Electoral Studies 44 (2016) 329–340

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

Evaluating partisan gains from Congressional gerrymandering: Using computer simulations to estimate the effect of gerrymandering in the U.S. House[☆]

Jowei Chen [a], David Cottrell [b, *]

[a] Department of Political Science, University of Michigan, USA
[b] Program in Quantitative Social Science, Dartmouth College, USA

A R T I C L E I N F O

Article history:
Received 22 December 2013
Received in revised form
11 June 2016
Accepted 22 June 2016
Available online 20 August 2016

Keywords:
Gerrymandering
Electoral bias
Redistricting
Electoral geography

A B S T R A C T

What is the effect of gerrymandering on the partisan outcomes of United States Congressional elections? A major challenge to answering this question is in determining the outcomes that would have resulted in the absence of gerrymandering. Since we only observe Congressional elections where the districts have potentially been gerrymandered, we lack a non-gerrymandered counterfactual that would allow us to isolate its true effect. To overcome this challenge, we conduct computer simulations of the districting process to redraw the boundaries of Congressional districts without partisan intent. By estimating the outcomes of these non-gerrymandered districts, we are able to establish the non-gerrymandered counterfactual against which the actual outcomes can be compared. The analysis reveals that while Republican and Democratic gerrymandering affects the partisan outcomes of Congressional elections in some states, the net effect across the states is modest, creating no more than one new Republican seat in Congress. Therefore, the partisan composition of Congress can mostly be explained by non-partisan districting, suggesting that much of the electoral bias in Congressional elections is caused by factors other than partisan intent in the districting process.

1. Introduction

How does the gerrymandering of United States Congressional districts affect parties' control over legislative seats? In 2012, Democratic Party candidates managed to win only 201 of 435 US House of Representatives elections despite receiving an overall majority of the total combined votes in nationwide House election races. The prevailing presumption among media pundits and political commentators was that such a disparity between a party's legislative seat share and its underlying vote share reflects a concerted district gerrymandering effort by Republican state legislatures. Such gerrymandering, it was presumed, enabled Republicans in many states to win more legislative seats than warranted by their underlying vote support. How accurate are these claims regarding gerrymandering, and how extensive is gerrymandering's overall effect on each party's control over Congressional seats? Isolating and precisely measuring the impact of gerrymandering on partisan control of Congressional seats requires us to analyze a counterfactual: How many legislative seats would each party control in the complete absence of any gerrymandering?

Previous studies of redistricting have produced mixed conclusions with respect to the partisan electoral effects of gerrymandering. Some studies have shown that partisan redistricting produces substantial partisan biases in the outcomes of legislative elections (Abramowitz, 1983; Erikson, 1972; King, 1989; Cox and Katz, 1999; McDonald, 2004). While other studies have found that partisan redistricting produces minor, mixed or null partisan effects (Ferejohn, 1977; Glazer et al., 1987; Squire, 1985; Gelman and King, 1994; Abramowitz et al., 2006). This lack of consensus is indicative of the challenge that researchers face in isolating and measuring gerrymandering's effect on election outcomes.

To accurately measure the effect of gerrymandering, we would ideally analyze the partisan control of each district in the same

☆ Replication data and code for the results and figures in this article are online at: http://www.umich.edu/~jowei/gerrymandering/.

* Corresponding author.
E-mail addresses: jowei@umich.edu (J. Chen), David.Cottrell@dartmouth.edu (D. Cottrell).

http://dx.doi.org/10.1016/j.electstud.2016.06.014
0261-3794/

election under two different conditions, with and without gerrymandering. But unfortunately, in any given election, we are only able to observe the outcome of each district under a single condition: Either a district is gerrymandered, or it is not. Empirically, we do not observe the partisan control of the districts under a non-gerrymandered counterfactual. Therefore, in order to measure the effect of gerrymandering, scholars have been forced to *estimate* this counterfactual.

This estimation has been done in one of two different ways in past literature. The first method has been to leverage temporal variation in the electoral outcomes of districts before and after redistricting. Under this approach, non-gerrymandered districts at one point in time are used as a counterfactual for gerrymandered districts at another point in time. Any deviation between the two districts is, therefore, considered bias due to gerrymandering. The second method has been to simply posit that a non-gerrymandered counterfactual is one in which the districts produce a "fair" votes-to-seats curve, where the mapping between votes and seats is the same for both parties. Under this approach, any deviation from a "fair" curve is defined as bias due to gerrymandering.

While both methods attempt to measure the bias due to gerrymandering, they are both subject to potential confounding factors. For example, in the former approach, one must make the assumption that there are no temporal differences between the two outcomes other than those associated with gerrymandering. Therefore, shifts in public opinion, district demographics, and other variables that similarly affect electoral outcomes may falsely generate bias that is attributed to gerrymandering. Likewise, in the latter approach, one must assume that "unfair" votes-to-seats curves are caused only by biased districting. However, votes-to-seats curves can be affected by changes in the underlying partisan and racial distribution of voters across geographic space. For example, the geographic clustering of Democratic voters may shift the votes-to-seats curve in a way that tends to favor Republicans (Erikson, 1972, 2002; Jacobson, 2003).

Therefore, any measurement that claims to isolate the biasing effect of gerrymandering must successfully eliminate both temporal and spatial confounding factors. This requires that the estimated counterfactual share the same temporal and spatial conditions as the observed outcome. In this manuscript, we construct and analyze such a counterfactual: We develop a new method of simulating how districts would have been constructed under a non-gerrymandered, partisan-neutral process. We then compare these non-gerrymandered outcomes to the redistricting plans enacted in each state, allowing us to comprehensively measure the partisan effects of gerrymandering.

Unlike counterfactuals established in previous research, the counterfactual used in this paper fully accounts for the electoral bias in a districting plan that would have likely been achieved had the districts been drawn without partisan intent. By using this counterfactual, we account for the possibility that partisan bias in a districting plan can be produced as an unintended consequence of innocuously drawing equally apportioned, contiguous, and reasonably compact districts. Therefore, we avoid erroneously assuming that partisan bias in a state's districting plan is necessarily due to gerrymandering when it may actually be a natural outcome of a non-partisan districting process. We also account for the possibility that the absence of partisan bias in a districting plan does not necessarily imply the absence of gerrymandering. Instead, gerrymandering may be used to reduce the partisan advantage that the opposing party would have received naturally. Therefore, we also avoid erroneously concluding that gerrymandering was not present when it actually was. As a result of establishing this counterfactual, we are able to make stronger claims about the likelihood that a plan was gerrymandered, as well as the effects of such gerrymandering.

2. Identifying gerrymandering: improving on previous methods

To infer with confidence that partisan gerrymandering is responsible for the boundaries of a given state's Congressional districts, we first need to establish a non-gerrymandered counterfactual against which the actual districts can be compared. If the actual districting plans produce bias that is significantly different from the non-gerrymandered counterfactual, then we can infer that the districts have been gerrymandered. The key is in establishing this counterfactual precisely.

Typically scholars have attempted to analyze this counterfactual using one of two methods. One approach has been to perform a pre-post test, with post-redistricting plans as the treatment group and pre-redistricting plans as the counterfactual. A second approach has been to analyze the vote-to-seat relationship, where the current vote-to-seats relationship is compared to a function that translates a toss-up election into a two-party split of the legislative seats.

2.1. Pre-post test

The pre-post test takes advantage of the temporal variation in election outcomes, comparing a state's districts before and redistricting. In this test, the pre-redistricting plan is assumed to be the non-gerrymandered counterfactuals against which the post-redistricting plan is to be compared. Differences before and after redistricting are attributed to bias caused by gerrymandering. For example, Campagna and Grofman (1990), Herron and Wiseman (2008), and Goedert (2014) all find significant biasing effects of partisan redistricting by comparing electoral outcomes across time.

However, other scholars have cautioned against drawing such conclusions from the observational comparison between pre- and post-redistricting outcomes. Abramowitz et al. (2006) have argued that change during the redistricting cycle may simply be the result of concurrent demographic changes and ideological realignments within the electorate. Relatedly, Masket et al. (2012) observe districting effects on polarization in California's Assembly and find that the effect of redistricting on polarization pales in comparison to the effect of shifts in electoral partisanship.

These scholars suggest that the difference in electoral results between pre- and post-redistricting is, in part, a function of something other than redistricting itself. As a result, these scholars express a concern that pre-post comparisons potentially fail to distinguish gerrymandering from a number of possible confounding factors. Their concerns reflect the need for identifying the true counterfactual to gerrymandered districts: A counterfactual that establishes how a state's districts would hypothetically have been drawn in the absence of intentional partisan gerrymandering while controlling for or separating out the effects of demographic and ideological shifts in the electorate across both space and time.

The salient feature of these scholars' pre-post redistricting comparisons is that such comparisons are designed to measure the change in electoral bias due to one particular redistricting cycle. But they are not designed to measure the cumulative effect of gerrymandering across all past redistricting cycles.

Hence, for a given state, we might observe a negligible difference in electoral bias before and after a particular redistricting cycle. This lack of a pre-post difference, however, does not necessarily imply the absence of gerrymandering: The newly enacted districting plan may simply have preserved what was already a heavily gerrymandered districting plan in the first place. As Cox and Katz (1999, 2002) remind us in their discussion of reversionary plans, we must be careful not to ignore the significance of legislatures that do not change the composition of the state's districts, but instead revert to the status quo. Consequently, there may be no differences between pre- and post-redistricting, but the resulting map may nonetheless be a product of past gerrymandering strategies.

For instance, instead of choosing to redraw districts entirely without respect to partisanship, government majorities may instead draw a plan similar to the previous plan, which may itself be a product of gerrymandering. Though such a strategy would not marginally increase partisan bias in the state's districts relative to the previous plan, it does preserve the compounded biases that may have accumulated over decades of redistricting. A proper analysis of redistricting should identify such a strategy as creating significant partisan bias relative to a non-gerrymandered plan.

Given the concerns with identifying gerrymandering using pre-post tests, we develop a measurement that accounts both for the underlying partisan changes in the electorate as well as the compounding effects of redistricting over time. Our measure simply determines the likelihood that a state's districts have been drawn with respect to partisanship, regardless of the prior enacted plan.

2.2. Vote share to seat share relationship

Another baseline typically used by researchers to identify gerrymandering among the states is the state's relationship between a party's vote share and its legislative seat share. For example, Tufte (1973) famously measures this votes-to-seats curve across U.S. states and partially attributes electoral bias and a declining swing-ratio to gerrymandering. Similarly, Gelman and King (1994) estimate electoral bias by analyzing the functional relationship between partisan vote share and the partisan control of legislative seats during the years 1968–1988, while Erikson (1972) analyzes electoral bias in Congressional districts in a similar manner.

However, some scholars have recently noted that in various states, this votes-to-seats relationship may be inherently biased due not to gerrymandering, but to the underlying residential asymmetries in the geographic distribution of Democratic and Republican voters. McDonald (2009) notes that Democratic voters in some states are less efficiently distributed across districts, leading to electoral bias favoring Republican control of legislative seats. Ansolabehere et al. (2006) illustrate the theoretical basis for this geographically-induced bias, explaining how an asymmetric distribution of within-district median voters can cause long-term electoral disadvantage against a party.

Hence, these studies suggest the need to account for the underlying geographic distribution of Democratic and Republican voters. In this manuscript, we develop a method of simulating how a non-gerrymandered districting plan might have been drawn, but we analyze such plans in the context of the actual geographic distribution of voters in each state. This method allows us to isolate the effect of gerrymandering on electoral bias and separate it from the electoral bias due to underlying spatial distributions of voters.

2.3. Controlling for geographic factors that lead to bias

In quantifying the amount of bias in a districting plan caused by gerrymandering, we must control for certain geographic factors - like the spatial distribution of partisans and the geometric shape of district boundaries - that lead to this bias independent of gerrymandering. The bias due to these geographic factors occurs due to the interaction of two conditions:

First, common residential patterns produce spatially clustered, partisan populations. This means that Democrats and Republicans are often distributed non-uniformly across geographic space, creating a complex spatial distribution of partisans that Kendall and Stuart (1950) described as being the product of powerful geographical and historical as well as political and demographic factors (195). Second, states draw districts according to certain geographic criteria that make this spatial distribution of partisans salient. For example, when designing district boundaries, states apportion their districts with equal populations and follow the basic governing principles of contiguity and compactness. In other words, districts are drawn to include communities of individuals who reside in similar locations rather than individuals who live far apart from one another. Moreover, many states have additional standards requiring district lines to maintain the integrity of certain geographic entities (i.e. municipalities).

Therefore, independent of gerrymandering, basic geographic limitations on redistricting increase the likelihood that two neighbors will share a Congressional district. Natural residential clustering of partisan voters causes voters of the same party to be more likely to share a Congressional district, even when districts are drawn in a non-partisan fashion. While this result is observationally consistent with the packed partisan districts that often arise from gerrymandering, it is instead the product of nonpartisan geographic factors that influence the districting process. Therefore, to identify gerrymandering, one must account for the confounding effects of partisan clustering. Chen and Rodden (2013) make this point with respect to Florida, providing evidence that some of Florida's bias toward Republicans can be explained by its overwhelming clustering of Democrats in urban centers, rather than intentional partisan gerrymandering.

3. Establishing the baseline: simulating the absence of gerrymandering

We can attempt to construct this non-gerrymandered baseline through computer simulations. The benefit of using a computer to simulate districting is that it allows us to draw the same number of districts from the same complex geographic distribution of voters using the same geographic criteria for redistricting as the state boundary-makers themselves. The only distinction is that the computer is indifferent to partisan outcomes. Therefore, the partisan bias that arises from these simulated districting plans can be identified as the bias that arises simply by chance alone. Ultimately, we obtain a distribution of hypothetical election outcomes for a given state as if the lines were drawn randomly with respect partisanship. Using this distribution as a baseline, we can then compare the actual election outcome within the state to this distribution of simulated non-gerrymandered outcomes. If the actual outcome and the simulated outcome are the same, then it can be said that the districts produce a result that is no different from a result that would have been produced had the districts been drawn without partisan intent. However, the more the actual outcome deviates from the baseline distribution, the more confidence we can have that the state engaged in gerrymandering.

3.1. The automated districting algorithm

We conduct congressional districting simulations designed to draw geographically compact, contiguous, and equally apportioned districts in each state using precinct-level maps and voting results from the 2008 McCain-Obama election.[7] This section explains this algorithm by illustrating its implementation in Florida.

As of the November 2008 election, Florida consisted of 6984 voting precincts - the smallest geographic unit at which election results are publicly announced. We map the votes of these precincts to Florida's 484,481 Census blocks according to population and then aggregate the votes into a set of 15,640 similarly-populated square polygons so as to produce a geographically-precise spatial grid of the state. These 15,640 "squares" of the grid are then used as the building blocks for the districting simulations.[1] Hence, a complete districting plan consists of assigning each one of Florida's building blocks to a single legislative district, such that all districts are equally populated, compact, and geographically contiguous.

The simulation proceeds as follows. Suppose Florida is to be divided into 27 districts, the current size of Florida's Congressional delegation. First, one of the 15,640 squares is randomly chosen as the "seed" of the first district. The first district is then created by assigning as many geographically closest neighboring squares to the seed as is necessary to comprise a fully populated first district. The second district is similarly created by randomly selecting an unassigned square and combining it with as many geographically nearest unassigned squares as needed to comprise a full second district. This process is repeated until 27 fully populated districts are formed in Florida. At any step along the way, the plan is abandoned, and the simulation algorithm starts anew, if the plan reaches a situation in which 27 contiguous, equally populated districts cannot possibly emerge.

At this point, Florida has been divided into 27 geographically contiguous districts. These districts are equally apportioned by population and reasonably, but not perfectly, compact. The computer then iteratively makes compactness improvements to the plan through the following process: One of the 15,640 squares is chosen at random. If the square belongs to district A but borders district B, then the algorithm determines whether reassigning the square from district A to district B would improve the compactness of both districts, while nevertheless leaving both districts within 5% of the ideal district population. If all of these conditions are satisfied, then the randomly-chosen square is reassigned from district A to B. This process is repeated until no further compactness-improving reassignments can be identified. Once complete, the computer will have divided Florida into 27 equally populated, contiguous, and reasonably compact districts. As illustrated in Fig. 1, we begin with the spatial grid of Florida, where precinct-level presidential votes have been mapped onto 15,640 square polygons. These square polygons are depicted in the left map of Fig. 1 and shaded from blue to red, indicating the McCain vote share of each square. The computer simulation process results in a set of 27 randomly-drawn, computer-generated districts, and the right map of Fig. 1 displays an example of a complete districting plan produced by these simulations.[2]

[1] By mapping precinct votes onto this smaller-scale spatial grid, we better address the Modifiable Areal Unit Problem (MAUP) (Openshaw, 1984). MAUP refers to the biases that can occur from using spatially aggregated data. Both the scale and the shape of the spatially aggregated units can distort measurements of the underlying population. One solution to this problem is to reduce the spatial scale of the precincts into smaller, more uniformly shaped units. We do this by overlaying precinct boundaries onto their constituent Census blocks, and dividing the presidential votes among the blocks according to their share of the voting-aged population in that precinct. We then scale up slightly, combining blocks into a spatial grid, where the units are uniformly square and small in size (approximately 2000 people each).

[2] Our simulation algorithm differs in several important ways from the methodology employed by Chen and Rodden (2013). First, Chen and Rodden (2013) use precincts as the building blocks for their districts. Our algorithm, in order to address the Modifiable Aerial Unit Problem discussed in Footnote [1], first creates a grid of square polygons, each of which is no greater than 2000 in population, by combining census blocks. These squares are then the building blocks for simulated districts. Second, Chen and Rodden (2013) create all districts simultaneously and with unequal populations, using later iterations to adjust district populations to an acceptable threshold. Our algorithm, by contrast, creates districts one-by-one in order to make districts as equally populated and as compact as possible, rather than simply reaching a target threshold of population equality. The motivation for our different algorithm is that we seek to analyze how districts would have been drawn when strictly following traditional districting criteria of population equality and compactness.

We iterate our procedure 200 times for each state with more than one House district.[3] After completing this simulation procedure, we aggregate the precinct-level McCain-Obama vote counts to calculate the partisanship of each of the computer-simulated districts. The following section explains how we use these simulated district calculations to measure the extent of gerrymandering in each state.

4. Measuring gerrymandering using districting simulations

Gerrymandering is measured by comparing the electoral outcomes from these simulations to the electoral outcomes of the actual districts in each state. The simulations are intended to provide us with a baseline for what a state's Congressional districts would look like in the absence of partisan gerrymandering. Attempting to mimic the districting procedure under the same minimal constraints faced by the actual boundary-makers, the simulations estimate the distribution of potential outcomes under the null hypothesis that districts are drawn at random with respect to partisanship. Therefore, if there is no significant partisan difference between the simulated and actual electoral outcomes, then we cannot reject the claim that the districts were drawn without partisan intent. However, if a significant partisan difference does exist, we can infer that the actual maps, unlike the simulations, are the result of a districting procedure that has intended partisan consequences. In other words, assuming that the state boundary-makers follow the same basic rules as the automated districting algorithm in drawing compact, contiguous, and equally apportioned congressional districts, we can attribute the partisan difference between the simulated and actual plans to gerrymandering.

An illustration of what gerrymandering might look like using this measure is presented in Fig. 2. Here we focus on Florida's presidential election results in 2008. The first plot in the figure compares John McCain's share of the two-party vote to what his vote share would have looked like in each of Florida's 27 Congressional districts had the districts been drawn under a non-gerrymandered simulation.

To produce this plot, we first perform 200 independent districting simulations, each of which divides the 15,640 polygons, derived from Florida's 484,481 census blocks, into 27 hypothetical districts. For each simulation we calculate the McCain-Obama (November 2008) vote by aggregating all recorded votes for McCain and Obama across all the polygons contained in each of the simulated districts. Then we arrange the districts in order from least Republican (least McCain share of the two-party vote) to most Republican (greatest McCain share of the two-party vote) and plot the vote shares in grey. Hence, in the figure, each district has 200 grey dots, one from each of the simulations. The 200 grey dots for the 1st district represent McCain's vote share in his least favorable district, whereas the 200 grey dots for the 27th district represent McCain's vote share in his most favorable district. In total, this gives us a visualization of the distribution of McCain votes across the simulated districts.

We can then compare these simulations to the McCain vote share within the actual district boundaries enacted by Florida in 2010. These vote shares are calculated by aggregating all recorded votes for McCain and Obama across all the polygons contained within the actual boundaries (as identified by a GIS map of the current congressional districts provided by the Census). They are ordered from least to most and are plotted in red.

[3] The simulations tend to converge quickly, such that 100 simulations tend to produce a similar distribution of district outcomes as 200 simulations.

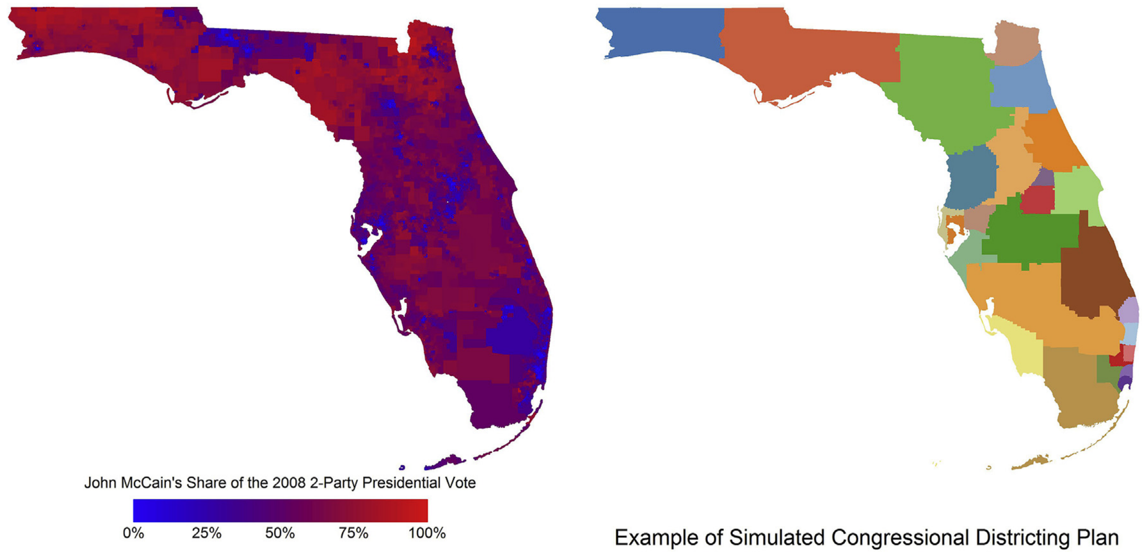

Fig. 1. Example of simulated districting plan in Florida (27 districts).

One benefit of the graphic is that it helps us to visualize the distribution of partisanship across Florida's districts. Although Florida is a battleground state where McCain lost by only a slim margin in the 2008 presidential election, we can immediately see that there is notable variation in the presidential vote share across the actual congressional districts (the red dots). For example, there are districts where McCain is so outnumbered by Obama supporters that he fails to win even 20% of the vote. Whereas, there are other districts where Republican support is strong enough that McCain received substantial voting majorities. In addition to the variance in vote-share, the district-level distribution is skewed such that the median district slightly favors McCain while the statewide average slightly favors Obama. In fact, McCain managed to win a majority of the vote in well over a majority of the districts, overtaking Obama in 17 of the total 27 constituencies.

From this distribution of presidential votes across districts, we can make inferences about the resulting partisanship of Florida's congressional districts. Yet, we can be more precise about how presidential votes translate into congressional outcomes by using congressional election data to inform our predictions. We do this by performing a simple logit transformation, where a binary indicator for whether a congressional seat was won by a Republican is regressed on McCain's share of the two-party vote for that district. We estimate the model by matching the electoral outcomes from the 2006, 2008, 2010 and 2012 congressional elections across every district in every state to the McCain share of the two-party vote contained in the district. As a result, the ith district's McCain vote share is transformed into the likelihood that a Republican wins the congressional election in that district using the following estimated model:

$$\Pr(district_i = Republican) = \text{logit}^{-1}(\beta_0 + \beta_1 McCainVoteShare_i)$$

This transformed data is then plotted in the second plot of Fig. 2. Here, the y-axis is simply the predicted probability that the elected Congressman from a given district is Republican. The districts are then aligned along the x-axis by the magnitude of the probabilities, from the least likely to elect a Republican to the most likely to elect a Republican.

We can see from this district-level distribution of partisanship that Florida's congressional delegations might produce Republican majorities even if it might lack a majority of support for Republicans statewide. It is simply the case that the current districts in Florida divide the partisan vote in a way that returns more seats per vote for Republicans than for Democrats. The median district, for example, is more likely to elect a Republican delegate than a Democratic delegate, whereas the state, at-large, would be less likely to do so. This is because Democratic supporters are not distributed across the districts in a way that most efficiently transforms votes into seats. Instead, large coalitions of Democratic support are contained in a small number of districts where the additional support has almost no impact on increasing the already strong odds that a Democrat wins the seat. As a result, valuable support that could be used to swing marginal districts in the favor of Democrats is lost in districts that are already overwhelmingly Democratic. Therefore, Republicans see favorable returns in seat share by maintaining a slight advantage in marginal districts, an advantage that they might not have received had the Democratic voters been distributed across the districts in a more efficient manner.

Yet, is this inefficient distribution of Democrats a result of gerrymandering? To properly assess whether, and to what extent, gerrymandering is the cause of this inefficiency we must be able to identify whether, and to what extent, this observed distribution is different from the set of potential distributions that are likely to result in the absence of gerrymandering. By replicating the districting process using a non-partisan procedure, the simulations allow us to do just this.

We can compare the actual district-level results against the simulated results to identify the effect of gerrymandering. If the simulations are unable to replicate the actual results, then the districts we observe are likely to have been designed with partisan intent and the difference between the actual and simulated results would be attributable to gerrymandering. However, if the simulations do replicate the actual results, then this would imply that the districts we observe are indistinguishable from districts that are designed without partisan intent. Any claim that the Florida districts are gerrymandered would be unsubstantiated.

Fig. 2 allows us to visualize the similarities and differences between the actual results (red dots) and the simulated results (grey dots). In some ways, the two distributions are similar. For example, they both produce significant variance in partisanship across the

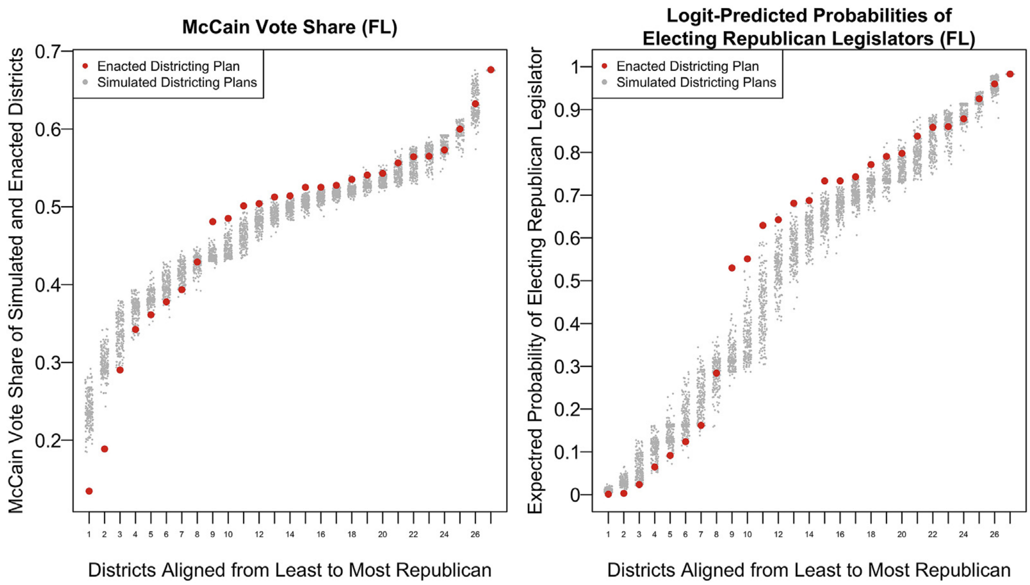

Fig. 2. The effect of gerrymandering on Florida's congressional delegation.

districts, where some districts are strongly Democratic and others are strongly Republican. Moreover, the actual median district and the average simulated median district are both marginally Republican despite marginal statewide support for Democrats. While it is commonly assumed that gerrymandering is responsible for parties winning a majority of the seats without a majority of the votes, the simulations suggest that the majority Republican delegation that we observe in Florida is one we should expect to observe even in absence of gerrymandering. Such an outcome is no different from a typical outcome produced by a non-partisan procedure.

However, there are some critical differences between the actual and simulated districts in Florida that indicate that the districts were gerrymandered. For example, there is greater variance in partisanship across the actual districts than across the simulated districts. The seven most Democratic districts are more Democratic than their simulated counterparts, while the nineteen least Democratic districts are less Democratic than their simulated counterparts. This is a consequence of a major jump in Republican support between the 7th and 9th most Democratic districts. It is a jump that occurs in the actual congressional districts but not in the simulations. As a result, the actual plans pack Democratic support into a few districts while the simulated plans distribute the support more uniformly across the districts. Consequently, the actual districts improve the Democrats' chances of winning a minor set of safe seats but reduce their chances of winning the remaining, more competitive seats.

This difference in the district-level distribution of partisan support between the actual and simulated plans has an important consequence for Florida's congressional delegation. By summing across the likelihoods that a district will elect a Republican (the area under of curve in the second plot of Fig. 2), we can compute the expected total number of Republican delegates that will emerge from each distribution. In Florida, for example, although every simulated plan produces a congressional delegation with a Republican majority, the actual plan produces a larger Republican majority than every simulation. In other words, given the current partisan conditions in Florida, we would expect a non-partisan districting procedure to produce a majority Republican delegation, but we would not expect such a procedure to produce the total number of Republican delegates that it actually does. Since a non-partisan districting procedure is unable to explain the current partisan outcome to its full extent in Florida, we can reject the hypothesis that the districts were drawn without partisan bias. Instead, as the alternative hypothesis suggests, it is likely that the additional seats were produced through gerrymandering.

5. State-level results

We apply this same procedure across 42 multi-district states and plot the results in Fig. 3. We exclude those states that have only one district (Alaska, Delaware, Montana, North Dakota, South Dakota, Vermont, and Wyoming), as well as Oregon, for which we do not have precinct level data. The figure gives a visual of the expected partisanship of each state's congressional delegation, ordering the states by the percent of the total congressional districts that we would expect to elect a Republican. To achieve this measure, we follow the same basic steps we took in Florida: aggregating the McCain-Obama vote for every district (both actual and simulated) using 2008 precinct-level election returns and then converting the vote into the likelihood that a Republican member of Congress is elected in that district. For each state, we calculate the sum of these likelihoods to find the expected number of Republican delegates produced by each districting plan and plot them in Fig. 3 as a percentage of the total number of seats.

In the figure, a red X indicates the expected Republican seat share produced by the actual post-2010 districting plan in each state. The grey dots represent the expected Republican seat share produced by each of the 200 simulated plans (they are jittered vertically to improve visualization). The values range from 0% to 100% along the x-axis and the states are ordered vertically along the y-axis according to the Republican seat share produced by the actual plan. Those states toward the top of the axis are those where Republicans win the greatest share of the total seats while those states toward the bottom of the axis are those where Republicans win the smallest share.

For each state, the simulated results (the grey dots) attempt to estimate the true distribution of the expected seat share won by Republicans in the absence of gerrymandering. Comparing these

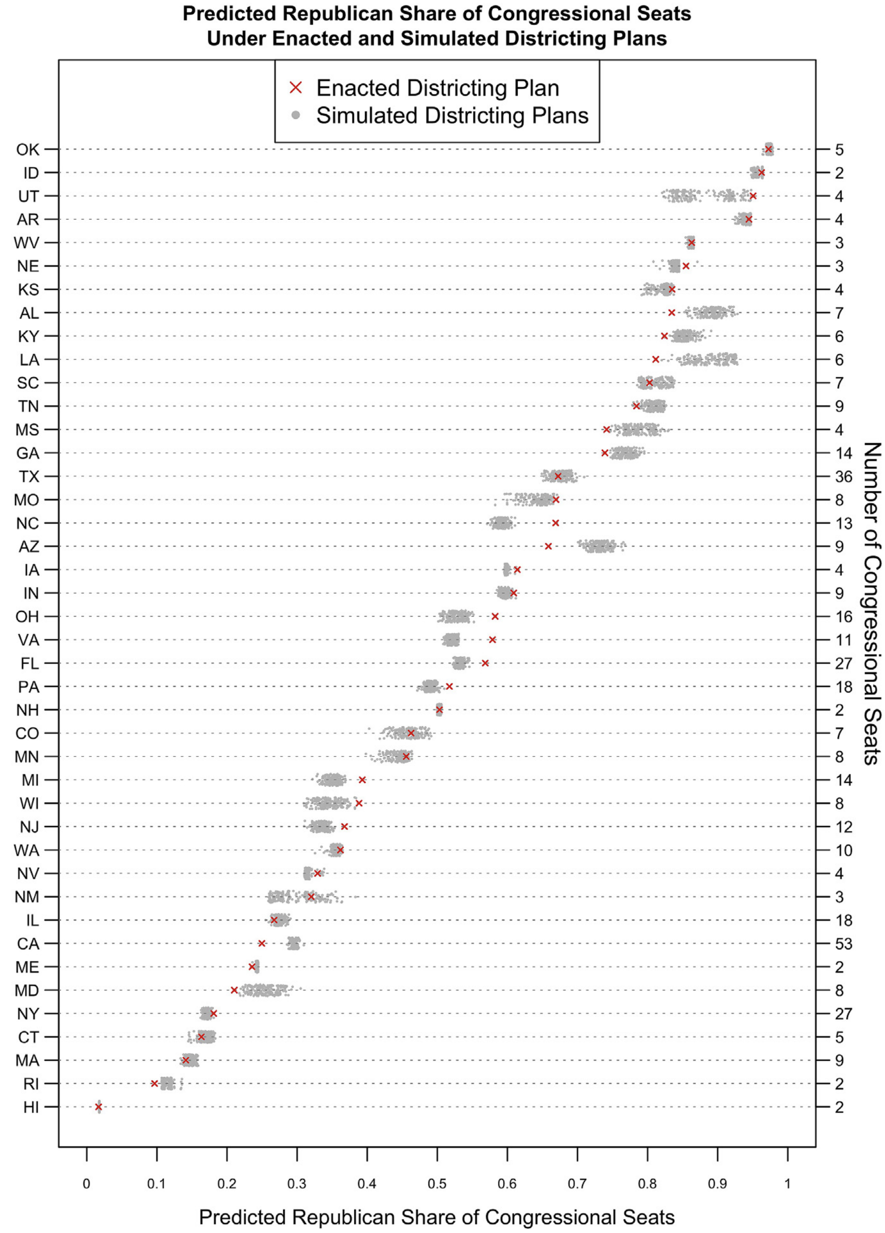

Fig. 3. Gerrymandering across the states.

simulated seat shares to the actual, observed seat share allows us to make inferences about the likelihood that a state's current districting plan is the result of non-partisan districting. The more the actual outcome deviates from the set of simulated outcomes the more confidence we can have that the current plan was influenced by partisan motivations.

Looking at Fig. 3, we can immediately see that the simulations do a pretty good job at predicting the expected Republican seat share generated by the enacted plans in each state. For any given state, the difference between the expected seat share from the enacted plan and that from the median simulation is no more that 9% and, in half of the states, the median simulation is within 2% of the actual expected seat share. Hence, where gerrymandering is present, the expected partisan gain in seat share that it generates is often relatively small. Although many of these states are in the outer tails of the simulated distribution, suggesting that the districting process was, at least in part, influenced by partisan consideration, the national aggregate effect size of partisan gerrymandering is relatively small.

Moreover, for states that fall well within the simulated distribution, we are unable to claim that the actual results are any different from that which would occur naturally, through non-partisan districting. In these states there is very little distinction between the actual results and the simulated results and, therefore, the evidence simply does not support the assertion that the enacted districting plans in these states are a product of gerrymandering.

However, there are a number of states where the enacted plans produce an expected Republican seat share that does deviate significantly from the simulated plans. In some of these states, the deviation favors Republicans, while in others it favors Democrats. In Wisconsin, for example, the enacted plan is about 4% (or a third of a seat) more Republican than the median simulation in the state, whereas, in California, the enacted plan is about 4% (or just over 2 seats) less Republican than the median simulation in the state. In both cases, the actual expected share of the seats won by Republicans is well outside the distribution of expected seat shares predicted by the simulations. As a result, we can confidently claim that the actual districting plan was not generated by the nonpartisan districting procedure used in the simulations. Rather, assuming that the districts are drawn to be reasonably compact, contiguous, and equally apportioned (as they are in the simulations), the best explanation for the partisan difference in expected seat share is that the districts were designed in a way that intentionally advantages one party over the other. In other words, the partisan difference can be explained by gerrymandering.

In addition to Wisconsin and California, there is noticeable variation across the states in the direction and degree to which the actual plans differ from the simulations. A partial explanation for this variation is that it depends on the partisan control of the districting process. Under this explanation, we should be more likely to observe districting plans that are designed to benefit a particular party when that party is unobstructed in implementing the plan of their choice. For example, in states where the legislature is responsible for approving district boundaries, we would expect to see a map that favors a particular party when that party controls both chambers of the legislature as opposed to when it does not.

As a quick test to see whether this explanation corresponds with our measure of gerrymandering, we divide the states according to who had decisive control over the redistricting process for the 2011–12 congressional redistricting cycle. To make this determination, we follow Justin Levitt's survey of redistricting institutions, in which he identifies the party - if any - that was in control of the redistricting process for each of the states.[4] Then, in Fig. 4, we plot the difference in Republican seat share between the enacted districts and the simulated districts according to whether Republicans or Democrats control the districting process.

In addition, we single out pre-clearance states. Pre-clearance states are required by Section 5 of the Voting Rights Act (VRA) to seek approval from the Department of Justice for any modifications to their election laws.[5] Under the VRA, the Department of Justice is responsible for protecting minority representation in these states by blocking any proposal that might have the effect of diluting minority votes. Hence, congressional districts in these states are intentionally designed to avoid splitting blocs of minority voters in a way that might reduce their representation. As a result, there exists a number of majority-minority districts that might not have existed otherwise had the districts been drawn at random.

Because of this, the resulting partisanship of the actual districts in these states will likely deviate from the partisanship of our simulated districts. Since minority groups protected under the VRA tend to vote overwhelmingly Democratic, districts produced by racial gerrymandering will likely be as distinct from our simulations as those produced by partisan gerrymandering. Therefore, in these pre-clearance states, our measure of partisan gerrymandering may actually be capturing the partisan effect of racial gerrymandering instead.

Fig. 4 displays all four categories: states where Republicans control the districting process, states where Democrats control the districting process, states with split partisan control of the districting process, and states that are required to obtain pre-clearance under the VRA.[6] As we might expect, the difference between the actual and simulated Republican seat share corresponds with these categories. For example, among partisan controlled states, the direction of this difference can be partly explained by which party is in control of the districting process. The difference favors Republicans in 10 of the 12 states where Republicans have control and the difference favors Democrats 4 of the 6 states where Democrats have control. However, the average effect is much stronger for Republicans than it is for Democrats. This result suggests that although both parties may use control over the districting process to their advantage, Republicans produce greater gains than Democrats.

Interestingly, however, there is evidence that Democrats are able to reclaim some of this Republican advantage through the race-conscious districting that occurs in pre-clearance states. Democrats do not control a single pre-clearance state and, therefore, we would not expect Democrats to gain an advantage through partisan gerrymandering here. Yet, five of the eight states are significantly more Democratic than they would have been had their districts been drawn without respect to race or partisanship. Even Louisiana, Alabama and Georgia, where districting was completely controlled by Republicans in the 2011–2012 cycle, produced districting plans that were more Democratic than the non-gerrymandered counterfactual plans. Racial consideration in the districting process, such as maintaining minority-majority districts, seems to produce stronger Democratic delegations in these states than they might otherwise produce without such considerations. The result suggests that the standards imposed by the Voting Rights Act may actually have the effect of not only improving minority representation but Democratic representation as well.

6. Total additional partisan seats due to gerrymandering

In our state-by-state analysis, we can see that both Republicans and Democrats make partisan gains through gerrymandering. Yet, what is the aggregate partisan effect of this gerrymandering on Congress as a whole? Running 200 simulated elections across the congressional districts of every state, we calculate the expected total number of Republicans that are elected for each simulation. Every simulated total can then be compared to the number of Republicans that the actual congressional boundaries of the 113th Congress are predicted to produce. Fig. 5 displays the results for 430 congressional districts (we exclude Oregon's 5 districts due to the lack of precinct-level data in the state).

Under the actual enacted plans, Republicans and Democrats are predicted to nearly split the House. Republicans take 217.3 of the 430 seats in expectation. Although this is a larger Republican delegation than 91% of the simulations, it is only slightly larger in degree. In fact, the actual expected Republican seat share is within the range of the 95% confidence intervals as indicated by the grey vertical lines. The simulated counterfactual produces an average of 216.4 Republican seats, which is approximately one fewer

[4] Justin Levitt's coding can be found at http://redistricting.lls.edu/who-partyfed.php.

[5] These are states subject to pre-clearance prior to the ruling in *Shelby County v. Holder*, 133 S. Ct. 2612 (2013).

[6] Although California and New Jersey both enlist independent commissions to approve their maps and, thus, are coded as being nonpartisan in Levitt's coding, we include them as partisan gerrymanders in the figure. Because many have suggested that California's Independent Citizen's Commission was heavily influenced by the Democrats, we have coded the state's districting process as being controlled by the Democratic party (Cain, 2012; Pierce and Larson, 2011). And because a Republican (Former New Jersey Attorney General, John Farmer, Jr.) was appointed to be the tie-breaking vote for New Jersey's independent commission, we have coded the state's districting process as being controlled by Republicans. These states are indicated by an asterisk in Fig. 4.

[7] Shapefiles of precinct-level McCain-Obama votes were collected from the Harvard Election Data Archive available at http://projects.iq.harvard.edu/eda/data.

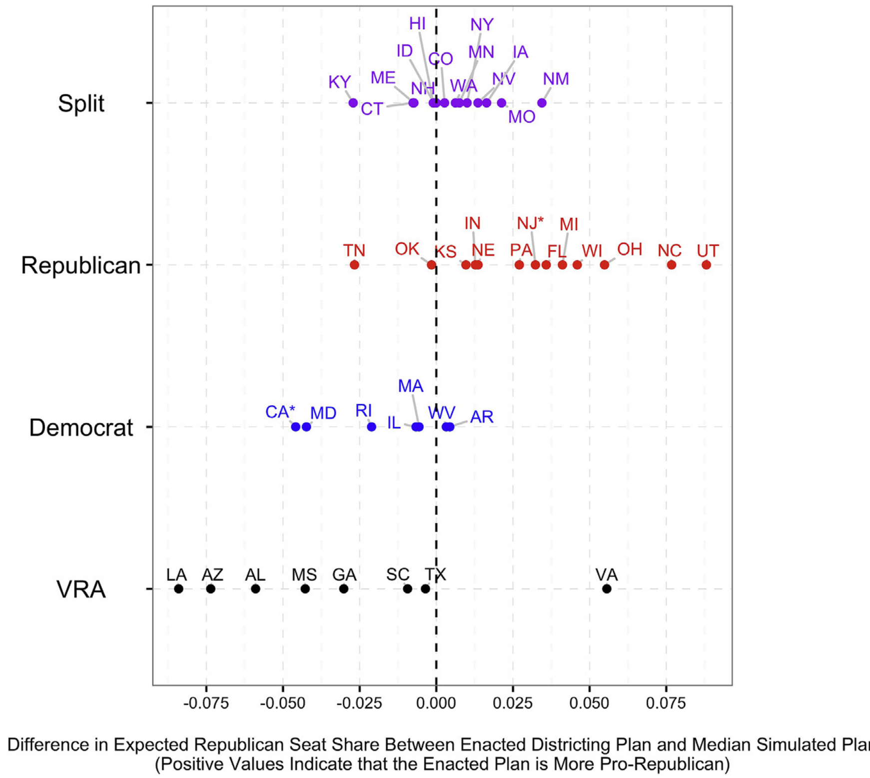

Fig. 4. Gerrymandering and political control over districting.

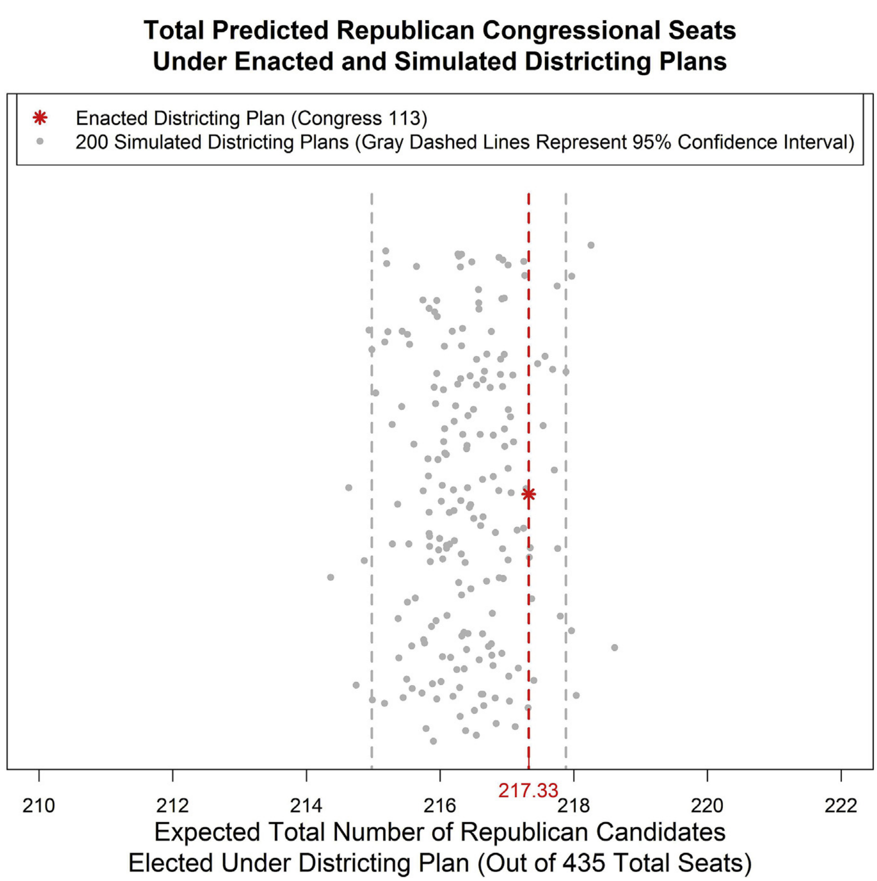

Fig. 5. The effect of gerrymandering on the total number of Republican seats in Congress.

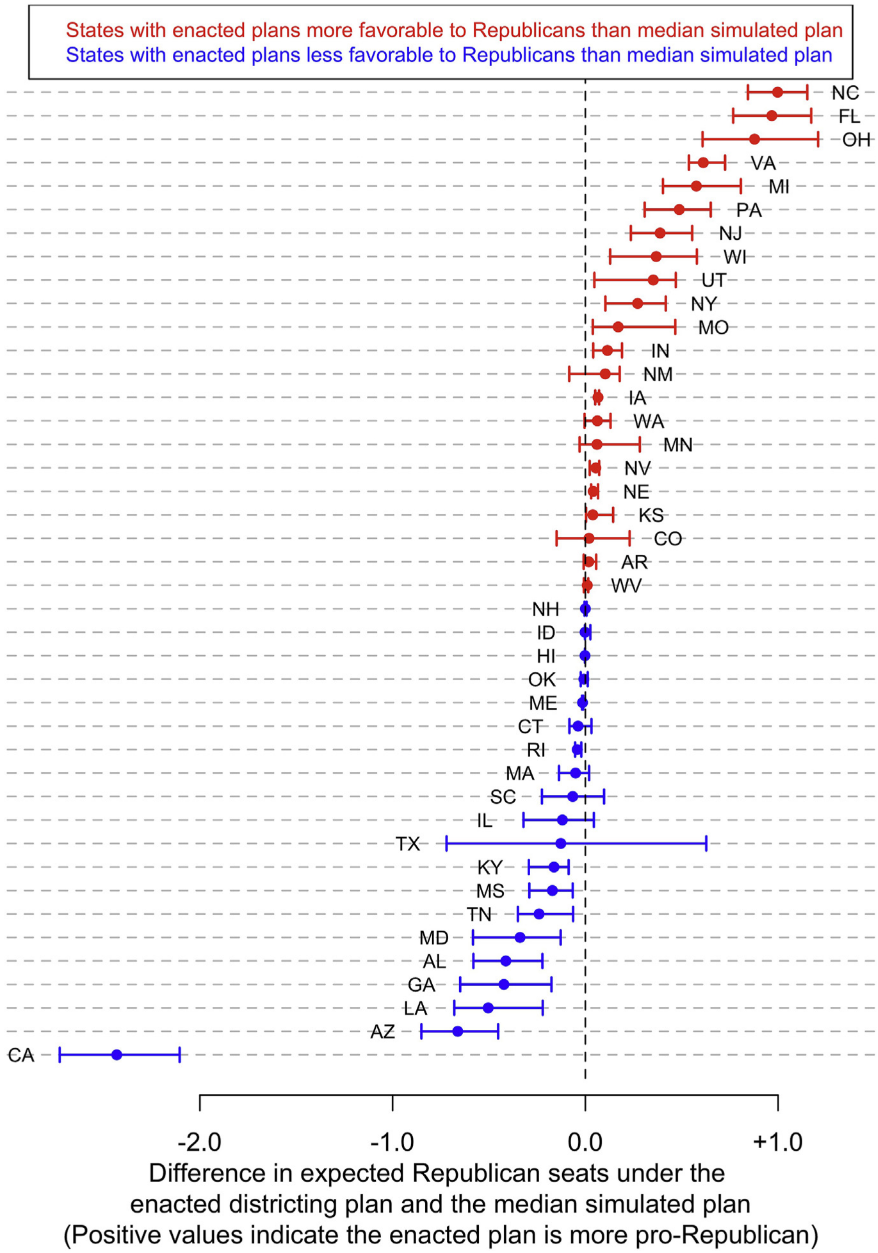

Fig. 6. The legislative consequences of gerrymandering by state.

Republican seat than the enacted plan is predicted to produce. This suggests that if districts were drawn randomly with respect to partisanship and race, Republicans would only expect to lose a single seat in Congress to the Democrats. Therefore, although we identify the partisan gains from gerrymandering in a number of states, these gains tend to be small and generally cancel out in the aggregate.

Such a result is meaningful because it contradicts a common perception that major partisan gains in seat share are made in the House of Representatives through gerrymandering. Instead, the evidence suggests that the partisan makeup of the House would be almost no different if gerrymandering - both partisan and racial - were altogether eliminated. Although some state delegations would see significant change, the aggregate advantage received by a particular party in Congress would be almost zero.

Fig. 6 disaggregates the effect of gerrymandering by state. It displays the magnitude of the expected partisan seat gain (or seat loss) that is due to gerrymandering in each state. It is calculated as the difference between the expected Republican seat share of the actual districting plans and the median Republican seat share across the state's 200 simulations. First, we can see that the state-level effect is limited. For any given state other than California, gerrymandering causes no more than one seat to switch parties. In most of these individual states, the partisan effect of gerrymandering is negligible.

The greatest seat gain for Republicans occurs in states where partisans control the districting process. In fact, eight of the nine largest gains for Republicans were in states where Republicans were in control. On the other hand, most of the seat gain for Democrats occurs either in California or pre-clearance states. California contributes as many seats to the Democrats through gerrymandering as the top three Republican gerrymanders

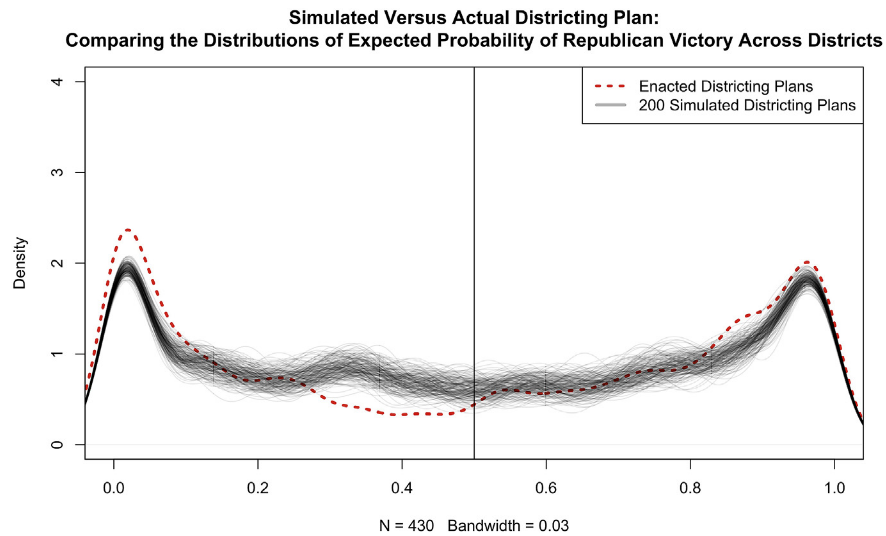

Fig. 7. The distibution of the likelihood of a Republican victory across Congressional districts: Simulated districts vs actual districts.

combined. Democrat-controlled Maryland and Illinois are also among the additional ten largest gains for Democrats. Yet, remarkably, six of the eight others are pre-clearance states. The pre-clearance states act to bolster Democratic representation, swapping Republican seats for Democrats. Without these pre-clearance states, Republicans would experience a slightly larger bias in aggregate seat share in Congress. We can aggregate the total number of seats gained by gerrymandering across the four regime types mentioned in the previous section. This gives us a sense of how the regimes contibute to the net effect of gerrymandering in Congress. We find that among states controlled by Republicans, about five Republican seats are gained through gerrymandering. And among states controlled by Democrats, about three Republican seats are lost by Gerrymandering. Moreover, Democrats gain about 1.75 seats from states subject to preclearance. This suggests that the Republican seat gain from Republican controlled states is counterbalanced by the seat loss in Democratic controlled and pre-clearance states.

We can get a clearer picture of the how the full set of simulated electoral outcomes differ from the actual outcomes by comparing the distributions. Fig. 7 plots each of the 200 simulated densities of the 423 congressional districts belonging to the 42 multimember states (minus Oregon) on which we ran the simulations against the density of the actual congressional districts among those states. It is immediately clear from the distribution of the actual districts that the most liberal districts tend to be outliers. In other words, there is a noticeable bump toward the left of the distribution reflecting the collection of the most liberal districts that overwhelmingly support Democrats.

These districts are usually districts with large metropolitan areas where Democrats tend to concentrate. These packed liberal districts waste Democratic votes and thus allow Republicans to efficiently win a larger number of moderate districts by relatively small margins. Where the simulated distributions tend to have high numbers of marginally Democratic districts relative to the numbers of marginally Republican districts, the actual distribution has high numbers of marginally Republican districts relative to the numbers of marginally Democratic districts. Therefore, liberal districts are getting far more Democrats than they should be getting under a random draw. They are grossly overrepresented in a few districts while the remaining districts get a thin supply of the residual Republicans. This is enough to swing marginal Democratic districts in the favor of Republicans. The effect of gerrymandering, therefore, is clear: Democrats gain safety in some of their most liberal districts, which, in effect, reduce their chances of winning more marginal districts. The Democrats lose seats in Congress as a result.

7. Conclusion

We have argued that there is a need to better identify when a state engages in gerrymandering, so as to not confuse partisan gerrymandering with the type of partisan bias that results as an unintended consequence of drawing districts around a geographically-concentrated, partisan population. To avoid such confusion, we need to establish a baseline for what a state's districts would look like if they had not been intentionally gerrymandered. Using such a baseline, we could then compare the actual electoral districting plans to a non-gerrymandered hypothetical. If the plans differ significantly from the baseline, then it is likely that the actual plans were gerrymandered.

In this paper, we have established such a baseline. We use computer simulated districting plans that mimic the districting decisions performed by the states. The simulations draw compact, contiguous and equally apportioned districts around the unique geographic distribution of partisans in each state. The only difference between the simulated and actual districts is that the simulated districts are not subject to intentional partisan or racial gerrymandering. Thus, any difference between the computer-produced plans and the actual plans should reliably be due to gerrymandering.

By comparing the enacted post-2012 congressional districting plans to their simulated counterfactual, we are able to more precisely identify the effect of gerrymandering on partisan seat share for each state's congressional delegation. While we find that gerrymandering does play a role in altering electoral outcomes across the states, in most states gerrymandering has little to no effect on the partisan outcome of congressional elections. Instead, most of the outcomes can easily be explained by an unbiased districting procedure, where gerrymandering has been completely removed from the districting process.

Moreover, we find that in states where gerrymandering does have a significant effect on congressional elections, the effect is relatively small. For example, other than in California, the partisan gain from gerrymandering amounts to no more than a fraction of a seat in any given state. While the total number of seats gained by Republicans is greater than the total number of seats gained by Democrats, the net effect of gerrymandering in Congress is only marginal. In fact, we find that Republicans are expected to net no more than one additional seat as result of it.

However, the type of gerrymandering that benefits Republicans appears to be different from the type of gerrymandering that benefits Democrats. For example. Republicans appear to make most of their gains as a result of having full institutional control over the districting procedures in their states. This allows them to etch partisan bias into their districts without any effective opposition. Our analysis reveals that Republicans tend to take advantage of these opportunities.

On the other hand, because there are fewer states where Democrats have institutional control over the districting process, Democrats gain fewer seats from partisan gerrymandering than Republicans. In fact, counterintuitively, we observe that much of the significant state-level bias toward Democrats occurs in states where Republicans hold legislative majorities. It turns out that these states are also pre-clearance states that engage in racial gerrymandering to meet the standards set out by the Voting Rights Act. Therefore, Democrats tend to gain seats as a result of racial bias instead of partisan bias.

If we were to remove the pre-clearance states from the analysis, then we would observe that gerrymandering tends to favor Republicans. However, the additional seats that Republicans would gain on average from gerrymandering - even without pre-clearance states - is marginal. It is simply the case that the effect of gerrymandering is small and that removing bias from districting process - whether it is racial or partisan - is not likely to change the partisan composition of Congress.

References

Abramowitz, Alan I., 1983. Partisan redistricting and the 1982 congressional elections. J. Polit. 45 (03), 767–770.
Abramowitz, Alan I., Alexander, Brad, Gunning, Matthew, 2006. Incumbency, redistricting, and the decline of competition in US House elections. J. Polit. 68 (1), 75–88.
Ansolabehere, Stephen, Rodden, Jonathan, Snyder, James M., 2006. Purple America. J. Econ. Perspect. 20 (2), 97–118.
Cain, Bruce E., 2012. Essay and feature: redistricting commissions: a better political buffer? Yale LJ 121, 1808–2405.
Campagna, Janet, Grofman, Bernard, 1990. Party control and partisan bias in 1980s congressional redistricting. J. Polit. 52 (04), 1242–1157.
Chen, Jowei, Rodden, Jonathan, 2013. Unintentional gerrymandering: political geography and electoral bias in legislatures. Q. J. Political Sci. 8 (3), 239–269.
Cox, Gary W., Katz, Jonathan N., 1999. The reapportionment revolution and bias in US congressional elections. Am. J. Political Sci. 812–841.
Cox, Gary W., Katz, Jonathan N., 2002. Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution. Cambridge University Press.
Erikson, Robert S., 1972. Malapportionment, gerrymandering, and party fortunes in congressional elections. Am. Political Sci. Rev. 1234–1245.
Erikson, Robert S., 2002. Sources of partisan bias in US congressional elections: an update stimulated by Ron Johnston's essay. Polit. Geogr. 21 (1), 49–54.
Ferejohn, John A., 1977. On the decline of competition in congressional elections. Am. Political Sci. Rev. 166–176.
Gelman, Andrew, King, Gary, 1994. Enhancing democracy through legislative redistricting. Am. Political Sci. Rev. 88 (03), 541–559.
Glazer, Amihai, Grofman, Bernard, Robbins, Marc, 1987. Partisan and incumbency effects of 1970s congressional redistricting. Am. J. Political Sci. 680–707.
Goedert, Nicholas, 2014. Gerrymandering or geography? How Democrats won the popular vote but lost the Congress in 2012. Res. Polit. 1 (1), 2053168014528683.
Herron, Michael C., Wiseman, Alan E., 2008. Gerrymanders and theories of law making: a study of legislative redistricting in Illinois. J. Polit. 70 (01), 151–167.
Jacobson, Gary C., 2003. Terror, terrain, and turnout: explaining the 2002 midterm elections. Political Sci. Q. 118 (1), 1–22.
Kendall, Maurice G., Stuart, Alan, 1950. The law of the cubic proportion in election results. Br. J. Sociol. 1 (3), 183–196.
King, Gary, 1989. Representation through legislative redistricting: a stochastic model. Am. J. Political Sci. 787–824.
Masket, Seth E., Winburn, Jonathan, Wright, Gerald C., 2012. The gerrymanderers are coming! Legislative redistricting won't affect competition or polarization much, no matter who does it. PS Political Sci. Polit. 45 (01), 39–43.
McDonald, Michael P., 2004. A comparative analysis of redistricting institutions in the United States, 2001–02. State Polit. Policy Q. 4 (4), 371–395.
McDonald, Michael, 2009. Seats to Votes Ratios in the United States (Unpublished Paper).
Openshaw, Stan, 1984. The Modifiable Areal Unit Problem. Geobooks, Norwich, England.
Pierce, Olga, Larson, Jeff, 2011. How Democrats Fooled California's Redistricting Commission. http://www.propublica.org/article/how-democrats-fooled-californias-redistricting-commission.
Squire, Peverill, 1985. Results of partisan redistricting in seven US states during the 1970s. Legis. Stud. Q. 259–266.
Tufte, Edward R., 1973. The relationship between seats and votes in two-party systems. Am. Political Sci. Rev. 540–554.