Assessing Robustness of Findings About Racial Redistricting’s Effect on Southern House Delegations

Abstract: We assess whether racial redistricting increases the number of Southern representatives to the left of the US House median. Our results, which are based on Monte Carlo simulations and an alternative measure of representatives’ preferences, are generally null findings. The data do not support the claim that racial redistricting promotes liberal policy outcomes; nor do they support the claim that it promotes conservative policy outcomes. At a methodological level, we suggest techniques that researchers can use to assess how robust their findings are to noise in variables based on estimated values like DW-Nominate scores.

1 Introduction

Since the 1990s, political scientists have debated how majority-minority districts affect representation and policymaking. The conventional wisdom is that racial redistricting has perverse effects: minority voters are packed into a few districts, other districts in a state are made more conservative, and the conservative legislators from those districts enact policies that are contrary to the preferences of most African-American and Latino voters.

Elections for the US House of Representatives in the 1990s were widely interpreted as evidence supporting the perverse effects claim. After the 1990 census, the federal government mandated a substantial increase in the number of majority-minority districts. In 1992, a record number of African-American and Latino House members were elected, and in 1994 the Republican party won control of the House in a landslide election.
Of course, major rightward national partisan tides also occurred in 1994, and Republicans did well in the Rocky Mountain West and the Senate, neither of which was affected by racial redistricting. However, social scientists developed methods to project the effect of different districting plans, and generally came to the conclusion that racial redistricting was bad for minority voters’ policy interests. The typical approach (Grofman, Griffin and Glazer 1992; Cameron, Epstein and O’Halloran 1996; Lublin 1997) was to use election results and roll-call voting data to project the type of the representative who would be elected from a district with a particular demographic composition. Using the projections, scholars simulated the effect of different districting plans and concluded that majority-minority districting did not serve the policy interests of liberal minority voters.

In a series of papers, Shotts (2001, 2002, 2003a) took issue with the reasoning behind the perverse effects claim. He argued that because there are many different ways of drawing districts, the crucial question is how a strategic gerrymanderer would craft districts when unconstrained and when subject to a federal mandate to create majority-minority districts. Shotts developed theoretical models which predicted that federal mandates would force conservative redistricters to elect liberal minority representatives, something they would not otherwise do, especially in states where a majority of the voters were conservative. As long as liberal redistricters in other states had enough flexibility to draw majority-minority districts that did not waste the votes of liberal white voters, the models predicted that the net effect of racial redistricting would be to promote liberal policy outcomes.

To test this theoretical intuition, Shotts (2003a), analyzed DW-Nominate scores (Poole and Rosenthal 2001) for southern House members elected between 1986 and 1996. He found that the number of southern members to the left of the US House median increased after racial redistricting in the 1990s. A simple regression analysis that controlled for other factors that could affect a state’s propensity to elect liberals similarly found a relationship between racial redistricting and the election of liberals to the House. Thus, Shotts concluded that within the canonical median legislator model of policy choice, racial redistricting promotes liberal policy outcomes.¹

¹ Subsequently, Washington (2012) conducted a difference-in-differences analysis of the average liberalism of elected representatives. In most of her specifications, an increase in the number of majority-minority districts was associated with increased liberalism of a state’s delegation. This finding is implicitly grounded in a different theory of policymaking, namely a mean legislator model, but it is broadly consistent with Shotts (2003a).
This finding was surprising, because the conventional wisdom was that Southern House elections in the 1990s were the prototypical example of the perverse effects of racial redistricting. However, Shotts’s analysis left several questions unanswered, due to the limited set of data available. For example, although his theoretical models focused on differences in how majority-minority mandates affect Republican and Democratic gerrymanderers, his empirical analysis did not separate out these effects; given that there was only one Southern redistricting plan crafted by Republicans, it would have been impossible to do so.

Now, with over a decade of additional data, including several southern redistrictings controlled by Republicans, it is possible to estimate separately how mandates affect Republican and Democratic gerrymanderers. This is what Simons and Mallinson (2015) do. Using more nuanced versions of the empirical specifications in Shotts (2003a), they find null results for essentially every coefficient of interest, including the original one that Shotts focused on in his analysis.

Simons and Mallinson also replicate Shotts’s original analysis of House members elected between 1986 and 1996, and obtain a null result. As explained in their paper, they use Shotts’s data for every variable except the DW-Nominate scores that are used to calculate the key dependent variable, Left of Median, i.e. the proportion of a state’s representatives to the left of the House median in a Congress. The reason for the discrepancy between Simons and Mallinson’s findings and Shotts’s findings is that DW-Nominate scores are continually re-estimated. Not only does a representative’s score change from Congress to Congress, his or her score for a particular Congress was different in 2012 (when Simons and Mallinson downloaded DW-Nominate scores) than it was in 2000 (when Shotts downloaded DW-Nominate scores). The dependent variable Left of Median was sensitive to small movements in the DW-Nominate of individual representatives.

Although this paper is a reply to Simons and Mallinson (2015), we want to be clear that our goal is not to dispute their analysis or results. We assume that the more recent DW-Nominate estimates are better than the older ones, because they are based on additional information about the voting behavior of members of Congress. Although at the time that the Shotts (2003a) article was published, the data appeared to provide compelling evidence that majority-minority mandates lead to liberal policy outcomes, the evidence in support of that claim should no longer be considered to be compelling. However, it is also important to note that this does not mean the data provide compelling evidence in support of the opposite claim embodied in the conventional wisdom about perverse policy effects. Rather, we will show that the data provide no strong evidence in support of either claim.

The fact that Shotts’s results disapparated as DW-Nominate scores were re-estimated and updated suggests that applied researchers should be cautious when using estimated dependent variables based on DW-Nominate scores or other
noisy inputs. Our primary goal in writing this paper is methodological: to identify, with the benefit of hindsight, things that an applied researcher like Shotts could have done to take into account the ways that measurement error could pose a threat to his results. Our hope is that this analysis will be useful for other scholars, and as a preliminary step in that direction, we apply some of our techniques to Simons and Mallinson’s analysis.

In the next section, we focus on Shotts (2003a). First, we look at the raw DW-Nominates to get a rough sense of whether Shotts’s results appear to be fragile, depending on the precise values of the scores for southern representatives. Second, we re-estimate Shotts’s regression model, using data from Monte Carlo simulations that allow for errors in DW-Nominates. Third, we re-do Shotts’s analysis, constructing his Left of Median dependent variable using a completely different measure of each House member’s ideology, from Bonica (2014).

In the subsequent section, we turn our attention to Simons and Mallinson’s specification, and we use our Monte Carlo technique to assess whether their results are robust to measurement error in DW-Nominate scores. We find that their null results are robustly null.

Before getting into the details of our analysis, we wish to make one other methodological point. Although our analysis is not experimental, we pre-registered our analysis with the Experiments in Governance and Politics network (EGAP). This document is included as an Appendix. We believe pre-registration is particularly useful for replies, rejoinders, and other papers that focus on results that are contested and controversial. In such situations, scholars whose work is under scrutiny have especially strong incentives to hunt for whatever results best comport with their previous arguments. In the context of our analysis, the fact that we pre-registered our research design is meaningful. Some of our results (namely the analysis using a dependent variable based on Bonica’s estimates of legislators’ ideology) could be used to buttress Shotts’s (2003a) claim that racial redistricting promotes liberal policy outcomes. However, the rest of our results do not support that claim. Thus, following our research design, our overall conclusion is that the results from Shotts (2003a) were not robust to the possibility of measurement error in his dependent variable. In terms of the crucial applied question about whether majority-minority districting leads to perverse effects, the evidence is inconclusive.

2 Re-Analysis of Shotts (2003a)

The data that we use in this section come from three sources. Old Nominates are DW-Nominate scores downloaded from Keith Poole’s website on February 16,
2000, and used in Shotts (2003a). *Current Nominates* are DW-Nominate scores downloaded from Keith Poole’s website on July 16, 2015. All other variables come from the data set used in Shotts (2003a).

### 2.1 Raw DW-Nominates

We begin by plotting the *Old Nominates* for southern representatives and comparing them to the House median. This procedure, which was not done by Shotts (2003a), could potentially raise a red flag about his analysis. For example, consider the following scenario. Suppose several southern representatives are estimated by the DW-Nominate procedure to be just to the right of the median before redistricting and several of them are estimated to be just to the left of the median after redistricting. Such a scenario would suggest that Shott's findings were driven by luck; namely, the estimates of those representatives' ideal points, which easily could have been either to the left or right of the median, happened to be to the left pre-redistricting and to the right post-redistricting, thereby leading Shotts to mistakenly conclude that racial redistricting increased the number of liberals elected.²

To assess this possibility, we focused on the 102nd Congress (the last one before redistricting) and the 103rd and 104th Congresses. The reason we use two post-redistricting Congresses is that many scholars believe that perverse effects of the 1990 redistricting arrived in the 104th Congress. Note that the relevant median is estimated using *Old Nominates* and is different for each Congress; as argued by Shotts (2003a), this is actually useful because it makes it possible to control for national partisan tides.

As can be seen from visual inspection of Figure 1, there is no evidence supporting either part of the red flag scenario. There is no clump of southern representatives with DW-Nominate just to the right of the median in the 102nd Congress, and neither the 103rd nor 104th Congress features a clump just to the left of the House median. In fact, within 0.05 units of each median, the southern representatives are roughly balanced on either side before and after redistricting. This is important given that Carroll et al. (2009) estimate the average first-dimension standard error for DW-Nominate estimates to be 0.047.

We also applied this technique to *Current Nominates*, to check whether the lack of statistically significant results in Simons and Mallinson’s analysis might

---

² Along these lines, Lublin and Voss (2003) criticized Shotts by noting that there were some southern House members just to the right of the House median before redistricting. However, they failed to note that there were also some Southern representatives just to the left of the median.
be driven by a group of southern representatives whose DW-Nominates are just to the left of the median before redistricting or just to the right of the median after

Figure 1: Southern representatives’ DW-Nominates, compared to House Median. This figure displays DW-Nominate scores for southern representatives, as well as the House median for the 102nd, 103rd, and 104th Congresses. The top panel uses Old Nominates, as used in Shotts (2003a), whereas the bottom panel uses Current Nominates.
redistricting. However, as shown in Figure 1, the data provide no evidence sup-
porting either of these scenarios.

Our conclusion from this part of the analysis is that there is no obvious red
flag in the DW-Nominates.

2.2 Monte Carlo Reestimation

The second part of our analysis is to re-estimate Shotts’s model, using data pro-
duced by Monte Carlo simulations to allow for errors in DW-Nominates. This
technique enables us to check whether Shotts’s findings are robust to reasonable
levels of error. As in the original article, the specification is:

\[
\text{Left of Median}_{st} = \alpha + \beta_1 \text{Left of Median}_{st-1} + \beta_2 \text{Fraction Majority-Minority}_{st} \\
+ \beta_3 \text{Unified Democratic Control}_{st} + \beta_4 \text{Voter Liberalism}_{st} \\
+ \beta_5 \text{Percentage Black}_{st} + \beta_6 \text{Percentage Latino}_{st} + \epsilon_{st},
\]

The dependent variable is the proportion of districts in state \( s \) in Congress \( t \) that
elected a representative to the left of the House median. The sample consists of
the 10 southern states that had majority-minority districts for years ranging from
1988 to 1996.

For our first set of Monte Carlo simulations, we allow for different levels of
error in the DW-Nominate scores that were used as the basis for \( \text{Left of Median} \).
All variables except the dependent variable and its lag are exactly as in Shotts
(2003a). To construct \( \text{Left of Median}_{s,t} \), each iteration of the simulation draws for
each House member in each Congress an ideal point from a normal distribution,
with mean at his/her DW-Nominate score for that Congress.\(^3\) These draws are then
used to calculate the House median for Congress \( t \) as well as the proportion of
representatives in state \( s \) who were to the left of the median in that Congress.

We allow the standard deviation of the draws to vary from 0 to 0.2, in 0.01 unit
increments. For a sense of this scale, in the 102nd Congress, 0.02 is roughly the
distance in \( \text{Old Nominate} \) between the median members, James Chapman (D-TX)
and Michael Andrews (D-TX), both of whom were at \(-0.108\), and Stephen Neal
(D-NC) who was at \(-0.131\), or Martin Lancaster (D-NC) and Lewis Payne (D-VA),
both at \(-0.09\). The upper end of our scale is the distance between the median and
Ronald Coleman (D-TX, at \(-0.313\)) or the distance between the median and the

\(^3\) This method is far from perfect, because errors in DW-Nominate are presumably are corre-
bated both across time and across members. However, to the best of our knowledge, no technique has
been developed that would deal with this issue in DW-Nominate error. Hence, we treat the errors
as being independent.
most conservative Democrats in the House, like Charles Stenholm (D-TX, at 0.08) or Gene Taylor (D-MS at 0.10). Carroll et al. (2009, pp. 270–271) estimate a standard error for DW-Nominates that is around 0.05, so we believe that is the most relevant region of standard deviations for our Monte Carlos. However, we set up the simulations to include a broader range of values as well.

We performed 1000 iterations for each standard deviation between 0 and 0.2. The product of this analysis is Figures 2–4, with the horizontal axis showing the standard deviation for the Monte Carlos.

The vertical axis of Figure 2 is the proportion of iterations that resulted in a statistically-significant (at the conventional 0.05 level, using two-tailed tests) positive coefficient on \( \text{Fraction Majority-Minority} \). With no error, the proportion is 1.0, which simply means that estimating the exact same regression 1000 times with no variation in the dependent variable always yields the same output. As the error increases from zero, the proportion of coefficients that are positive and statistically significant declines rapidly, down to around 10%. This means that to believe that the data provide compelling support for Shotts’s original finding, one would have to believe a very implausible assumption, namely that there is almost no measurement error in DW-Nominates.

Figure 2: Significant positive coefficient on \( \text{Fraction Majority-Minority} \).
This figure displays the proportion of 1000 Monte Carlo simulations that yielded statistically significant positive coefficients on the \( \text{Fraction Majority-Minority} \) variable, using different levels of error in DW-Nominates.
Assessing the Effect of Racial Redistricting

Figure 3: Negative coefficient on Fraction Majority-Minority.
This figure displays the proportion of 1000 Monte Carlo simulations that yielded negative coefficients on the Fraction Majority-Minority variable, using different levels of error in DW-Nominates. The figure also shows the proportion of coefficients that were negative and significant (which was zero for all error levels).

Figure 4: Average coefficient and confidence interval.
This figure shows the average coefficient on Fraction Majority-Minority from the Monte Carlo simulations, using different levels of error in DW-Nominates.
Figure 3 provides a similar plot of the proportion of coefficients that are negative, as well as the proportion that are negative and statistically significant. The proportion that are negative is always very low (less than 10%), and even for the highest amounts of error none of the iterations produced coefficients that were negative and statistically significant. Thus, although the data provide little support for Shotts’s claim, they provide even less support for the conventional wisdom about perverse effects of racial redistricting.

Figure 4 plots, for each level of the assumed standard deviation for DW-Nominate scores, the mean and a 95% confidence interval from our iteration estimates. For all standard deviations the mean coefficient estimate is positive, but it declines as the Monte Carlo incorporates higher levels of error in DW-Nominates. For very low levels of the standard deviation, the confidence interval excludes zero, but it crosses zero for a standard deviation somewhere between 0.01 and 0.02. This is a very small level of error, especially compared to the standard error estimate from Carroll et al. (2009). Thus, based on this analysis, as well as Figure 2, we conclude that Shotts’s original finding was not robust to the possibility of reasonable measurement error in DW-Nominate scores. However, as can be seen from Figures 3 and 4, the analysis provides no support for the perverse effects claim.

We also did a second set of Monte Carlo simulations, using Carroll et al.’s (2009) estimated standard errors for individual members. The procedure was very similar to our first set of Monte Carlos. However, there is only one level of standard deviation in the DW-Nominate for each member. Unlike our first set of simulations, this standard deviation varies across members. One other difference to note is that our second set of Monte Carlos use Current Nominates because those are the ones for which estimated standard errors are available. We conducted 1000 iterations of the simulated regression. As shown in Table 1, in 2.7% of these iterations, the coefficient on Fraction Majority-Minority was positive and statistically significant. It was never negative and statistically significant (and in fact it never was negative in any of the iterations). The Monte Carlo based coefficient is 0.27, which is substantially smaller than Shotts’s original coefficient (0.47), and the coefficient is clearly statistically insignificant, with a standard error of 0.20.

As with our first Monte Carlo analysis, this finding is more consistent with Shotts’s argument than with the perverse effects claim. Nonetheless, it is clearly a null result, and we again conclude that Shotts’s finding was not robust to the possibility of measurement error in DW-Nominates. It is also important to note that our results strengthen Simons and Mallinson’s null findings. It is not simply the case that Shotts’s findings hold with Old Nominates and fail to hold with Current Nominates. Rather, even doing an analysis based on Old Nominates, Shotts’s results were predicated on an implicit assumption of a relatively-low level of
measurement error in DW-Nominates. Now that we know this, it is not surprising that the results failed to hold as the scores were re-estimated and updated.

### 2.3 Alternative Measure of Left of Median

The final way we reassess Shotts (2003a) is by using a different measure of legislators’ ideology as a basis for the variable Left of Median. Although DW-Nominates are the most commonly used measure of legislators’ ideal points, there are others: prior versions of Nominate scores like D-Nominate (Poole and Rosenthal 1997), ratings by general interest groups like Americans for Democratic Action (ADA) and the American Conservative Union (ACU), ratings by organizations that focus on issues of particular importance to minority voters like the Congressional Black Caucus (CBC) and the Leadership Council on Civil and Human Rights (LCCR), Bayesian Markov Chain Monte Carlo estimates (Clinton, Jackman and Rivers 2004), and measures based on campaign contributions to members of Congress (Bonica 2014).

Each of these types of measures has advantages and disadvantages. Most notably, despite the fact that almost all applied uses of ideal point estimates implicitly attribute cardinal meaning to them, there are strong reasons to suspect that they are not cardinal, even within a given Congress, because they stretch and shrink different portions of the ideological spectrum (Krehbiel and Peskowitz 2015). Shotts’s (2003a) analysis only relied on a much weaker assumption, i.e. that DW-Nominates are ordinally correct within a Congress. Nevertheless, it

---

Table 1: Monte Carlo results using Carroll et al.’s (2009) estimated standard errors.

<table>
<thead>
<tr>
<th></th>
<th>Old Nominates</th>
<th>Current Nominates</th>
<th>MC Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>S.E.</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Fraction Majority-Minority</td>
<td>0.4737</td>
<td>0.2092</td>
<td>0.2733</td>
</tr>
<tr>
<td>Unified democratic control</td>
<td>–0.18</td>
<td>0.446</td>
<td>–0.1191</td>
</tr>
<tr>
<td>Voter Liberalism</td>
<td>–0.0043</td>
<td>0.0039</td>
<td>–0.0033</td>
</tr>
<tr>
<td>% Black</td>
<td>–0.0029</td>
<td>0.0031</td>
<td>–0.0014</td>
</tr>
<tr>
<td>% Latino</td>
<td>0.0562*</td>
<td>0.0292</td>
<td>0.053*</td>
</tr>
<tr>
<td>Lagged Left of Median</td>
<td>0.4851***</td>
<td>0.1116</td>
<td>0.5306***</td>
</tr>
<tr>
<td>Constant</td>
<td>0.2123*</td>
<td>0.1111</td>
<td>0.2014*</td>
</tr>
</tbody>
</table>

*p<0.10, **p<0.05, ***p<0.01.

%Significant coefficients from two-sided t-test at 5% level.
is possible that the scores are not even ordinally correct. Indeed, the fact that the DW-Nominate procedure produces different ordinal rankings for the same representatives in the same Congress, depending on voting patterns in subsequent Congresses, shows that there is at least some degree of ordinal measurement error in DW-Nominate.

One potential antidote to the problem of either cardinal or ordinal measurement error is to check whether a result holds when using alternative measures of legislators’ ideal points. This is what Washington (2012) does, running her difference-in-differences analysis using a wide variety of different measures. Here, we focus on one particular alternative measure: Bonica’s “CFscores.” Bonica’s methodology uses campaign contribution data to estimate the ideal points of candidates and contributors, and yields ideology scores for members of Congress that can be used for the purpose of our robustness analysis.

As laid out in our research design document, we replicated the Shotts analysis, with the only difference being that we used CFscores as the basis for the dependent variable Left of Median. There are two versions of CFscores: one in which each legislator’s location is constrained to be constant over time and one in which it can move. Our research design document did not specify which version we would use, so we present results for both versions. As can be seen in Table 2, the results provide support for Shotts’s claim that majority-minority mandates increase the number of House liberals elected from a state. The coefficient on Fraction Majority-Minority is very similar to the coefficient reported in Shotts (2003a), and is significant at the 5% or 10% level, depending on which version of the CFscores is used.

Table 2: Regression results using Bonica scores as alternate measures of ideology.

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>S.E.</th>
<th>Coefficient</th>
<th>S.E.</th>
<th>Coefficient</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Fraction Majority-Minority</td>
<td>0.4737**</td>
<td>0.2092</td>
<td>0.513*</td>
<td>0.2728</td>
<td>0.5319**</td>
<td>0.2172</td>
</tr>
<tr>
<td>Unified democratic control</td>
<td>–0.18</td>
<td>0.446</td>
<td>0.057</td>
<td>0.6059</td>
<td>–0.3954</td>
<td>0.4768</td>
</tr>
<tr>
<td>Voter Liberalism</td>
<td>–0.0043</td>
<td>0.0039</td>
<td>–0.0124**</td>
<td>0.0056</td>
<td>–0.0066</td>
<td>0.0043</td>
</tr>
<tr>
<td>% Black</td>
<td>–0.0029</td>
<td>0.0031</td>
<td>–0.0095**</td>
<td>0.0045</td>
<td>–0.0056</td>
<td>0.0034</td>
</tr>
<tr>
<td>% Latino</td>
<td>0.0562*</td>
<td>0.0292</td>
<td>0.0323</td>
<td>0.0379</td>
<td>0.0356</td>
<td>0.0302</td>
</tr>
<tr>
<td>Lagged Left of Median</td>
<td>0.4851***</td>
<td>0.1116</td>
<td>0.5276**</td>
<td>0.1144</td>
<td>0.6525***</td>
<td>0.0904</td>
</tr>
<tr>
<td>Constant</td>
<td>0.2123*</td>
<td>0.1111</td>
<td>0.4504**</td>
<td>0.1756</td>
<td>0.2163*</td>
<td>0.1293</td>
</tr>
</tbody>
</table>

*p<0.10, **p<0.05, ***p<0.01.
To be clear, however, we do not take this to be strong evidence in support of Shotts’s argument that racial redistricting promotes liberal policy outcomes, for three reasons. First, Bonica’s CFscores are only one measure – a more thorough analysis would also make use of other measures of House members’ ideologies. Second, we only did the analysis for the Congresses in Shotts’s original article, and did not extend it to more recent ones. Finally, the finding has not been subject to sensitivity testing of the sort that we conducted in the previous subsection. The finding may well fail to be robust to the possibility of measurement error in the Bonica scores, just like the original Shotts finding failed to be robust to measurement error in DW-Nominites.

3 Application to Simons and Mallinson (2015)

In this section we reestimate Simons and Mallinson’s Conditional Effects Model so as to understand the implications of measurement error in their modified context. Following our research design, we implement the Monte Carlo simulation procedure with different levels of error. The Conditional Effects Model extends the original data from Shotts (2003a), replaces Shotts’s Unified Democrat Control dummy with dummies for Unified Republican Control and Bipartisan Control, and further interacts these dummies with the Fraction Majority-Minority variable. The results for this model can be found in Table 3 (fourth column) of Simons and Mallinson (2015), which shows coefficients that are not significantly different from zero for Fraction Majority-Minority, Unified Republican Control, Bipartisan Control, and for both of their interactions.

Our robustness checks in this section follow the same Monte Carlo procedures that we used in the previous section. We first introduce different levels of error in DW-Nominites; this allows us to recalculate the House median and shares of legislators to the left of this median. We then repeatedly estimate the coefficients of the Conditional Effects Model with differing levels of error in the ideology scores. To do this, we first obtained the same set of DW-Nominate scores used by the authors (updated up to the 111th Congress) and replicated their original regression results, which we present in Table 3. When presenting the results from our procedure we focus on their five key independent variables (Fraction Majority-Minority, Unified Republican Control, Bipartisan Control, Majority-Minority×Unified Republican, and Majority-Minority×Bipartisan), and plot both

---

4 Simons and Mallinson (2015) use a slightly different version of the voter liberalism variable compared to Shotts (2003a). We use their measure in our replication and simulations.
the share of positive and negative significant coefficients according to the level of error in DW-Nominates.

Our results can be found on Figure 5(i–v). The Monte Carlo simulation procedure does not change the results from Simons and Mallinson's Conditional Effects Model. While the coefficients on Bipartisan Control and its interaction with Fraction Majority-Minority start to more often become negative significant and positive significant (respectively) as we increase the standard deviation of the error, the switches do not happen often enough to be of concern to Simons and Mallinson's null results.

We also assess robustness of Simons and Mallinson's findings by using Carroll et al.'s (2009) estimated standard errors for individual members. As noted in the previous section, these are only available for the most recent version of DW-Nominates, so we use those for our analysis. Once again, we conducted 1000 simulated regressions, each with a different noise draw based on each member's standard error. We present our results in Table 3, and report how often each coefficient is significant at the 5% level (and further split the frequency of positive significant and negative significant coefficients). As in the original regression, the five coefficients of interest remain statistically indistinguishable from zero for most of our simulations, once again supporting the statistical validity of Simons and Mallinson’s null results.

Table 3: Monte Carlo results using Carroll et al.'s (2009) estimated standard errors for Conditional Effects Model.

<table>
<thead>
<tr>
<th>Conditional effects model</th>
<th>MC simulation results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient (1)</td>
<td>SE (2)</td>
</tr>
<tr>
<td>------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Fraction Majority-Minority</td>
<td>0.2261</td>
</tr>
<tr>
<td>Unified Republican Control</td>
<td>0.0291</td>
</tr>
<tr>
<td>Bipartisan Control</td>
<td>−0.0339</td>
</tr>
<tr>
<td>Fraction x Republican</td>
<td>−0.5072</td>
</tr>
<tr>
<td>Fraction x Bipartisan</td>
<td>−0.0233</td>
</tr>
<tr>
<td>Voter Liberalism</td>
<td>0.005**</td>
</tr>
<tr>
<td>% Black</td>
<td>−0.001</td>
</tr>
<tr>
<td>% Latino</td>
<td>0.0006</td>
</tr>
<tr>
<td>Lagged Left of Median</td>
<td>0.6384***</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.1576**</td>
</tr>
</tbody>
</table>

*p<0.10, **p<0.05, ***p<0.01.

%Significant coefficients from two-sided t-test at 5% level.
i. Fraction Majority-Minority coefficient

ii. Unified Republican Control coefficient

iii. Bipartisan Control coefficient

iv. Interaction: Fraction Majority-Minority x Republican

v. Interaction: Fraction Majority-Minority x Bipartisan

Figure 5: Monte Carlo simulation results for Simons and Mallinson (2015). This figure displays the proportion of 1000 Monte Carlo simulations that yielded positive and negative coefficients that were significant at the 5% level for key variables and interaction effects in Simons and Mallinson's model. The horizontal axis is the level of noise assumed for DW-Nominate.
4 Conclusion

During the past decade, Shotts’s argument that racial redistricting promotes liberal policy outcomes have been subject to a variety of criticisms. Nakao (2011) develops a very nice theoretical model of strategic gerrymandering in the presence of electoral uncertainty, and shows that within his model racial redistricting can have perverse effects. Lublin and Voss (2003) criticize Shotts’s use of a median legislator model. They also argue that Democrats could have done better in the 1990s by selectively undoing racial gerrymanders and reallocating votes; however, as noted by Shotts (2003b), their empirical argument presumes knowledge of electoral outcomes and thus basically amounts to Monday morning quarterbacking.

Simons and Mallinson (2015) provide the most compelling empirical criticism to date of Shotts’s arguments about the effects of majority-minority mandates. Our analysis here largely supports their findings. We show that Shotts’s original results were fragile, in the sense that they relied on an implicit assumption that there was an implausibly-low level of measurement error in DW-Nominates. We also show that Simons and Mallinson’s null results are robust to measurement error in DW-Nominates. Taken together, these findings imply that the jury is still out on the question of whether racial redistricting moves the US House median to the left or to the right.

At a broader level, the fact that Shotts’s results failed to be robust should serve as a cautionary tale for scholars who use DW-Nominates and other estimates of legislators’ ideal points. Given that all ideal point measures are noisy, it is important to check whether empirical findings are robust to using alternative measures and to reasonable levels of noise. The techniques that we used in this paper are one such way of assessing robustness.

Appendix: Registered Research Design

On July 9, 2015, before beginning any analysis, we pre-registered our study with EGAP, under the project title “Assessing Robustness of Findings About Racial Redistricting and Southern House Members” (ID #20150709AA). What follows is the content of that document.

Background information: Shotts (2003a) analyzed the effect of racial redistricting in the 1980s and 1990s, using a dependent variable that was measured using DW-Nominates. Simons and Mallinson (2015) extended this analysis to include more recent Congresses. Along the way they replicated Shotts (2003a). To
do this, they used the current DW-Nominate estimates, and the key result in Shotts (2003a) no longer holds (i.e. is no longer statistically significant). The reason for the discrepancy is that as more votes happen over the years, each representative’s DW-Nominate score for previous years moves a bit. The Shotts (2003a) dependent variable – the proportion of members of a state’s delegation to the left of the House median – apparently was sensitive to small movements in DW-Nominate estimates for individuals near the median.

Our goal is to lay out simple tools that can be used to assess whether Shotts’s results are robust to a small amount of error in the DW-Nominate. The motivation for our analysis is methodological, i.e. to identify things that an applied researcher like Shotts (or Simons and Mallinson) could and should have done to take into account the possibility of measurement error.

Our analysis will use two different techniques.

The first technique is reestimation of the Shotts (2003a) model using data produced from Monte Carlo simulations, allowing for errors in DW-Nominate. This technique will enable us to check whether the results in Shotts (2003a) are robust to reasonable levels of error. This method is far from perfect, because errors in DW-Nominate presumably are correlated both across time and across members. However, to the best of our knowledge, no technique has been developed that would deal with this issue in DW-Nominate error. Hence, we will treat the errors as being independent.

The second technique is to use a different measure of the ideology of elected officials, based on the Bonica (2014) estimates.

Most of the input data for this study already exist and are readily available. All data used in Shotts’s original study are available in his replication archive. Current DW-Nominate estimates, along with estimated standard errors are available from Keith Poole’s website. The Bonica scores are available from his website. Simons and Mallinson have graciously provided the data they used in their extension of Shotts.

As of the time that this document is being submitted to EGAP, we have not conducted any analysis of the data.

Part 1: Re-Analysis of Shotts (2003a)

1-1. Plot of raw DW-Nominate. To give a sense of the basic data before and after the 1990s redistricting, we will plot the Old and Current DW-Nominate for Southern Representatives and the House Median in the 102nd, 103rd, and 104th Congresses. This will give a rough sense of whether the count of the number of Southern representatives to the left of the median is sensitive to errors in measurement for DW-Nominate.

1-2. Reestimation of the original model using samples obtained from Monte Carlo simulations that allow for different levels of error in Old-DW-Nominate.
(i.e. the ones used in Shotts 2003a). We will perform regressions for Shott's key empirical specification (Table 1, p. 221) in a fashion similar to bootstrap resampling procedures. All variables except the dependent variable \textit{Left of Median} and its lag will be exactly as in Shotts (2003a). For the \textit{Left of Median} variable, each iteration of the simulation will draw for each House Member in each year a DW-Nominate score from a normal distribution, with mean at his/her estimated mean. The standard deviation of the draws will range from 0 to 0.2, in 0.01 unit increments. Based on the work of Carroll et al. (pp. 270–271) we believe that the most relevant standard deviation will be one around 0.05. However, we are examining smaller and larger values as well.

We will perform 1000 iterations for each standard deviation between 0 and 0.2. The product of this analysis will be a figure. On the horizontal axis will be the standard deviation. On the vertical axis will be the proportion of the 1000 iterations that resulted in a statistically-significant (at the conventional 0.05 level) positive coefficient on the key variable of interest in Shotts (2003a), namely \textit{Fraction Majority-Minority}. We also will plot the proportion of iterations that resulted in a statistically-significant negative coefficient on \textit{Fraction Majority-Minority}, and report the mean and standard error from our iteration estimates.

Given that Shott's results were no longer statistically significant in Simons and Mallinson's analysis (which used DW-Nominates that they downloaded recently from Poole's website), we expect that a small amount of noise will dramatically reduce the proportion of the time that there is a statistically-significant positive coefficient on \textit{Fraction Majority-Minority}. However, to be clear, we do not have strong priors on how large the effect will be for different levels of the standard deviation.

We do not expect that there will be many statistically-significant negative coefficients on \textit{Fraction Majority-Minority}, at least for small standard deviations.

1-3. Monte Carlo using Carroll et al. (2009) estimated standard errors in Current-DW-Nominates for individual members. This analysis will be very similar to part 1-2. However, there will be only one level of standard deviation in the DW-Nominate for each member of Congress. This will vary across members. Also, the DW-Nominates that we will use will be Current-DW-Nominates, because those are the ones for which estimated standard errors are available.

We will conduct 1000 iterations of the simulated regression, and will report the proportion of the time that the coefficient on \textit{Fraction Majority-Minority} was positive and statistically significant and the proportion of the time that it was negative and statistically significant at the conventional 0.05 level.

1-4. Using Bonica scores to measure \textit{Left of Median}. Another, very different, way of assessing the robustness of Shott's results is to use a different measure of politicians' ideal points. We will do this using Bonica's (2014) measures, which
are based on interest group contributions. For this analysis we will replicate Table 1 from Shotts (2003a), using the same data for every variable except Fraction Left of Median and its lag, both of which will be calculated using Bonica's estimates of members’ ideological positions.

Part 2: Application to Simons and Mallinson

2-1. Monte Carlo simulation for different levels of error in Current-DW-Nominate. As above, in part 1-2, we will re-run Simons and Mallinson’s model (the “Conditional Effects Model” in Table 3 of the March 2015 version of their paper) using a Monte Carlo Simulation with the standard deviation of the DW-Nominate estimates allowed to vary from 0 to 0.2 in 0.01 unit increments. For each of the key independent variables (Fraction Majority-Minority, Unified Republican Control, Bipartisan Control, Majority-Minority X Unified Republican, and Majority-Minority X Bipartisan), we will plot the proportion of iterations that resulted in a statistically-significant positive coefficient as well as the proportion that resulted in a statistically-significant negative coefficient.

2-2. Monte Carlo simulation using Carroll et al. (2009) estimated standard errors. As in part 1-3, we will re-run Simons and Mallinson’s model using a simulation with estimated standard errors for Current-DW-Nominate. For each of the key independent variables, we will report the proportion of iterations in which the coefficient was positive and statistically significant as well as the proportion in which it was negative and statistically significant.

References


**Article note:** We thank Joseph Simons and Daniel Mallinson for graciously sharing their data with us and for engaging in a collegial and productive scholarly exchange. Any omissions or errors are our own. Part of Ken Shotts’s work on this project was completed during a sabbatical at the Center for Advanced Study in the Behavioral Sciences at Stanford University. Replication data and code are available at http://dx.doi.org/10.7910/DVN/RX0WGF.