. They have much greater freedom to redistrict for political purposes. States that did gain or lose seats are forced to redistrict to accommodate the change in seats, and may be more constrained in their ability to draw districts for political gain. Results are reported in the second panel of Table 7. Once again, the basic pattern of results holds.

Finally, we split the sample by the authority responsible for redistricting: state legislative and executive branches versus courts or redistricting commissions.[12] Existing research has documented that the authority responsible for redistricting can have an impact on competitiveness of Congressional elections, with courts and commissions leading to more competitive elections than legislative redistricting (Carson & Crespin, 2004). Results are reported in the third panel of Table 7, and – again – reveal the same pattern of results as our main specifications regardless of the redistricting authority (albeit with reduced precision).

5.4 Robustness Check - Response to shifts in racial composition of districts

Thus far, our results document that an increase (decrease) in the predicted Democrat share of a Congressional district leads to more liberal (conservative) roll-call voting behavior, and that this change occurs both because of an increased (decreased) likelihood of electing a Democrat

[12] We combine courts and commissions into a single category in the analysis due to the smaller numbers of observations in those categories. Combined, they account for roughly 40% of observations, with legislature/governor redistricting accounting for the remaining 60%.

("extensive margin") and because of changes in how individual representatives vote ("intensive margin"). This response at the intensive margin, which contrasts with some existing empirical literature, is consistent with predictions of the classic Downsian model, wherein politicians' policy positions move towards the preferences of voters within their districts. There is, however, an important alternative explanation for the observed change in roll call voting: rather than responding to voter preferences, politicians may simply be responding to increased or decreased competition. Decreased competition (e.g., a Democratic district becoming more Democratic) may lead politicians to feel free to indulge in their own preferred policy outcomes.

In this subsection, we provide a related empirical test to assess whether politicians are genuinely responding to the composition of their districts. Specifically, we test whether an increase in the percent of residents of a district who are black impacts how representatives vote on issues related to race and civil rights. If our previous results were genuinely driven by a reaction to voter preferences within the district, one would expect that racial composition would also directly impact voting on issues related to race. If, on the other hand, our previous results were strictly driven by changes in electoral strength and competition within the district, then shifts in Democrats and Republicans within the district should matter for roll-call voting, but shifts in other types of district composition (e.g., race) should not.

To test this hypothesis, we modify our existing empirical approach. We replace DW-Nominate scores with Leadership Conference on Civil Rights (LCCR) ratings as our outcome variable. The LCCR ratings capture the extent to which representatives vote in favor of African Americans and civil rights issues during a particular session of Congress. We use an adjusted version which is comparable across Congresses (Groseclose, Levitt, and Snyder, 1999). The measure is increasingly positive when voting records on civil rights issues are more in line with

the policy positions of the LCCR. On the right hand side, rather than measuring shifts in predicted Democrat share, we consider the impact of shifts in percent black within a district caused by redistricting.[13] As noted in the data section, the adjusted LCCR scores are only available through the 110th Congress; the analyses in this subsection are therefore based on just the 1990s and 2000s waves of redistricting. Beyond those changes, the structure of the specifications are otherwise similar.

Results are reported in Table 8. Column 1 reports results from a specification including only district fixed effects; it therefore captures the overall effect (combining intensive and extensive margin effects) of a change in percent black within a district. Not surprisingly, an increase in percent black within a district leads to a more supportive record on civil rights issues. To provide a sense of the magnitude of the coefficient in Column 1, a 10 percentage point increase in percent black is associated with a 0.25 standard deviation increase in the LCCR score. Of course, the percent black within a district is correlated with the (predicted) Democrat share[14], so one concern is that we are simply controlling for a proxy for partisan composition (and therefore potentially a proxy for electoral strength and/or competition). The specification in Column 2 differs from that of Column 1 only in that it includes a control for predicted Democrat share. The strong influence of the LCCR score survives even when controlling for predicted Democrat share.

Columns 3 and 4 restrict the sample to districts where the party of the elected representative does not change during the redistricting wave, and therefore eliminates the extensive margin effect of electing a representative from a different party. Still, there is a large and significant effect of an

[13] Our measure of "percent black" is constructed in the same way as our "predicted Democrat share" measure. That is, for each redistricting wave, we hold fixed the demographic data we use (using the most relevant Census data) and simply lay pre- or post-redistricting maps over Census block groups to obtain changes driven only by redistricting. The main difference is that percent black can be measured directly, so there is no need for "predicted percent black".
[14] In our sample, the correlation coefficient between predicted Democrat share and percent black is 0.5817.

increase in black population within a district and the LCCR score. The coefficient in Column 3 is roughly half the size of the coefficient in Column 1 suggesting that the extensive and intensive margin effects contribute roughly equally to the overall effect. As in Column 2, Column 4 adds a control for predicted Democrat share. Again, the significant influence of percent black survives. (In fact, in this specification eliminating the influence of partisan extensive margin effects, *only* percent black is predictive. The predicted Democrat share does not significantly influence LCCR scores in this specification.)

Columns 5 and 6 report the results of including individual congressperson fixed effects, thereby eliminating any further intraparty extensive margin effects.[15] As in previous analysis, the result survives even with the inclusion of individual fixed effects. In Column 6, we include both percent black and predicted Democrat share as controls. We find a result that is nearly identical to Column 5.

In short, the pattern of results from Table 8 are remarkably similar to our main results. In both cases, the roll call voting behavior of representatives appears to genuinely respond to the composition of their district.

6. Discussion and Concluding Remarks

The question of how voter preferences are translated into policy is especially pertinent in the current electoral environment. Congressional races in the US have become less and less competitive in recent decades (Friedman & Holden, 2009). In many districts, then, a very large

[15] Doing so may be particularly relevant here, as we do not control for the race of the member of Congress; it may be that a district that is always Democratic leaning continues to elect Democrats after redistricting, but is more likely to elect an African America representative if percent black within the district increases. Assuming black members of Congress have a more favorable voting record on civil rights issues, this could explain the increase in the LCCR score even in observations where the party of the representative does not change after redistricting.

shift in voter preferences would be required to elect a candidate from the opposing party or unseat the incumbent. If it is indeed the case that voters' preferences only filter into policy decisions by impacting *who* is elected, then this decrease in competition in Congressional elections may imply an increasing disconnect between voter preferences and representation in Congress. On the other hand, if sitting Congressional representatives respond to even small shifts in preferences of their voters (as our results suggest), the problem – while still important – may not be so severe.

One could reasonably expect that an increase in the number of Democrats in a district would lead to more leftist representation in Congress. Consistent with earlier work (Levitt, 1996; Gerber & Lewis, 2004), we find strong evidence in support of this linkage. However, the main focus of our analysis is on understanding what drives this relationship. Broadly speaking, the theoretical and empirical literatures put forth two hypotheses. Either: (1) politicians aim to maximize votes by adopting policies that please their electorates, in which case a shift in the preferences of the electorate will directly impact the way politicians vote in Congress, or (2) once elected, politicians enact their personally preferred policies and voters only impact legislative voting through their choice of candidate. The first of these two hypotheses stems from the Downsian model of electoral competition (Downs, 1957). One prediction of this model – that politicians from different parties adopt identical policies – has been subjected to substantial empirical scrutiny, to mixed results. Much less empirical work has focused on whether politicians' roll call voting responds in any way to shifts in electoral strength or the electorate's preferences, even if evidence of pure policy convergence is not observed. That is, even if the "median voter theorem" does not strictly hold, do politicians still respond to the median voter, albeit in an attenuated fashion? Our study provides new evidence on this issue.

We take advantage of variation in the partisan composition of Congressional districts which stems from Census-initiated redistricting in the early 1990's, 2000's, and 2010's. Using this variation, we assess how partisan composition impacts representative's roll call voting behavior. Ultimately, we find that an increase in the share of Democrats within a district impacts roll call voting in two ways. First, an increase in electoral strength for Democrats indeed leads to a higher likelihood of Democrats being elected, which in turn is associated with more leftist representation. This result is consistent with the second hypothesis discussed above, wherein politicians enact their preferred policy if elected and voters choose between them. However, in contrast to existing empirical work implemented using different approaches (Lee, Moretti, and Butler, 2005; Fedaseyeu et al., 2015), we find that this "extensive margin" or party effect does not explain the entire relationship between partisan composition and roll call voting. Instead, there is also a direct effect reminiscent of politicians' responses in the Downsian model. We find that an increase in Democrat share leads to more leftist roll call voting even when we isolate within-party and/or within-representative changes.

Thus, we find that both the extensive and intensive margins are important. Specifically, our estimates suggest that a 10 percentage point increase in predicted Democrat share leads to a roughly 1/3rd standard deviation leftward shift in the DW-Nominate score of that district's representation. Restricting our attention to observations where the party representing a district does not change, the estimated impact of a 10 percentage point increase in predicted Democrat share shrinks to a leftward shift of roughly 0.13 standard deviations or one third of the within-party standard deviation in DW-Nominate score. Compared to the overall effect, this result suggests that party switching accounts for approximately 60% of the overall effect. The remaining 40% is driven

by a combination of extensive margin effects occurring within the party (e.g., more liberal candidates being chosen in primaries) and changes in individual Congressperson voting behavior. Including Congressperson fixed effects, we find that 2/3rds of the within-party effect is driven by the selection of more liberal candidates at the primary stage, with the remaining 1/3rd explained by movement to the left of individual members of Congress who retain their seat after redistricting.

Finally, extending our analysis to evaluate the impact of the size of a district's black electorate on legislator voting relative to the priorities of the Leadership Conference on Civil Rights, we find a very similar pattern of results - both in terms of the magnitudes and relative importance of the extensive and intensive margin effects.

References

Albouy, D. (2011). Do voters affect or elect policies? A new perspective, with evidence from the US Senate. *Electoral Studies*, *30*(1), 162-173.

Albouy, D. (2013). Partisan representation in Congress and the geographic distribution of Federal funds. *Review of Economics and Statistics*, *95*(1), 127-141.

Alesina, A. (1988). Credibility and policy convergence in a two-party system with rational voters. *The American Economic Review*, 796-805.

Avery, J. M., & Fine, J. A. (2012). Racial composition, white racial attitudes, and black representation: Testing the racial threat hypothesis in the United States Senate. *Political Behavior*, *34*(3), 391-410.

Beland, L.P. (2015) "Political Parties and Labor Market Outcomes. Evidence from U.S. States," *American Economic Journal: Applied Economics*, 7(4): 198-220.

Bertelli, A. M., & Carson, J. L. (2011). Small changes, big results: Legislative voting behavior in the presence of new voters. *Electoral Studies*, *30*(1), 201-209.

Besley, T., & Coate, S. (1997). An economic model of representative democracy. *The Quarterly Journal of Economics*, 85-114.

Boatright, R. G. (2004). Static Ambition in a Changing World: Legislators' Preparations for, and Responses to, Redistricting. *State Politics & Policy Quarterly*, *4*(4), 436-454.

Caughey, D., & Sekhon, J. S. (2011). Elections and the regression discontinuity design: Lessons from close US house races, 1942–2008.*Political Analysis*, *19*(4), 385-408.

Carson, J. L., & Crespin, M. H. (2004). The effect of state redistricting methods on electoral competition in United States House of Representatives races. *State Politics & Policy Quarterly*, *4*(4), 455-469.

Crespin, M. H. (2010). Serving two masters: redistricting and voting in the US House of Representatives. *Political Research Quarterly*, *63*(4), 850-859.

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

Fedaseyeu, V., Gilje, E., & Strahan, P. E. (2015). *Voter Preferences and Political Change: Evidence from Shale Booms* (No. w21789). National Bureau of Economic Research.

Ferreira, F., & Gyourko, J. (2009). Do Political Parties Matter? Evidence from US Cities. *The Quarterly Journal of Economics*, *124*(1), 399-422.

Friedman, J. N., & Holden, R. T. (2009). The Rising Incumbent Reelection Rate: What's Gerrymandering Got to Do With It?. *The Journal of Politics*,*71*(02), 593-611.

Gerber, E. R., & Hopkins, D. J. (2011). When mayors matter: estimating the impact of mayoral partisanship on city policy. *American Journal of Political Science*, *55*(2), 326-339.

Gerber, E. R., & Lewis, J. B. (2004). Beyond the median: Voter preferences, district heterogeneity, and political representation. *Journal of Political Economy*, *112*(6), 1364-1383.

Grose, C. R. (2005). Disentangling Constituency and Legislator Effects in Legislative Representation: Black Legislators or Black Districts?. *Social Science Quarterly*, *86*(2), 427-443.

Groseclose, T., Levitt, S. D., & Snyder, J. M. (1999). Comparing interest group scores across time and chambers: Adjusted ADA scores for the US Congress. *American political science review*, *93*(01), 33-50.

Hill, A., & Jones, D.B. (2016). Governors and the distribution of education spending. *Unpublished manuscript.*

Lee, D. S., Moretti, E., & Butler, M. J. (2004). Do voters affect or elect policies? Evidence from the US House. *The Quarterly Journal of Economics*, 807-859.

Leigh, A. (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.

Leip, D. (2013). Voter Registration & Turnout Data by County. *Dave Leip's Atlas of US Presidential Elections*.

Levitt, S. D. (1996). How do senators vote? Disentangling the role of voter preferences, party affiliation, and senator ideology. *The American Economic Review*, 425-441.

Lewis, J.B., DeVine, B., Pitcher, L. and Martis, K.C.. (2013). *Digital Boundary Definitions of United States Congressional Districts, 1789-2012.* Retrieved from http://cdmaps.polisci.ucla.edu.

Llavador, H. (2006). Electoral platforms, implemented policies, and abstention. *Social Choice and Welfare*, *27*(1), 55-81.

McCarty, N., Poole, K. T., & Rosenthal, H. (2009). Does gerrymandering cause polarization?. *American Journal of Political Science*, *53*(3), 666-680.

Nokken, T. P., & Poole, K. T. (2004). Congressional party defection in American history. *Legislative Studies Quarterly*, *29*(4), 545-568.

Osborne, M. J., & Slivinski, A. (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics*, 65-96.

Poole, K. T., & Rosenthal, H. (2007). *Ideology and Congress: of Congress: A Political Economic History of Roll Call Voting*. Transaction Publishers.

Reed, W. R. (2006). Democrats, republicans, and taxes: Evidence that political parties matter. *Journal of Public Economics*, *90*(4), 725-750.

Whitby, K. J., & Krause, G. A. (2001). Race, issue heterogeneity and public policy: the Republican revolution in the 104th US Congress and the representation of African-American policy interests. *British Journal of Political Science*, *31*(03), 555-572.

Tables

Table 1: Summarizing set up of data, timing of redistricting, and source of Census and mapping data

Years	Congress	Redistricting wave	Pre/Post Redistricting	Census data	District mapping
1989-1990	101	1990's	Pre	1990	102nd
1991-1992	102	1990's	Pre	1990	102nd
1993-1994	103	1990's	Post	1990	103rd
1995-1996	104	1990's	Post	1990	103rd
1999-2000	106	2000's	Pre	2000	107th
2001-2002	107	2000's	Pre	2000	107th
2003-2004	108	2000's	Post	2000	108th
2005-2006	109	2000's	Post	2000	108th
2009-2010	111	2010's	Pre	2010	112th
2011-2012	112	2010's	Pre	2010	112th
2013-2014	113	2010's	Post	2010	113th

Table 2: Summary statistics: Outcome variables

	Mean	Std. dev.	Min.	Max.	Obs.
DW-Nominate Score (1st dimension)					
Full Sample	0.08	0.48	-0.79	1	4228
Republicans	0.55	0.21	-0.07	1	2028
Democrats	-0.35	0.15	-0.79	0.282	2200
LCCR Score (Groseclose-Levitt-Snyder adjusted for intertemporal comparison)					
Full Sample	44.95	41.17	-24.31	102.14	2994
Republicans	5.53	16.81	-24.31	96.00	1396
Democrats	79.39	19.61	-1.08	102.14	1598

Table 3: The impact of predicted Democrat share on electoral outcomes

VARIABLES	(1) Pr(Democrat holds seat)	(2) Democrat vote share
Pred. Dem. Share	1.447***	0.807***
	(0.328)	(0.136)
Dist.*RD Wave FE's	X	X
Congress FE's	X	X
Observations	4,228	3,781
R-squared	0.852	0.778

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: The impact of predicted Democrat share on representatives' roll call voting – Overall effects

VARIABLES	(1) DW-Nom. (1st dim.)	(2) DW-Nom. (1st dim.)
Pred. Dem. Share	-1.612***	-0.547***
	(0.308)	(0.135)
Democrat		-0.736***
		(0.0201)
Dist.*RD Wave FE's	X	X
Congress FE's	X	X
Observations	4,228	4,228
R-squared	0.891	0.978

Robust standard errors (clustered at level of state-by-redistricting wave) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: The impact of predicted Democrat share on representatives' roll call voting – Intensive margin effects

VARIABLES	(1) DW-Nom. (1st dim.)	(2) DW-Nom. (1st dim.)	(3) DW-Nom. (1st dim.)	(4) DW-Nom. (1st dim.)
Pred. Dem. Share	-0.661***	-0.209**		
	(0.162)	(0.0975)		
Dem. X Pred. Dem. Share			-0.707***	-0.184**
			(0.203)	(0.0872)
Repub. X Pred. Dem. Share			-0.570***	-0.241
			(0.213)	(0.176)
Sample restriction	No party change		No party change	
District*RD wave	X		X	
District*RD wave*Rep. FE's		X		X
Congress FE's	X	X	X	X
Observations	3,347	4,228	3,347	4,228
R-squared	0.985	0.992	0.985	0.992

Robust standard errors (clustered at level of state-by-redistricting wave) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: The impact of predicted Democrat share on representatives' roll call voting – Intensive margin effects with time trends

VARIABLES	(1) DW-Nom. (1st dim.)	(2) DW-Nom. (1st dim.)
	District trends (and FEs) in districts with no party change	Individual specific trends and FEs
Pred. Dem. Share	-0.601**	-0.225*
	(0.264)	(0.129)
District-specific trends	X	
Person-specific trends		X
District*RD wave FE's	X	
District*RD wave*Person FE's		X
Congress FE's	X	X
Observations	3,347	4,228
R-squared	0.994	0.996

Robust standard errors (clustered at level of state-by-redistricting wave) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Assessing heterogeneity in impact of Democrat share by nature of redistricting

VARIABLES	(1) DW-Nom. (1st dim.)	(2) DW-Nom. (1st dim.)
Partisan redist. X Dem. share	-0.635***	-0.159
	(0.226)	(0.139)
Not partisan redist. X Dem. share	-0.695***	-0.260*
	(0.213)	(0.135)
Observations	3,347	4,228
R-squared	0.985	0.992
Gained/lost seats X Dem. share	-0.577***	-0.191*
	(0.162)	(0.115)
No gained/lost seats X Dem. share	-1.010**	-0.285*
	(0.411)	(0.160)
Observations	3,345	4,226
R-squared	0.985	0.992
Leg. & Gov. redist. X Dem. share	-0.628***	-0.223
	(0.200)	(0.142)
Court/Commission redist. X Dem. share	-0.679**	-0.165
	(0.273)	(0.132)
Observations	3,228	4,083
R-squared	0.985	0.992
Sample restriction	No party change	
District*RD wave	X	
District*RD wave*Rep. FE's		X
Congress FE's	X	X

Robust standard errors (clustered at level of state-by-redistricting wave) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Impact of African American population within district on LCCR score

VARIABLES	(1) Adj. LCCR Score	(2) Adj. LCCR Score	(3) Adj. LCCR Score	(4) Adj. LCCR Score	(5) Adj. LCCR Score	(6) Adj. LCCR Score
Pct. black in pop.	102.610***	56.628**	51.122***	45.156***	18.635*	19.273*
	(21.648)	(22.357)	(15.412)	(17.059)	(9.731)	(11.240)
Pred. Dem. share		89.454***		11.985		-1.657
		(31.764)		(18.386)		(11.665)
Sample restriction			No party change	No party change		
District*RD wave	X	X	X	X		
District*RD wave*Rep. FE's					X	X
Congress FE's	X	X	X	X	X	X
Observations	2,994	2,994	2,391	2,391	2,994	2,994
R-squared	0.873	0.874	0.965	0.965	0.975	0.975

Robust standard errors (clustered at level of state-by-redistricting wave) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figures

Figure 1: Kernel density estimate of predicted Democrat share

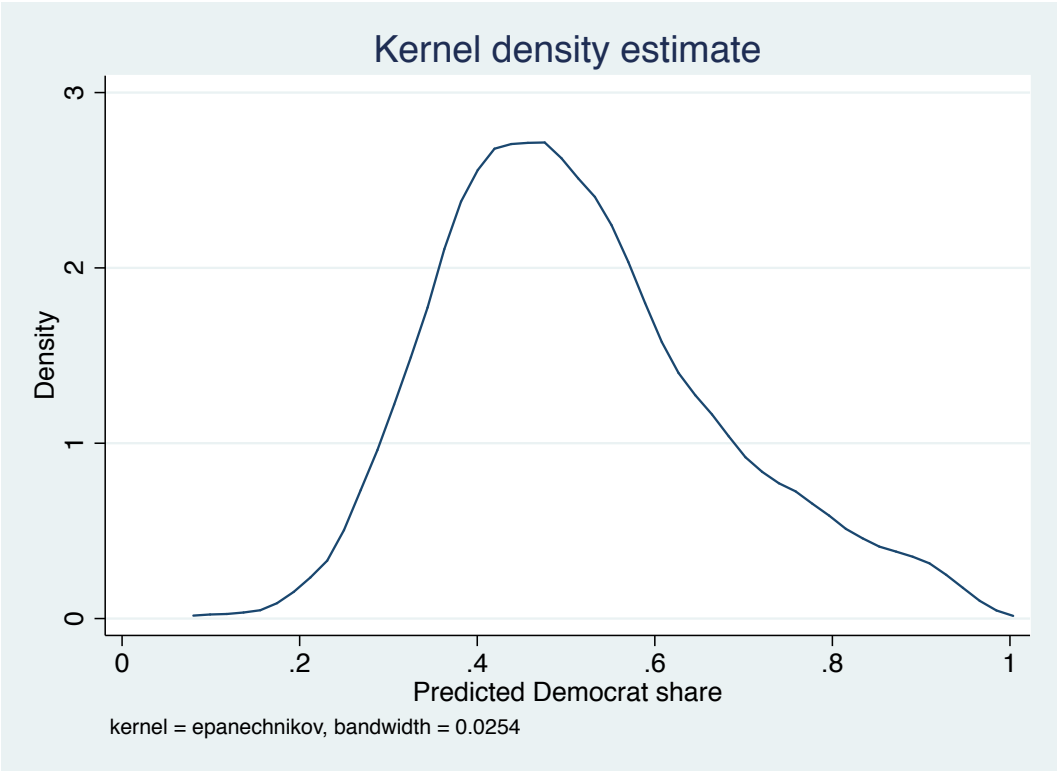


Figure 2: Two-dimensional kernel density estimate of *Pre-redistricting democrat share* and redistricting prompted *Change in Democrat share* pairings

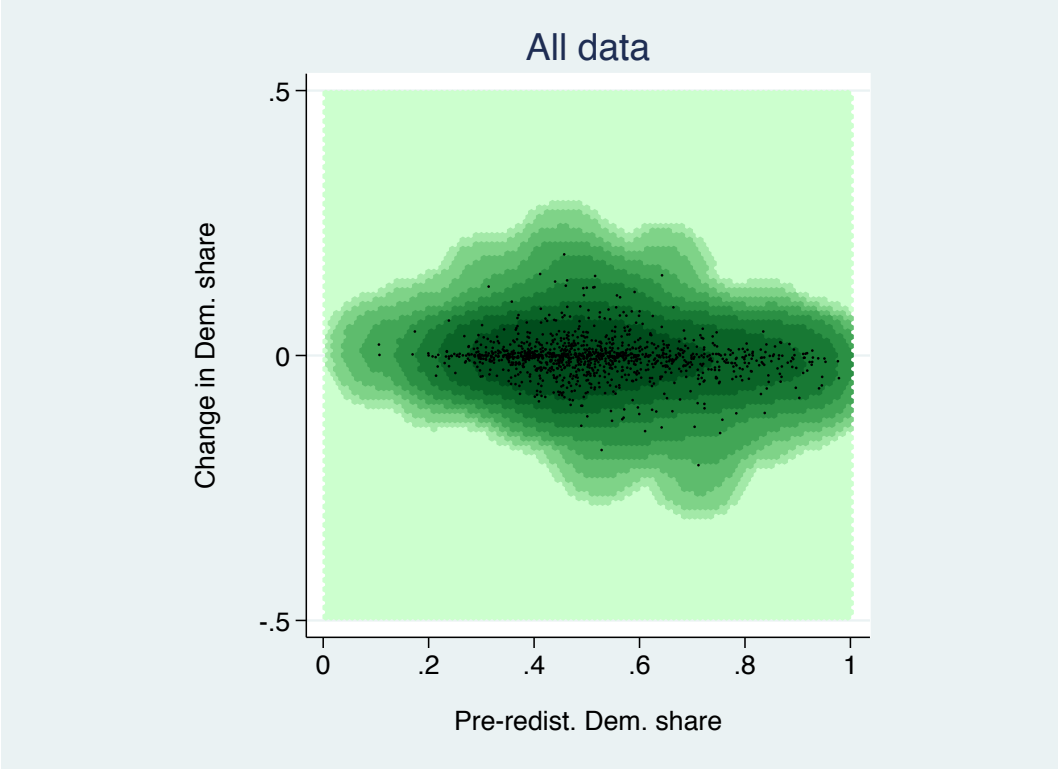
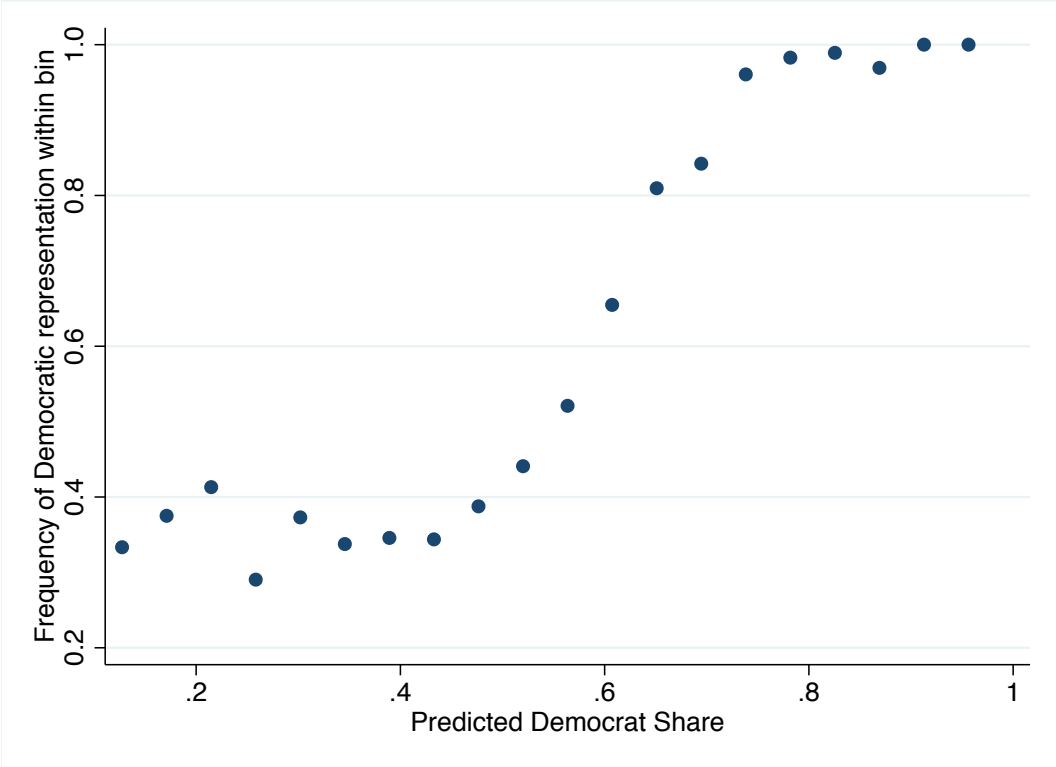


Figure 3: Binned scatterplot: Frequency of Democrat representation as a function of predicted Democrat share



Appendix 1: Data

1. Constructing predicted Democrat share

As noted in the main text, there are four steps in constructed our *DemShare* measure. First, we use the 2010 county-level voter registration data and the 2010 county-level demographic data to estimate coefficients that we will eventually use to predict Democrat share at other geographic levels. Second, we map Census block groups into pre- and post-redistricting Congressional districts. Third, we use those mappings to aggregate Census block group demographic data up to pre- and post-redistricting Congressional district-level demographics for each wave of redistricting. Finally, we use the coefficients from the estimation in the first step to construct *predicted* Democrat share at the Congressional district level, for every Congressional district immediately before and after every wave of redistricting.

The four steps are described in detail in this Appendix.

1.1 Estimating the relationship between observed Democrat share and demographics at the county-level

Using the voter registration data, we construct *actual* Democrat share. This is measured as the number of voters registered as Democrats in a district divided by the number of voters registered as Democrats or Republicans. Recall that only a subset of states maintain statistics on the partisan affiliation of registered voters. Thus, of the 3,186 counties in our data, we observe partisan affiliation statistics and can calculate "Democrat share" in 1,362 counties.

We then estimate fractional logit models taking "Democrat share" on the left hand side, and a set of county-level demographic characteristics on the right hand side. (Fractional logit models ensure that the result predicted values of Democrat share fall within 0 and 1, which is

important when we predict Democrat share at other geographic levels and in other years.) Although there are relatively rich demographic characteristics that we can observe at the county-level, we are restricted in the demographic characteristics we can use in our estimation; because the resulting coefficients will ultimately be taken to Congressional district-level data constructed from block group-level data, we can only use variables also available in the block-group level data.

On the right hand side of the fractional logit, we include the following covariates: share of population that is: in an urban area, male, over 18, over 65, black, white, Hispanic, Asian/Pacific Islander, or Native American. We also include share of households that consist of: one male, one female, a married couple with children, a married couple without children, a single male head of household family with children, a single male head of household family without children, a single female head of household family with children, a single female head of household family without children, a non-family with a male head of household, or a non-family with a female head of household. We include covariates measuring share of houses that are: vacant, owned (renter occupied), or owned (owner occupied). Finally, we also include population density in the estimation.

Because particular demographic characteristics may have different consequences in different parts of the country, we actually estimate four fractional logits (and store four sets of coefficients), one for each of the four major Census regions (Northeast, South, Midwest, West).

After estimating the four models, we store the coefficients as they will be used to predict Democrat share of Congressional districts, where we can observe all of the same demographics listed above, but do not observe voter registration statistics.

1.2 Mapping Census block-groups into Congressional districts

Next, we use historical Congressional district geographic boundary files to map Census block groups into Congressional districts. Block groups are of interested because they are the largest geographic area defined by the Census Bureau that, with some very rare exceptions, nest into Congressional districts.

Congressional district geographic boundary files are available for every Congress. For the most part, we use the files associated with Congresses just before and after redistricting. We pair these with Census defined geography from the relevant Census period. Thus, we take – for instance – the 1990 block group boundaries and map them into the 102nd Congress district boundaries to obtain the pre-redistricting mapping for the 1990's. We then use the same 1990 block group boundaries, but pair them with the 103rd Congress district boundaries to obtain *post*-redistricting mappings for the 1990's. We repeat the process for 2000's redistricting and 2010's redistricting. Table 1 in the main text summarizes the relevant periods of Congress paired with each decade's Census geography.

Thus, this process ultimately assigns Census block groups to Congressional districts for every redistricting wave, with different assignments before and after redistricting within each wave.

1.3 Matching pre- and post-redistricting districts (accounting for name changes)

Of course, it is often the case that the post-redistricting district that most resembles a given pre-redistricting district bears a different name. This is especially common when a state gains or loses seats. So that district fixed effects are meaningful, we must pair pre- and post- districts based on something other than name. To do so, we first use our block-group level population data to identify the overlap between pre- and post-redistricting districts. Suppose we are focusing on the

1990's wave of redistricting (though the procedure is identical for all three decades.) Overlap between some pre-redistricting district X_{pre} and some post-redistricting districting Y_{post} is defined as:

$$Overlap(X_{pre}, Y_{post}) = \frac{Sum\ of\ 1990\ Census\ pop.\ in\ block\ groups\ assigned\ to\ X_{pre}\ AND\ Y_{post}}{Sum\ of\ 1990\ Census\ pop.\ in\ block\ groups\ assigned\ to\ X_{pre}\ OR\ Y_{post}}$$

Thus, if X_{pre} contains exactly the same set of block groups as Y_{post}, then Overlap=1. If none of the block groups in Y_{post} are contained in X_{pre}, then Overlap=0.

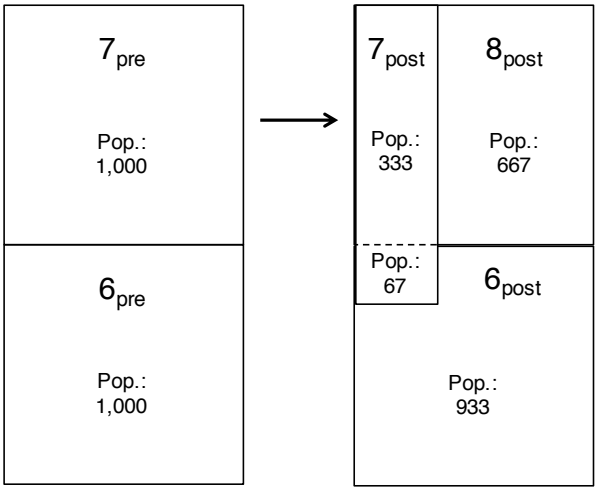
After calculating Overlap for every pair of pre- and post-districts, we identify unique pairing of districts such that:

$$Overlap(X_{pre}, Y_{post}) = max_i[Overlap(i_{pre}, Y_{post})] = max_j[Overlap(X_{pre}, j_{post})]$$

for all pre-redistricting districts i_{pre} and for all post-redistricting districts j_{post}. That is, in order for us to consider X_{pre} and Y_{post} the "same" district in our empirical analysis, it must be that (1) Y_{post} shares more population with X_{pre} than with any other pre-redistricting district *and* (2) X_{pre} shares more population with Y_{post} than any other post-redistricting district. This mutual matching scheme allows us to find a *unique* pairing of pre- and post- districts; it does, however, imply that not every district can be matched. This is not a surprise: we would not expect *every* post-redistricting district to be cleanly matched to some pre-redistricting district (or vice versa), nor would we want to pair districts that are not cleanly matched. Unmatched districts are therefore not used in our empirical analysis.

After finding a unique pairing, we rename districts so that paired districts have the same name (in our data) before and after redistricting. That is, if X_{pre} and Y_{post}, we name both of them X_{match} in our data. Thus, when we use district fixed effects in our empirical analysis, we our matched identifier for districts (e.g., "X_{match}").

To make sure this is clear, consider the hypothetical redistricting in the figure below. There, what was District 7 is split into Districts 7 and 8. The new District 8 is entirely contained within the old District 7. The new District 7 is mostly contained within the old District 7 but includes a small portion of land that was previously part of District 6.



Because Census block-groups (with very rare exceptions) nest into both pre- and post-redistricting districts, we can identify the population of each of these areas (including the subsets of districts that were transferred from other districts). For instance, we can identify that the Census block groups in the new District 7 that were previously part of District 6 have a total population of 67 people.

To pair districts, we first calculate overlap for every possible pairing:

$$\text{Overlap}(6_{pre}, 6_{post}) = 933/(1{,}000) = 0.93$$

$$\text{Overlap}(6_{pre}, 7_{post}) = 67/(1{,}333) = 0.05$$

$$\text{Overlap}(7_{pre}, 7_{post}) = 333/(1{,}067) = 0.31$$

$$\text{Overlap}(7_{pre}, 8_{post}) = 667/(1{,}000) = 0.67$$

Based on these calculations, we pair districts 6_{pre} and 6_{post}, and call them 6_{match} in our data. We pair 7_{pre} and 8_{post} and call them 7_{match}. District 7_{post} is unmatched and is not used in our analysis. This

illustrates the importance of mutual matching: of all possible pre-districts, District 7_{post} has the most overlap with 7_{pre} (Overlap(7_{pre}, 7_{post}) > Overlap(6_{pre}, 7_{post})). Thus, if not for the mutual matching condition, we might pair those two districts. However, of all possible pre-districts, District 8_{post} *also* has overlaps most with 7_{pre} (Overlap(7_{pre}, 8_{post}) > Overlap(6_{pre}, 8_{post})); so, without mutual matching (and some convention about whether to adopt pre-to-post matching or post-to-pre matching), District 7_{pre} could potentially be matched to two post- districts. Mutual matching prevents this, and identifies the unique best pair; in our example, 8_{post} clearly has more in common with 7_{pre} than 7_{post} does, so our matching scheme has selected the "correct" pairing.

1.4 Aggregating block-level demographics and constructing predicted Democrat share

The third and fourth steps of the construction of predicted Democrat share require less explanation. In the third step, using the mappings between block groups and Congressional districts, we aggregate block-level demographics to the relevant Congressional district-level for each pre- and post-redistricting period. In the final step, we simply use the result district-level demographics and the coefficients from step one to construct predicted Democrat share. Because the coefficients were drawn from a fractional logit, Democrat share is constructed by calculating:

$$DemShare_{dt} = \left[1 + e^{-\sum_i \bar{\beta}_i x_{idt}}\right]^{-1}$$

where *d* indexes the district (named according to our matching scheme discussed above), *t* indexes the time period (e.g., 1990's/pre-redistricting, 1990's/post-redistricting, 2000's/pre-redistricting, etc.), and *i* indexes the set of demographic variables used to predict Democrat share. x_{idt} is the value of particular demographic variable for district *d* at time *t*, and $\bar{\beta}_i$ is the relevant coefficient from the 2010 county-level fractional logit estimation.

Appendix 2: Additional tables and figures

In this appendix, we generate two-dimensional kernel density estimate plots (similar to Figure 2 in the main text). To assess whether the pattern of changes in partisan composition varies by circumstances surrounding redistricting, we split the sample by:

- (Figures A.1(a) & (b)) states that must redistrict because they gained/lost seats vs. those that did not gain/lose seats,
- (Figures A.2 (a), (b), and (c)) states where the redistricting authority (e.g., state legislature) is dominated by Democrats vs. dominated by Republicans.
- (Figures A.3(a) & (b)) districts with Democrat incumbents vs. Republican incumbents,
- (Figures A.4(a) & (b)) districts where DW-Nominate scores of elected representatives has trended left in the preceding decade vs. trended right
- (Figures A.5(a), (b), and (c)) redistricting that occurred in the 1990s, 2000s, or 2010s, and
- (Figures A.6(a), (b), and (c)) state-decade observations wherein redistricting was done by a redistricting commission, the courts, or a combination of the state legislature and governor.

In all cases, we test pairwise comparisons between slopes of lines fitted to the data to assess whether patterns of redistricting vary depending on the nature of the redistricting. Results are reported in Appendix Table 1. None of these tests reveal significant differences in slopes within categories.

Figure A.1(a): States that gained/lost seats

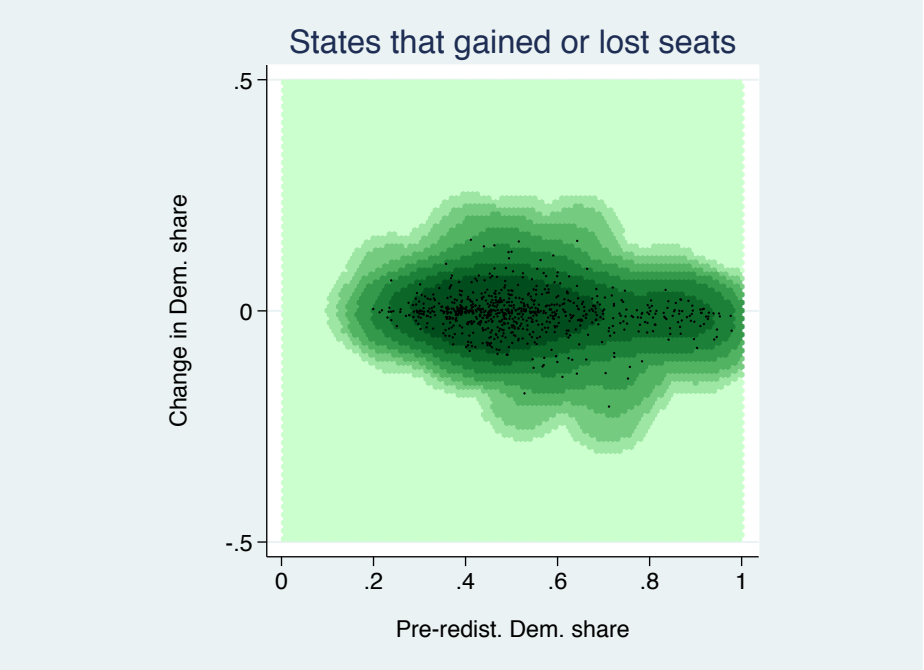


Figure A.1(b): States that did not gain/lose seats

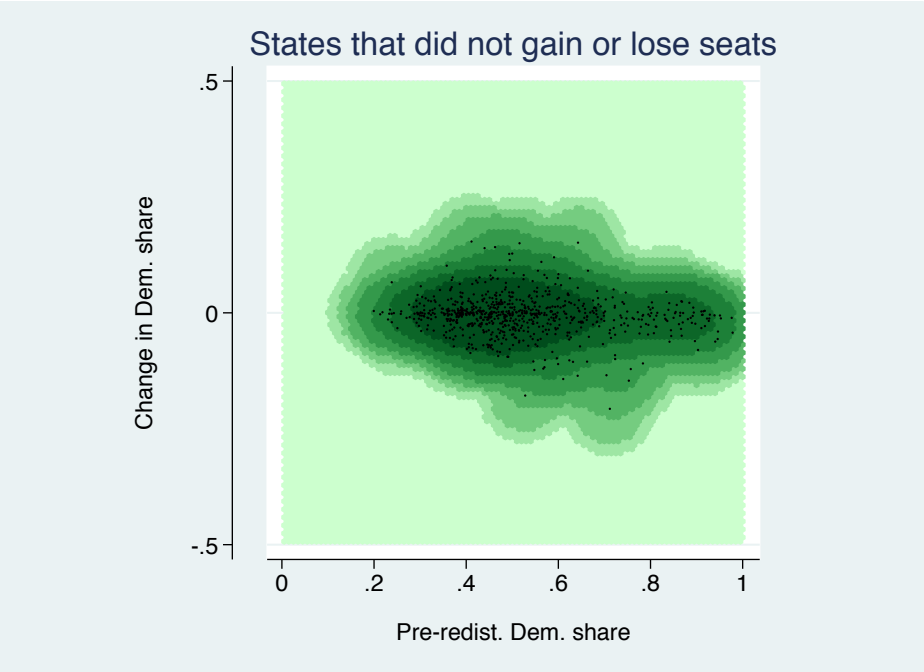


Figure A.2(a): States where redistricting authority is controlled by Democrats

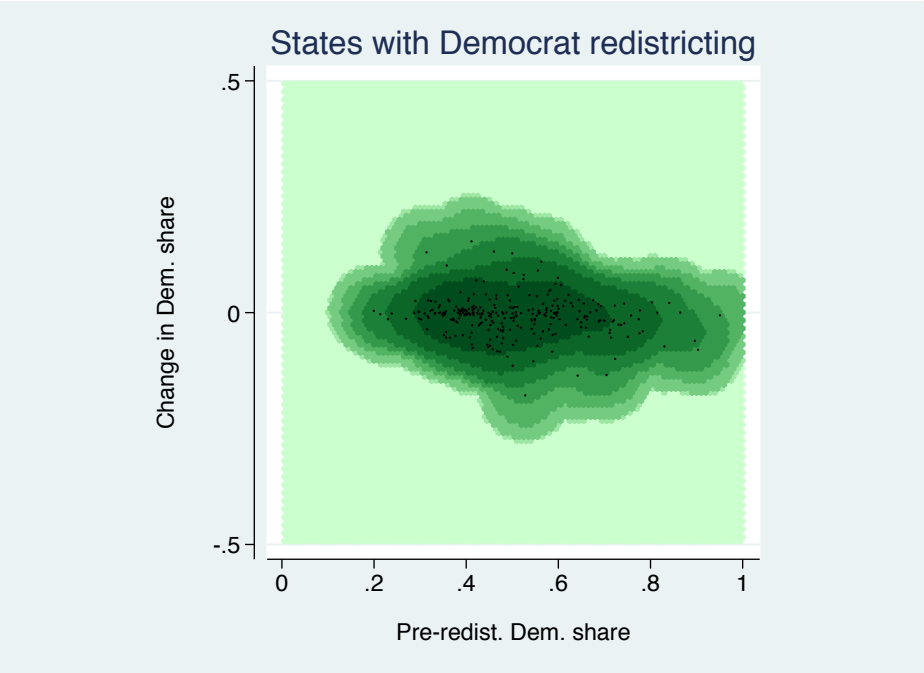


Figure A.2(b): States where redistricting authority is controlled by Republicans

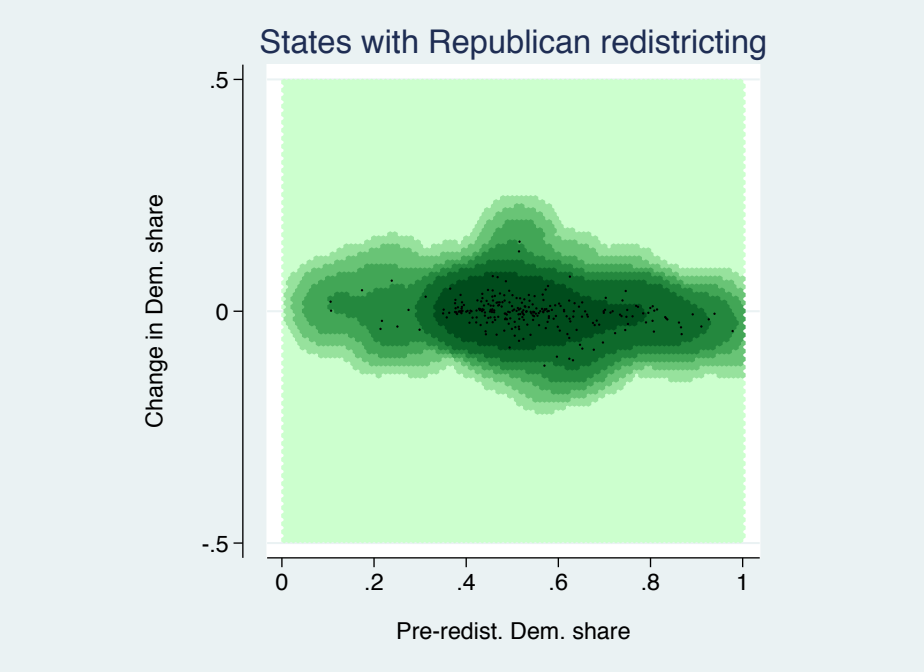


Figure A.3(a): Districts where the pre-redistricting incumbent was a Democrat

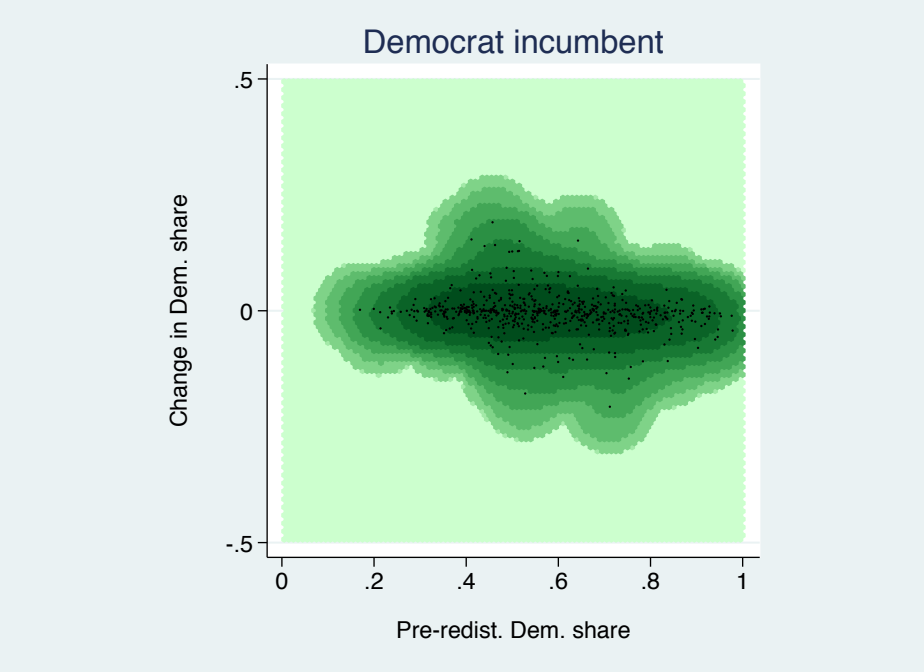


Figure A.3(b): Districts where the pre-redistricting incumbent was a Republican

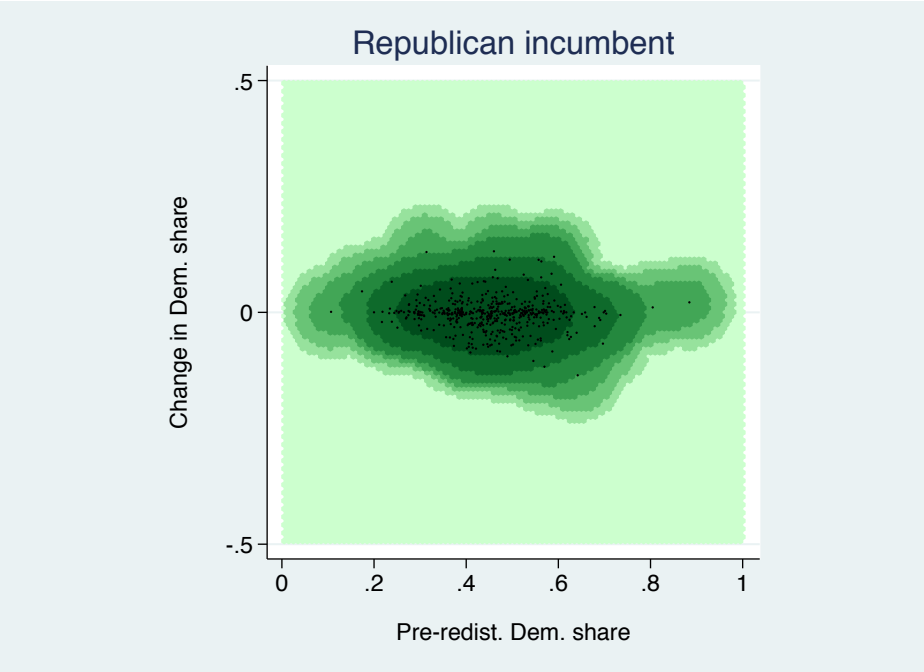


Figure A.4(a): Districts where DW-Nom. scores in pre-redistricting decade were trending left

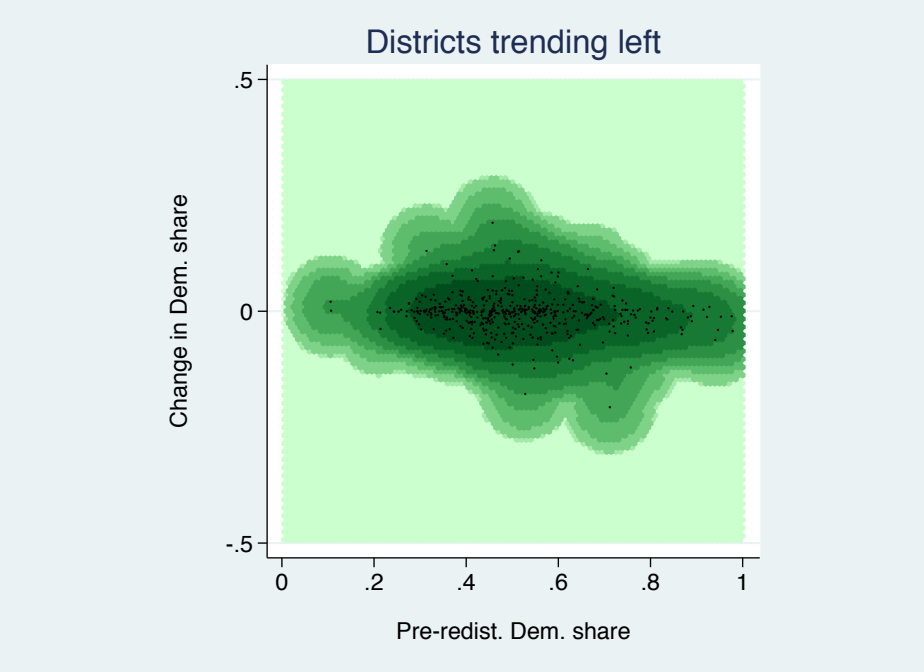


Figure A.4(b): Districts where DW-Nom. scores in pre-redistricting decade were trending right

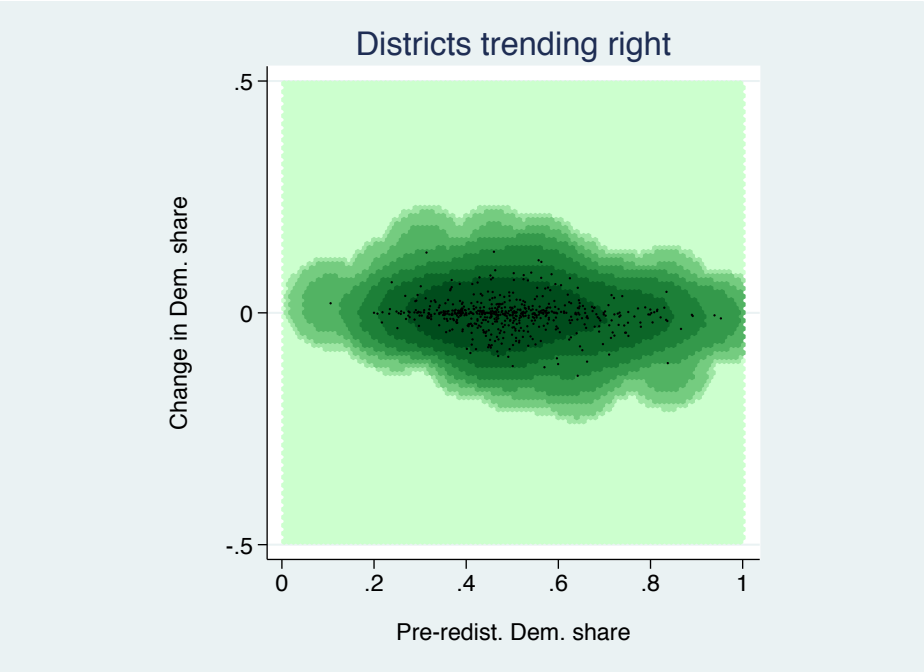


Figure A.5(a): 1990's post-Census redistricting

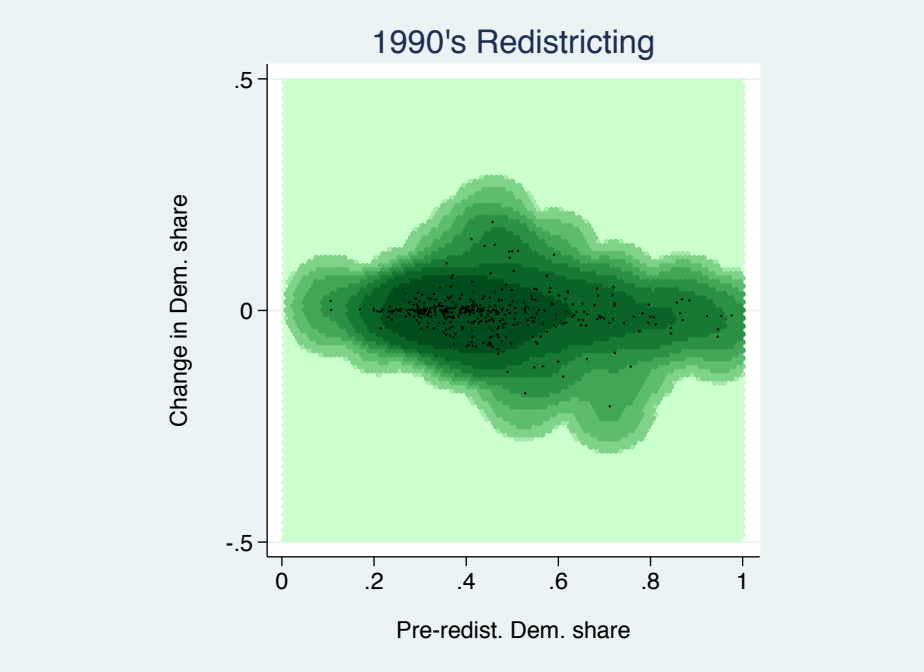


Figure A.5(b): 2000's post-Census redistricting

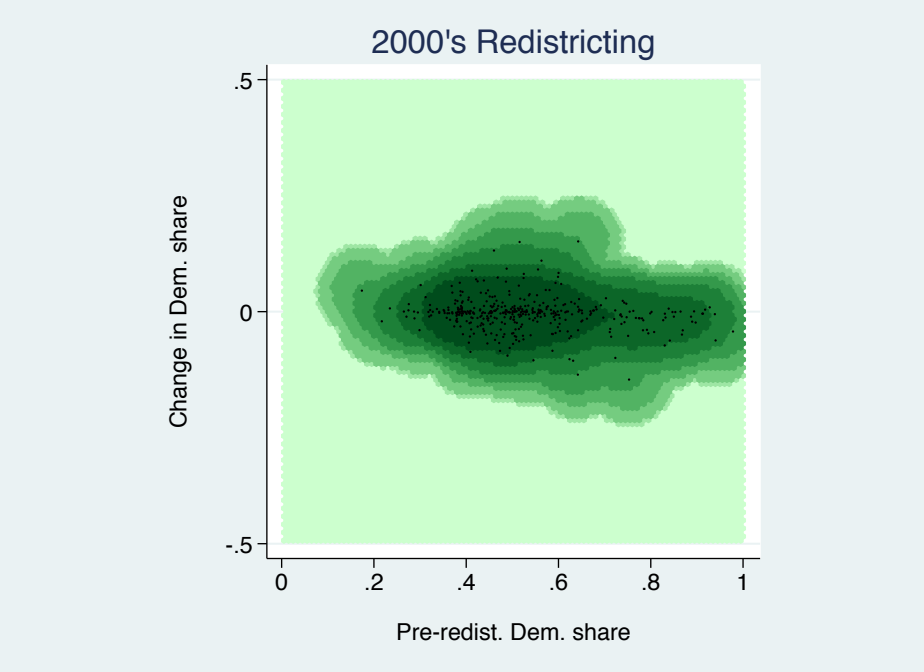


Figure A.5(c): 2010's post-Census redistricting

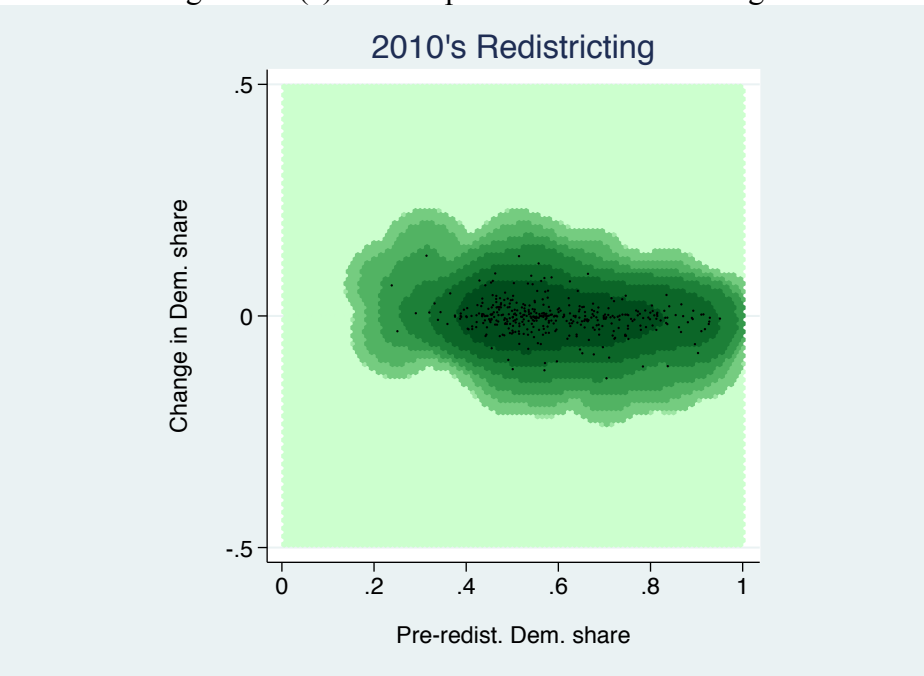


Figure A.6(a): Redistricting authority: Redistricting commission

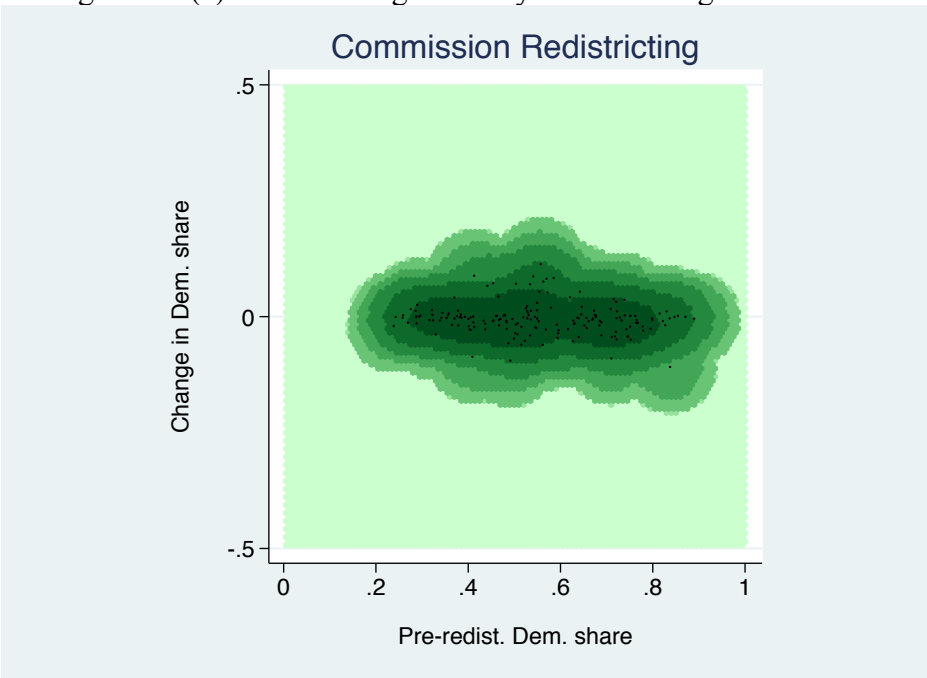


Figure A.6(b): Redistricting authority: Court

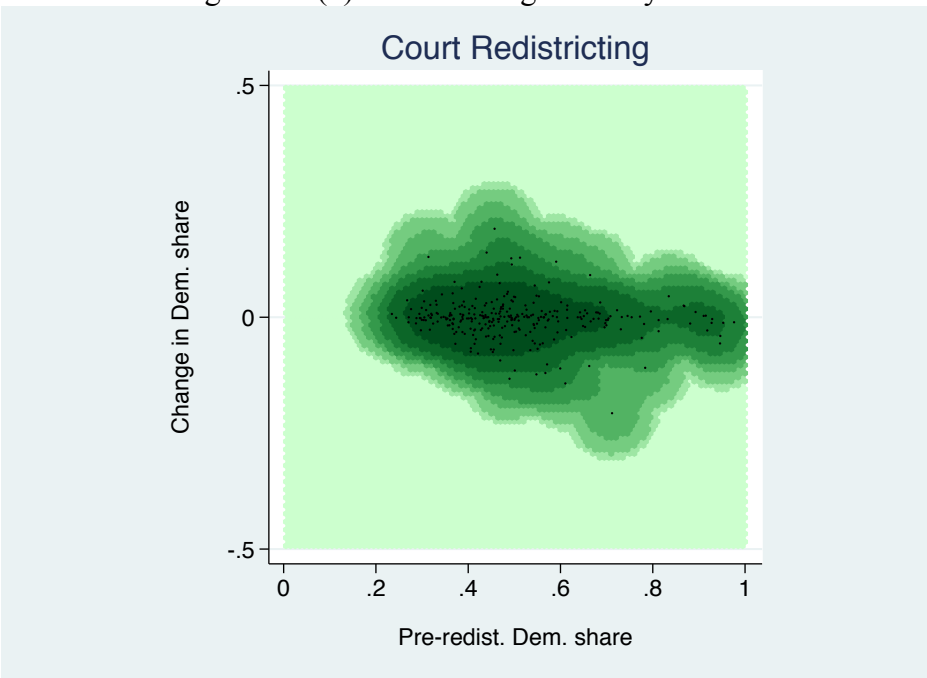
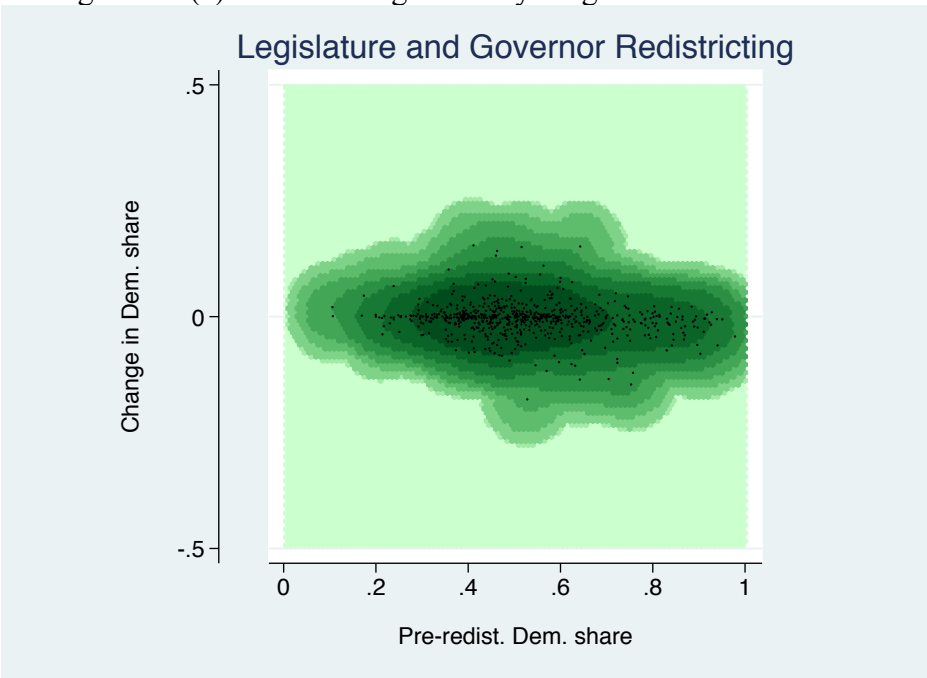


Figure A.6(c): Redistricting authority: Legislature and Governor



Appendix Table 1: Relationship between pre-redistricting predicted Democrat share and change in Democrat share in subsamples of the data

	Slope	Test of diff.	P-value
1. Redistricting cause			
(a) Forced	-0.035***		
	(0.006)		
(b) Not forced	-0.031***		
	(0.009)	(a) vs. (b)	0.739
2. Party of redistricters			
(a) Dem. redistricters	-0.044***		
	(0.012)		
(b) Repub redistricters	-0.050***		
	(0.012)	(a) vs. (b)	0.760
3. Party of pre-redist. incumb.			
(a) Repub. incumb.	-0.025**		
	(0.012)		
(b) Dem. incumb.	-0.041***		
	(0.007)	(a) vs. (b)	0.234
4. Pre-redist. DW-Nom. trend			
(a) DW-Nom trending left	-0.035***		
	(0.012)		
(b) DW-Nom trending right	-0.039***		
	(0.013)	(a) vs. (b)	0.820
5. Redistricting authority in state			
(a) Redist. by courts	-0.042***		
	(0.011)	(a) vs. (b)	0.414
(b) Redist. by commission	-0.028**		
	(0.013)	(a) vs. (c)	0.525
(c) Redist. by leg./gov.	-0.034***		
	(0.007)	(b) vs. (c)	0.701
6. Redistricting wave			
(a) 1990's	-0.045***		
	(0.010)	(a) vs. (b)	0.536
(b) 2000's	-0.036***		
	(0.010)	(a) vs. (c)	0.738
(c) 2010's	-0.049***		
	(0.009)	(b) vs. (c)	0.338

Table notes: This table reports slopes of lines fit through Appendix Figures A.1-A.6. Specifically, we run a simple regression of "change in Democrat share" on "pred. Dem. share". The column "slope" reports the coefficient on "pred. Dem. share". Each panel represents a distinct way of splitting the data (by cause of redistricting, party of redistricters, etc.), corresponding with the six sets of figures in this appendix. We conduct pairwise comparisons of slopes within each sample splitting category; p-values from these tests are in the final column of the table.

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1