Response to Chen and Rodden's "Report on Computer Simulations of Florida Congressional Districting Plans"

Nolan McCarty
Susan Dod Brown Professor of Politics and Public Affairs
Chair, Department of Politics
Princeton University

Nolan McCarty is the Susan Dod Brown Professor of Politics and Public Affairs and the chair of the Politics Department at Princeton University. He has published widely on topics related to legislative and electoral politics, especially the effects of electoral rules on legislative partisanship and polarization. He has used computer simulations of legislative districting to study the possible impact of gerrymandering on legislative polarization. He is also the co-editor-in-chief of the *Quarterly Journal of Political Science.*

In preparation, I consulted the February 15, 2013 "Report on Computer Simulations of Florida Congressional Districting Plans" prepared by Professors Jowei Chen and Jonathan Rodden. I also consulted current and prior versions of their forthcoming paper "Unintentional Gerrymandering: Political Geography and Electoral Bias." I am also familiar with this work through my editorial duties at the *Quarterly Journal of Political Science*, the journal in which this work is slated to appear.

The analysis described below relies primarily on reexamination of data provided by Professors Chen and Rodden. I may also supplement these analyses with additional ones based on data from other Florida elections drawn from the Harvard Election Data Archive. These data were compiled by Professor Rodden in conjunction with Harvard Professor Stephen Ansolabehere.

******

The central issue in Professor Chen and Rodden's academic work and expert report revolves around the empirical finding that electoral support for the Democratic Party and its candidates tends to be far more geographically concentrated than support for the Republican Party. This pattern results because citizens of densely populated urban areas overwhelming vote Democratic while suburban and rural voters tend to support Republicans but at lower rates. These geographic patterns in partisanship have important implications for the evaluation of legislative districting plans. If districts are required to be geographically compact and contiguous, Democratic voters will tend to be very concentrated in urban-centered districts while Republican voters will be more evenly distributed across rural and suburban districts. This effect

produces lopsided Democratic victories in cities causing many Democratic votes to be "wasted." In turn the Republicans benefit from an "unintentional gerrymander" where the party wins a greater share of legislative seats than its share of votes.[1]

The purpose of Chen and Rodden's computer simulation methodology is to establish that such natural geographic biases exist and to provide measures of their magnitude. The baseline simulations draw district boundaries based solely upon considerations of contiguity and compactness. Then using precinct-level electoral data, Chen and Rodden compute how many of the randomly generated districts produce a Republican majority. This number can be compared to the number of Republican majority districts in an enacted districting plan. When the number of Republican districts in the enacted plan exceeds that of the simulated plan, the authors count that as evidence of a Republican bias beyond that which can accounted for solely by the geographic concentration of Democratic voters.

In their academic work, Chen and Rodden find geographical sources of Republican advantage are more than sufficient to explain the observed number of Republican districts in the Florida congressional districting plan enacted in 2002. In their simulations, they use the 2000 George W. Bush presidential vote in Florida as the measure of Republican support. Approximately 25% of their simulated districting plans contain as many or more Bush majority districts as produced by the enacted plan (17 of 25 districts).[2] Under conventions for assessing

[1] This unintentional gerrymander might be undone by consciously linking a small cluster of urban areas to otherwise suburban or rural districts. But even were this to be done without violation of constraints on the contiguity and compactness of districts, it would tend not to respect municipal boundaries – cities would have to be carved up and placed in a large number of suburban-oriented districts.

[2] Chen and Rodden also considered hypothetical district plans with 27 seats – the number of districts in the 2012 plan. In these simulations, the average number of Bush majority districts was 17 out of 27 (see Figure 4).

statistical significance, Chen and Rodden's results would not lead to the rejection of the hypothesis that the enacted plans were constructed without partisan considerations.

In their expert report, Chen and Rodden use simulated districting to evaluate the congressional districting plan enacted in 2012. Their measure of Republican support is the 2008 John McCain presidential vote in Florida. In the simulations most comparable to those in the academic work, Chen and Rodden simulate 1000 districting plans and find that none of them produces as many McCain majority districts as the 17 out of 27 produced by the 2012 enacted plan. Chen and Rodden also conduct simulations that respect requirements to maintain the representation of racial minorities and respect county boundaries. But these simulations rarely produce plans with seventeen or more McCain majority districts. These findings lead them to conclude that it is statistically unlikely the 2012 congressional plan was formulated without partisan considerations benefitting the Republicans.

The findings of the expert report are difficult to reconcile with Chen and Rodden's academic work. The building blocks of both simulations are the same: geo-coded precinct-level voting data. Were their methodology robust and consistently applied, the difference between the expert and academic findings could only be accounted for by a major shift in the geographic distribution of partisanship in Florida between 2000 and 2008.[3] There is little evidence, however, for such shifts. It is my opinion therefore that the differences in findings are due to an inconsistent application and limitations of their simulation methodology.

Inconsistent Application of the Methodology

[3] A "robust" methodology would be one for which conclusions and inferences do not change dramatically in response to small changes in the underlying data.

In a first-past-the-post electoral system such as that of the United States, a party's legislative seat share is generally not proportional to its vote share.[4] Typically, the majority party obtains a greater percentage of seats than it receives in votes. Conversely, the smaller party receives a lower proportion of seats than it does votes. Because of this well-established pattern, scholars typically assess electoral bias using a counterfactual based on the seat share a party receives in a hypothetical tied election. More formally, scholars define a pro-Republican bias as the party's predicted seat share were it to win 50% of the vote.

In their forthcoming paper on "Unintentional Gerrymandering," Chen and Rodden adopt this convention. In the Florida case study, they tout Florida's 2000 presidential race because of its "unique quality as a tied election."[5] Thus, the Bush vote total precisely meets the assumption required to measure a pro-Republican bias. Moreover, in their cross-state study, Chen and Rodden simulate tied elections by applying a "uniform swing" which adjusts precinct level voting outcomes to produce a tied election at the state level.[6] In their expert report, however, Chen and Rodden fail to adhere to this convention. Their measure of Republican support is the unadjusted John McCain vote from 2008. As Senator McCain only received 48.6% percent of the major party vote, the baseline is no longer a tied election.

---

[4] A first-past-the-post or *plurality* electoral system is one in which the candidate who receives the largest number of votes in a district election wins a seat in the legislature. In such a system, the votes of smaller parties are often "wasted" as their candidates may receive as much as 49% of the district vote without winning a seat. So the larger party generally wins seats at a greater proportion than its vote share. Alternatively, in a *proportional* representation system such as used in much of continental Europe, legislative seats are allocated to parties in rough proportion to votes received.

[5] "Unintentional Gerrymandering" page 9.

[6] "Unintentional Gerrymandering" page 20.

Fortunately, it is straightforward to evaluate the robustness of the expert report to the correction of this methodological inconsistency. In order to simulate the effects of a tied election, I apply a "uniform swing" to the McCain vote. This involves adding 1.4% McCain (50 – 48.6) to simulated district vote totals.[7]

I then used Chen and Rodden's simulated district boundaries to recalculate the proportion of enacted and randomly generated districts for which the simulated McCain vote exceeds 50%. Based on the enacted district boundaries, seventeen districts had McCain majorities. Importantly, this is exactly the same as the total for the unadjusted McCain vote. But when the adjusted vote totals are used, the likelihood that random districting plans produced at least seventeen McCain majority districts was substantially larger. Whereas Chen and Rodden report that only 8 of 1000 random districting plans using three majority-minority districts produced seventeen or more McCain majority districts, I find almost 300 of 1000 random plans predicted seventeen or more McCain districts when the data are adjusted to simulate a tied election. Table 1 summarizes my results from replicating all of Chen and Rodden's districting scenarios after applying a uniform swing to the McCain vote. My findings suggest that there are reasonable odds that the actual districting plan could have produced seventeen McCain majority districts without any intention of maximizing the Republican share of congressional seats.

While scholars may reasonably disagree as to the importance of the "tied election" benchmark for analyses such as these, the fact that the results change so dramatically when it is employed raises concerns about the robustness of Chena and Rodden's methodology.

---

[7] Ideally one would add 1.4% to each precinct–level return. But the materials provided by Chen and Rodden do not include sufficient information to confidently link precinct returns to the district simulation. Fortunately, my procedure of adding 1.4% to each simulated district is roughly equivalent to adding 1.4% to each precinct total.

Robustness

My reanalysis raises important questions about the robustness of Chen and Rodden's methodology and the extent to which it is appropriate for the task of assessing the role partisan gerrymandering in the 2012 congressional districting plan. In their academic work, Chen and Rodden apply their methodology to show that geographic concentrations of Democratic voters along with requirements of contiguous and compact legislative districts produce natural biases towards the Republicans. Although the individual estimates of this natural bias may not be precise, the cumulative evidence across many states and districting plans is compelling. In the expert report, however, Chen and Rodden use their simulations to evaluate one specific districting plan as to its constitutionality. So precision in the estimates of the natural Republican bias are crucial. Claims of this degree of precision are undermined by the fact that small changes in the Republican vote totals used in the simulation produce substantial swings in the estimate of the probability that the actual districting plan was produced without partisan considerations.

To illustrate this problem, I performed the following analysis. First, I applied district-level swings to the McCain vote ranging from -1.5% to 2%. These swings are intended to show how the analysis might change given small changes in the measure of Republican support used. Second, I recalculated the number of McCain majority districts in the enacted districting plan. Third, I used Chen and Rodden's simulated plans with 3 majority-minority districts and computed the probability of a random plan generating as many McCain seats as the enacting plan.

As seen in Table 2, the probability of the number of McCain seats meeting or exceeding the prediction of the enacted plan varies greatly across the different swings. For example, if Chen and Rodden had analyzed an election where the Republican had uniformly underperformed McCain by as little as .5%, they would have concluded that there was a 23% chance that Republican performance under a random plan would meet or exceed that of the enacted plan. If they had used an election where the Republican candidate uniformly outperformed McCain by 1%, they would have found a 17% chance of a simulated plan leading to as good or better Republican performance than in the enacted plan. Indeed, it is Chen and Rodden's use of the unadjusted McCain vote that produces the outcomes most favorable to the claims of the expert report.

The analysis in Table 2 illustrates how Chen and Rodden's inferences might have been different had they used electoral data where the Republican candidate uniformly over- or underperformed John McCain. But these findings may well overstate the robustness of their conclusions. In particular, the use of the 2008 presidential election returns may be especially problematic for Chen and Rodden's inferences. As has been well documented, turnout among young, urban, and African-American voters was much higher than normal in 2008. This turnout surge will inflate the overconcentration of Democratic voters in urban areas and leads to Chen and Rodden's procedure to estimate a larger Republican bias. More generally the use of presidential vote outcomes to predict congressional elections is problematic. First, presidential vote is only a crude measure of partisanship and may not predict congressional voting patterns. Second, a presidential candidate may easily over- or under-perform partisan expectations.

Given these concerns, the best way to evaluate Professor Chen and Rodden's conclusions would be to use precinct-level vote returns from other Florida statewide elections. Unfortunately, Professor Chen and Rodden's submission currently lacks a template for confidently matching such returns to the output of their districting plan simulations. Were that template provided, I would be able to replicate the analysis on a variety of statewide elections using data from the Harvard Elections Data Archive.

Summary

In their academic work, Professors Chen and Rodden provide significant evidence for the proposition that the spatial distribution of partisan voters creates representational biases in favor of the Republicans. But in their expert report, they argue that this geographical bias is too small to explain the Republican success under the 2012 Florida districting plan. But their estimates in the expert report are inconsistent with those reported in their academic work. This difference can be accounted for entirely by differences in the methodology employed in each instance. Following Chen and Rodden's advocated procedure of focusing on tied election or using uniform swings to simulate them, my analyses show no evidence in favor of the proposition that the 2012 congressional districting plan was more Republican biased than required by the geographical constraints or minority representation considerations required by the Florida constitution. But the central issue may not be whether or not tied elections are the appropriate benchmark. My second analysis shows that any small swings in the baseline Republican vote can support a variety of conclusions about the enacted plan. This lack of robustness raises concerns about the conclusions of Chen and Rodden's expert report.

Table 1		
Percentage of Outcomes with Seventeen or More McCain-Majority Districts		
	Chen-Rodden	Uniform Swing
Three Majority Black Districts	.8%	28.1%
Three Majority Black Districts, 46 Intact Counties	.7%	19.9%
Two Majority Black Districts, 20th and 24th	.4%	14.6%
Two Majority Black Districts, 20th and 24th, 46 Intact Counties	.2%	10.6%
Two Majority Black Districts, 5th and 24th	.8%	22.6%
Two Majority Black Districts, 5th and 24th, 46 Intact Counties	.6%	15.4%
One Majority Black District, 24th	.8%	12.6%
One Majority Black Districts, 24th, 46 Intact Counties	.2%	10.9%
No Black Majority Districts	0%	2.1%
No Black Majority Districts, 46 Intact Counties	0%	1.9%

Table 2		
Swing	**McCain Majority Districts in Enacted Plan**	**Simulated Probability of Obtaining at Least as Many McCain Majority Districts as in the Enacted Plan**
-1.5%	13	41.7%
-1.0%	15	8.2%
-0.5%	15	23.0%
0.0%	17	.7%
0.5%	17	4.4%
1.0%	17	17.2%
1.5%	17	30.1%
2.0%	19	0.2%