Who is my Neighbor? The Spatial Efficiency of Partisanship

Nicholas Eubank*, Jonathan Rodden†

August 23, 2017

Abstract

Relative to its overall statewide support, the Republican Party has been over-represented in Congressional delegations and state legislatures over the last decade in a number of U.S. states. A challenge for courts is to determine the extent to which this can be explained by intentional gerrymandering vis-a-vis an underlying inefficient distribution of Democrats in cities. We explain the problem of “spatial inefficiency” in partisan support, and measure it by borrowing from the field of plant ecology, assessing the partisanship of the nearest neighbors of each voter in each U.S. state at the spatial scales relevant for Congressional delegations and both chambers of state legislatures. We demonstrate that as a result of urban-rural partisan polarization, much of the cross-state and cross-chamber variation in Republican advantage can be explained by the relative spatial inefficiency of Democrats. Moreover, our “pure” political geography approach to votes and seats provides a useful baseline against which to evaluate claims of partisan gerrymandering. We demonstrate that Republicans are often able to improve significantly on their underlying geographic advantage when they control the redistricting process, while Democrats are sometimes able to ameliorate it when they draw the lines.

*Post-Doctoral Fellow, Spatial Social Science Lab, Stanford University nick@nickeubank.com
†Professor, Department of Political Science and Senior Fellow, Hoover Institution, Stanford University jrodden@stanford.edu


Introduction

Since 2000, summing over all statewide elections for the United States Senate and all presidential contests, Republican candidates have received less than 48 percent of all votes cast. Yet during that same time period, they have received more than 52 percent of all seats contested in U.S. Congressional elections, 55 percent of all state lower chamber seats, and 56 percent of all seats in state upper chambers.

What accounts for this underrepresentation of Democrats in legislatures? On the one hand, it is clear that Republicans have controlled the redistricting process in several large states, where they made concerted efforts to draw favorable districts (McGann et al. 2016). And indeed, Democrats face an inefficient distribution of support across districts in those states, and hence suffer in the transformation of votes to seats (Royden and Li 2017; Stephanopoulos and McGhee 2015).

On the other hand, an inefficient geographic support distribution for a party of the left is nothing new in an industrialized democracy. Votes for left parties have been concentrated in densely populated districts in democracies using single-member districts ever since the industrial revolution and the rise of a geographically concentrated urban working class (Gudgin and Taylor 1979). In the contemporary United States, the growing correlation between population density and voting behavior has generated an inefficient clustering of Democrats in dense urban centers that is reminiscent of workers’ parties in the early 20th century (Rodden 2017). In some states, Democrats are sufficiently clustered that even a party-blind redistricting process would likely lead to the over-representation of Republicans (Chen and Rodden 2013).
It is difficult to disentangle these explanations, since the sharp rise in urban-rural political segmentation in recent decades coincided with the rise of Republican control over the redistricting process and the spread of access to sophisticated redistricting software. It is also perhaps not a coincidence that Republicans won both chambers of the legislature and gained control over the redistricting process in exactly those hotly contested states where their underlying geographic support was most efficient relative to that of Democrats. As plaintiffs attempt to convince state and federal courts to rule that partisan gerrymanders are unconstitutional in these same states, they must find a way to distinguish between these two explanations. When confronted with evidence of an unusual level of bias in the transformation of votes to seats, the easiest defense available to state legislatures is to lay the blame on the underlying spatial inefficiency of support for Democrats.

Yes these claims are ill-defined, and the nature of the Democrats’ spatial inefficiency across states has not yet been fully understood or measured. This paper attempts to make progress on both fronts. First, we argue that the spatial concentration of left voters in cities is a necessary but insufficient condition for under-representation in a system of winner-take-all districts. Spatial inefficiency emerges when cities are either too large or too small relative to the size of districts. In some states, like the New England states and Wisconsin, Democrats are dispersed relatively efficiently across medium-sized cities. In other states, including many of the large states of the Northeast and upper Midwest with a history of rail-based industrialization, Democrats are inefficiently clustered not only in large cities that they win with super-majorities, but also in smaller post-industrial outposts where they are outnumbered by surrounding Republicans. Moreover, in states that they lose by large margins, like Utah and Tennessee, concentrated support in cities can be a good thing for the
Democrats at the scale of Congressional districts, since it allows them to win districts in spite of low statewide support. More generally, the nature of the Democrats’ spatial inefficiency can change substantially as one moves between the spatial scale of state lower chambers, upper chambers, and the scale of the U.S. Congress.

The key contribution of this paper is to measure the relative spatial efficiency of the parties for each legislative chamber in 49 U.S. states without relying on districts drawn by politicians or computer algorithms. To achieve this, we borrow from spatial point pattern analysis that was developed in the field of plant ecology. We transform precinct-level data from the 2008 presidential election into a series of points representing voters, and calculate various quantities of interest for the nearest neighbors of each individual voter, where the relevant neighborhood corresponds to the spatial scale of state legislatures and the state’s Congressional delegation.

In general, we find that at any given level of statewide partisanship, Republicans tend to live in more “efficient” neighborhoods than Democrats. That is, the average Republican is more likely to live in a neighborhood with a small majority of co-partisans, while the average Democrat is more likely to live in an overwhelmingly Democratic neighborhood or a Republican neighborhood. Thus for any given level of state-wide partisanship, a larger share of all Republicans live in majority-Republican neighborhoods than is true for Democrats.

However, we also find that there is substantially heterogeneity in the level of relative spatial (in)efficiency across states, and within states across legislative chambers. The devil is in the details of the political geography of each state and legislative chamber. It is not the case that urban concentration creates a generic, uniform anti-Democratic bias in all situations.
Our measure of spatial efficiency can easily be transformed into an intuitive predictor of the seats won by parties in actual elections. Our approach treats every voter in the state as if she was at the center of her own district. For each state, we can thus calculate the share of voters living in a majority-Republican (nearest) neighborhood. We demonstrate that this is an excellent predictor of actual Republican seat shares in recent elections, and captures the extent to which the relative spatial efficiency of Republicans translates into a seat bonus.

We show that spatial efficiency is an especially good predictor of actual seats won by the parties in states where districts were drawn by independent commissions, courts, or with split partisan control. On the other hand, when districts were drawn by Republicans, observed Republican seat shares are often significantly higher than what we would anticipate based purely on the relative efficiency of Republicans’ support. Likewise, Democratic legislatures are able to increase Democratic seat shares beyond the geographic baseline established through the nearest neighbor analysis.

In other words, without drawing any districts, we are able to account for the underlying spatial efficiency of the parties’ support and isolate the advantage obtained through partisan control of the redistricting process. In states like Wisconsin, Pennsylvania, Michigan, and North Carolina, it is clear that unusually large Republican seat shares cannot be explained merely by the underlying efficiency of geographic support for Republicans. Thus we have provided a tool that can be useful to federal and state courts, who must disentangle the impact of partisan geography and intentional under-representation of a political party when evaluating claims of unconstitutional partisan gerrymandering. In this context, our approach provides a baseline against which to contrast a challenged redistricting plan. This approach is a complement to the use of automated districting simulations to generate such a baseline,
as proposed by Chen and Rodden (2015), Chen (2017), and Cho and Liu (2016).

**The Puzzle: Partisanship and Seats across States**

The United States is perhaps the purest two-party system in the world. Without the complications created by the national and regional small parties found in Canada and the UK, one might expect the transformation of votes to seats for the two major parties in the United States to come close to the standard “cube law” identified by Kendall and Stuart (1950). However, this is manifestly not the case. Figures 1 - 3 plot the share of seats won by Republicans during the period from 2008 to 2012 on the vertical axis, and the average statewide vote share of Republicans in presidential and senate elections during the same period on the horizontal axis. Figure 4 combines Figures 1 - 3 with lowess plots for each legislative chamber.

![State lower chambers](image)

**Figure 1: Partisanship and seats in lower state legislative chambers**
Figure 2: Partisanship and seats in upper state legislative chambers

Figure 3: Partisanship and seats in U.S. Congressional delegations
The last decade has been extremely competitive in American politics. Both the mean and median of average partisanship across states was .5 during this period, and the distribution of partisanship across states resembles a normal distribution. However, the distribution of partisanship within states is a different story. The cube law, portrayed in Figures 1 - 4 with a dashed line, is the expected translation of votes to seats when the distribution of partisanship across districts within states approximates a normal distribution. However, as demonstrated in Rodden (2010), the distribution of Republican support across districts within U.S. states typically has a pronounced left skew, such that Democrats are clustered in urban districts. This has been true for decades in the states of the upper Midwest, Northeast, and West Coast, and the phenomenon has grown over time and spread to the states of the South and Mountain West in recent years (Rodden 2017).

Partially as a result, most states in Figures 1 - 4 do not obey the cube law. One ob-
ervation is that for state legislatures, the translation of partisanship to legislative seats is much closer to proportionality than one might expect in a system of winner-take-all districts. For instance, Figure 4 suggests that in a state where around 40 percent of the vote goes to Republican presidential and Senate candidates, Republicans can expect around 40 percent of the seats in the state legislature. However, there is also a clear asymmetry favoring the Republicans. When Republican partisanship reaches 45 percent, the Republicans can expect to receive half of the seats in the state legislature or the Congressional delegation. In states that were evenly divided during this period, Republicans could expect comfortable majorities in each legislative chamber, and at 60 percent, Republicans can expect to receive more than a proportional share of seats, though still not as many as would be predicted by the cube law.

These rather striking graphs provide the motivation for this paper. Why are the graphs so much flatter than the predictions of anything like the cube law? What accounts for the fact that the transformation of statewide political support to seats is so favorable to the Republicans? To what extent does the spatial efficiency of partisanship explain these patterns? And to what extent are they explained by partisan gerrymandering? These graphs also foreshadow our main findings. It is already apparent that the shape of the graphs cannot be explained by gerrymandering alone. For instance, note the over-performance of Republicans relative to the cube law— and sometimes even relative to proportionality— in Democratic states like Minnesota, Colorado, Illinois, and New Jersey, where Republicans did not control the most recent redistricting process. At the same time, the extent of Republican over-representation is especially pronounced in states like Wisconsin, Michigan, North Carolina, Virginia, and Pennsylvania, where Republican gerrymandering has been
most overt and unapologetic.

**Cities and the Spatial Inefficiency of the Left**

One of the enduring contributions of Gudgin and Taylor (1979) is the observation that in a two-party system, when there are distinct geographic clusters of votes for the two parties—for example working-class and professional neighborhoods within towns—and when the size of districts is larger than the homogeneous clusters, the distribution of partisanship across districts should approximate a normal distribution, and the votes-to-seats relationship tends toward something like the familiar “cube law” identified by Kendall and Stuart (1950).

In the contemporary United States, where the urban core of virtually every city is overwhelmingly Democratic, suburbs are heterogeneous, and exurbs and rural areas are Republican, the analogous scenario is one in which urban core areas are uniformly small relative to the size of Congressional districts. Let us examine a stylized example of a polity, portrayed in Figure 5, with only 48 voters, 24 of whom are predisposed toward the left party (L) and 24 of whom typically prefer the right party (R). Voters living close to one another in city centers typically vote for the L party, suburbs are evenly split, and those living further from one another in rural areas typically vote for the R party. There are 3 L voters in each city, and 3 R voters in each surrounding rural periphery.

![Figure 5: Hypothetical polity with small cities](image-url)
Let us imagine that this polity must be apportioned into six districts of equal size, the boundaries of which are portrayed with vertical bars. Each district contains 8 voters, and is hence larger than the scale of the L and R clusters. The two parties are evenly matched in four of the resulting districts, which inevitably contain urban, suburban, and rural areas, while the L party can expect a majority in one district that ends up slightly more urban, and the R party can anticipate a majority in one district that ends up slightly more rural. Thus the distribution of expected partisanship across districts is symmetric with a large peak in the middle.

Next, let us consider scenarios in which one of the parties suffers from a scandal or benefits from a strong economy and the election is not tied. To do so, we simulate 10,000 elections in which one voter is randomly selected to switch from L to R, then do the same for two voters, three voters, and so on until the R party wins all of the votes. We conduct the same exercise in the opposite direction. For each scenario, we calculate the average seat share across all simulations that would be produced by the districting scheme displayed in Figure 5. The resulting vote-seat curve is displayed with the green line in Figure 6.
Figure 6: Vote-seat curves for two hypothetical polities, six districts

The green line in Figure 6 is a standard majoritarian vote-seat curve that approximates the cube law. For comparison, the dotted line represents proportionality, and the dashed black line represents an at-large system in which the party with a majority of votes wins all of the seats. It can also be understood as the vote-seat relationship in a hypothetical situation where, unlike Figure 5, partisanship is perfectly uniformly distributed across geographic space, such that 51 percent of the polity-wide vote yields a victory in every single district.

Figures 5 and 6 indicate that the concentration of voters in cities does not necessarily present a problem for the left. In this example, cities are smaller than the size of districts, but not so small that the left party is consistently overwhelmed. Cities are also similar in size and spaced with a uniform distance between them. As we will see below, at the spatial
scale of Congressional districts, this type spatial pattern reflects the political geography of some U.S. states that lack large cities.

Next, let us examine a political geography that more closely resembles U.S. states with large cities, where instead of being clustered in a series of small agglomerations that do not reach the size of districts, there is at least one cluster that reaches or surpasses that size, along with other smaller agglomerations that do not. For instance, the urban core of New York City is larger than the size of a Congressional district, while those of Rochester and Buffalo fall well short. Likewise, urban Memphis approaches the size of a Congressional district, while Knoxville and Chattanooga do not.

Such a polity is portrayed in Figure 7, which has one large city that votes overwhelmingly for the L party, surrounded by heterogeneous suburbs and a rural periphery that vote for the R party. It also contains a smaller city and a small town, both surrounded by right-leaning suburbs and rural periphery.

![Figure 7: Hypothetical polity with a hierarchy of cities](image)

Again, let us examine what happens when this evenly divided polity is partitioned into six districts of equal size. These are captured with the solid orange lines in Figure 7. Because its supporters are inefficiently packed into the large city, the Left party wins only 42 percent of the seats in spite of winning half the votes. This is an example of the classic case of electoral bias owing to an inefficient geographic clustering described by Brookes (1960), Johnston
Following the approach described above, the orange line in Figure 6 derives the vote-seat curve for this example. Several features of this curve are noteworthy. First, the curve is flatter and closer to proportionality than the green curve. Because of the greater relative clustering of the L party, it is able to win more seats when it performs badly (on the right side of the graph) than would have been the case with a more even distribution of support across cities. It is able to win seats in its urban core support area even when it performs very badly overall.

Likewise, because the support for the Left party is so concentrated in cities, the Right party is able to string together suburban and rural voters in districts that it wins with slim majorities, and relative to the more even distribution of cities portrayed by the green line, or the perfectly uniform distribution of voters captured by the dashed line, it is able to win a seat bonus when it is in the minority, and in the case of a tied election. Of course the latter case carries considerable normative importance in democratic theory.

The orange curve displays an asymmetry that is not present in the green curve. In general, the Right party can expect a larger seat bonus than the Left party. Most notably, the Right party can expect a majority of seats with only 45 percent of the votes. Likewise, it achieves proportional representation with 40 percent of the votes, while the Left party can only expect 30 percent of the seats with a similar vote share.

Figure 6 builds on the intuition of Calvo and Rodden (2015): A two-party system with perfect geographic dispersion of partisanship, such that a party with 45 percent of the overall vote receives 45 percent in each district, is equivalent to an at-large system, in that the election winner receives all of the seats. At the other extreme, if parties’ support is perfectly
segmented such that each district is homogeneous, representation is perfectly proportional to the vote share. Thus the vote-seat curve becomes flatter as partisan support becomes more clustered.

The wrinkle here is that when population density and voting behavior are highly correlated and dense cities are sufficiently large, the urban party is more concentrated within districts than the rural party, generating a skew in the distribution of party support across districts, and an asymmetric flattening of the vote-seat curve.

**Measuring Spatial Efficiency**

As indicated in the simple examples above, there is no one-size-fits all indicator of partisan spatial efficiency in a U.S. state. The interaction of district size and city size are such that an arrangement of partisanship that is efficient for the urban party at one scale, e.g. one of the state legislative chambers, might be inefficient at another scale like Congress. For instance, small Democratic cities like Fargo are too small and isolated to yield a Congressional victory, but the scale of state legislative districts is such that Democrats can expect to win several Fargo seats in the North Dakota state legislature. Moreover, even Democratic cities that are quite small can generate Democratic victories at the scale of Congressional districts if they are sufficiently close together, as with New England mill towns, or strings of old industrial towns like Appleton, Neenah, Oshkosh, and Green Bay, Wisconsin. Much also depends on the size and structure of suburbs, and the increasingly polycentric form of some U.S. cities.

It is also the case that a spatial distribution of support that is inefficient for a majority party can be quite efficient for a minority party (and vice versa). For example, a party
with 70 percent of the vote would rather not have a large right tail to its cross-district support distribution, whereas a party with 30 percent of the vote is much better off with a right-skewed distribution.

Our goal is to devise a measure of the parties’ relative support efficiency at different spatial scales in a way that facilitates comparisons across states with very different levels of support for Democrats and Republicans. That is, we wish to understand the likelihood that a districting scheme would generate departures from a normal cross-district distribution of partisanship such that one party is likely to end up with a larger share of its voters in districts that it wins by large margins and loses by smaller margins, while the other party is more likely to have its voters in districts that it wins by small margins and loses by large margins. Our ultimate goal is to explore how these differences in spatial efficiency translate into representation.

We achieve this by calculating the partisanship of the $k$ nearest neighbors of each voter in the state, where $k$ corresponds to the size of a district in either the state’s lower chamber, upper chamber, or its Congressional delegation. Estimation of the partisan composition of each voter’s neighborhood is accomplished through a three-step process. First, precinct-level election returns from the 2008 Presidential election are used to estimate the spatial distribution of voters in each state. This is done by creating a number of representative voter points within each precinct, where points are positioned uniformly at random within each precinct’s catchment area, and the number of points in each precinct’s catchment area is proportional to the number of votes cast for each party.\footnote{In particular, the number of points in each precinct for each party is determined by a draw from a binomial distribution, where $n$ is the number of voters for each party in the precinct. The binomial probability $prob$ varies by state-chamber, but is always equal to $prob = \frac{\text{number of districts}}{\text{number of voters in state}} \times 1000$, a probability that generates 1000 voters per district in expectation.}
Estimation of the partisan composition of the neighborhood around each of these representative-voter points is then calculated. In the nearest neighbor analysis, for each representative-voter point \( v \) of a given party \( p \in \{ D, R \} \), the partisanship of the neighborhood around \( v_p \) is equal to the share of the \( N_{state,chamber} \) nearest points who are also from party \( p \). The number of nearest neighbors considered – \( N_{state,chamber} \) – is set to ensure the included points represent the number of voters in the average district in state for chamber chamber. \(^2\) This estimate is analogous to asking “if a circular electoral district of average district population were centered on this voter, what share of people in that district would be co-partisans.”

This analysis generates an estimate for each representative-voter point of the share of neighbors who are co-partisans. These point-level estimates can then be aggregated in various ways, including (a) averaging these estimates separately for Democratic and Republican voters, or (b) calculating the share of Democratic and Republican voters who reside in neighborhoods that are \( X\% \) co-partisan.

When one party has a more efficient support distribution than its competitor, a larger share of its voters will find themselves in winning neighborhoods when receiving the same vote share. When a party’s support distribution is inefficient, too many of its supporters live in neighborhoods that it either 1) loses or 2) wins by large super-majorities.

We classify each individual’s neighborhood as majority Republican if its vote share is higher than John McCain’s 2008 national two-party vote share, and Democratic if it is below. For the period between 2008 and 2016, this cut-off is quite close to reality. In Congressional races, Democratic victories have been quite rare in districts where McCain’s

\(^2\)To illustrate, consider a state-chamber with 3 districts and 300,000 voters. The average district is home to 100,000 voters, and so the number of points considered in the nearest neighbor analysis should represent 100,000 voters. Note that because of how \( prob \) is constructed, this will always amount to examining the share of the 1,000 points around each person who are co-partisans.
2008 vote share was higher than 46.3 percent, and Republican victories have been quite rare in districts where Obama’s vote share was higher than 53.7 percent.

The next series of graphs allows for a cross-state visualization of this notion of spatial efficiency. The vertical axis captures the share of each party’s voters living in neighborhoods where the party is a local majority, defined using the 2008 presidential vote share, and the horizontal axis represents, once again, the party’s statewide support. For each state, the Democrats are displayed in blue, and the Republicans in red. Each state appears twice, once in red, and again in blue on the other side of the graph.

Figure 8: Share of a party’s voter’s living in winning neighborhoods, state lower chambers
Figure 9: Share of a party’s voter’s living in winning neighborhoods, state upper chambers

Figure 10: Share of a party’s voter’s living in winning neighborhoods, Congressional delegation

These graphs show very clearly the spatial efficiency advantage enjoyed by Republicans in
most states. Consider, for instance, a paired comparison of Minnesota and Virginia, where Democrats typically win between 54 and 56 percent of the vote in statewide offices, and Arizona and Georgia, where the same is true of Republicans. In the first two states, only a little more than half of Democrats reside in Democratic neighborhoods defined at the scale of state legislative districts. In Arizona and Georgia, however, over 80 percent of Republicans live in majority-Republican neighborhoods. Democrats are distributed somewhat more efficiently at the scale of Congressional districts, but the difference remains striking.

It is also instructive to examine the swing states right in the middle of the graph. For instance, in North Carolina, Missouri, and Iowa, at each spatial scale the share of Republicans living in Republican neighborhoods is larger than the share of Democrats living in Democratic neighborhoods.

Throughout the middle of the distribution of state-level partisanship, there is a large gap between the red and blue data markers, indicating that with the same vote share, a larger share of Republicans live in majority-Republican districts than is true for Democrats. The relative efficiency of support for Republicans is especially noteworthy in some of the largest states, including California, New York, and Texas.

This pattern is not universal, however. The Democrats have a relatively efficient support distribution in the New England states at each spatial scale, and at the level of Congressional districts, in New Mexico and Nevada. The pattern is also broken in Ohio and Wisconsin, where medium-sized 19th century Democratic industrial agglomerations are relatively well distributed throughout the state.

The lower chamber in Wisconsin is of particular interest, since it is the focus of *Gill v. Whitford*—a partisan gerrymandering lawsuit in which the defense has argued that un-
usually large pro-Republican electoral bias is driven by the relatively inefficient geographic
distribution of Democrats in cities. Figure 8 casts considerable doubt on the claim that
the Wisconsin Democrats are characterized by an unusually inefficient spatial distribution. In
fact, at the spatial scale of the Wisconsin lower chamber, Figure 8 shows that less than
half of Wisconsin Republicans live in Republican neighborhoods, while over 70 percent of
Democrats live in Democratic neighborhoods. Among states with relatively small long-term
statewide Democratic majorities, Wisconsin’s Democrats have a more efficient spatial distri-
bution than their co-partisans in comparable Democratic states like Colorado, Minnesota,
Virginia, Pennsylvania, and New Hampshire. At the scale relevant for lower-chamber dis-
tricts, Democrats typically form local majorities not only in larger cities like Milwaukee and
Madison, but also in small cities like Green Bay, Appleton, Kenosha, and Eau Claire.

The information conveyed in Figures 8 - 10 can be used to calculate something akin
to the notion of the “efficiency gap” as conceptualized by Stephanopoulos and McGhee
(2015). Using data from district-level election results, they calculate the difference between
the parties’ respected “wasted” votes as a share of total votes cast, where wasted votes are a
combination of lost votes cast for losing candidates and surplus votes for winning candidates
in excess of the threshold for victory.

We can do something similar using nearest neighborhoods rather than districts. We divide
the share of Democrats living in Democratic neighborhoods by the Democratic vote share,
and we divide the share of Republicans living in Republican neighborhoods by the Republican
vote share. We then subtract the latter quantity from the former, which provides us with a
measure that, like the Stephanopoulos and McGhee (2015) measure, takes on positive values
when Democrats have a more “efficient” support distribution, and negative values when this
is true for Republicans.

We have also applied the Stephanopoulos and McGhee (2015) formula to 2008 district-level presidential election results for each chamber. This “efficiency gap” that is calculated from actual enacted districts is correlated with our nearest-neighbor-based measure at around .8 for each legislative chamber. This very high correlation is driven in part by the fact that both constructs are mechanically correlated with state-level partisanship. For most realistic spatial distributions, a party inevitably wastes more votes as it finds itself losing by larger and larger margins.

Nevertheless, this very high correlation suggests that without some accounting for underlying political geography and overall partisanship, a surplus of wasted votes for one of the parties cannot easily be interpreted as evidence of intentional gerrymandering. The remainder of this paper attempts to provide such an accounting.

**Spatial Efficiency and Representation**

The next step is to understand how the parties’ relative spatial efficiency translates into legislative seats at varying levels of overall state partisanship. To examine this, we regress the average long-term Republican seat share in the state on the long-term Republican vote share, the share of Republicans residing in Republican districts, and the interaction of the two. The logic of this specification is simple: the translation of a party’s statewide partisanship to legislative seats is moderated by the underlying geographic (in)efficiency of support.
In Figure 11, we plot the predicted seat shares from these models. As can be seen by referring back to Figure 4, these simple models focusing on the interaction of overall partisanship and spacial efficiency do a remarkably good job explaining the shape of the cross-state relationship between statewide political support and seats. Without any information about the partisan identity of those drawing the districts, we can explain the shape of the cross-state vote-seat curve based purely on the geography of the parties’ support. In general, the fact that Republican seat shares exceed the predictions of the cube law, and indeed even

---

3Congressional models include only states with more than two Congressional seats. For more details about these models, see the Appendix.
exceed proportionality, can be explained quite well by the superior spatial distribution of Republicans in suburbs and heterogeneous rural areas. As in Figure 4 above, the bonus in predicted Republican seats owing to political geography is especially large in states where Republicans make up slightly less than half of the electorate, and continues into swing states where Republicans constitute small statewide majorities.

The Detection of Gerrymanders

This may seem at first blush like a refutation of the claim that gerrymandering is an important determinate of Republican over-representation in U.S legislatures. On the contrary, Figure 11 provides a powerful baseline against which to examine that claim. It is quite plausible that when Republicans control the decennial redistricting process, they are able to further “pack” and “crack” Democrats to as to push the spatial inefficiency of Democratic support well beyond that which is achieved by patterns of political geography alone, or to create an advantage in states like Wisconsin or Ohio where one otherwise would not be present. It is also possible that Democrats might be able to dampen this effect when they are able to draw districts.

To examine this possibility, we estimate the same simple models as above, focusing only on elections since the last round of redistricting, and we add dummy variables capturing the partisanship of those drawing the lines. One indicator variable takes on the value 1 if Democrats had unified control over the redistricting process for the chamber in question after the 2010 census, and zero otherwise. Another indicator variable takes on the value 1 if Republicans had unified control, and zero otherwise. Both of these variables take on the
value zero in instances of divided control, independent commissions, or court-drawn plans.\textsuperscript{4}

We also include an indicator variable that takes on the value 1 if the state was, at the time of redistricting, required by Section 5 of the Voting Rights Act (VRA) to seek approval from the Department of Justice for any modifications to its electoral laws. Section 5 of the VRA gave the Department of Justice the responsibility to block redistricting proposals that had potential to dilute the ability of minority voters to elect candidates of their choice, and thus constrained the choices of state legislatures.\textsuperscript{5}

The results of these models are displayed in Table 1. The clearest partisan effects are for Republican control. Controlling for the overall partisanship of the state and the spatial efficiency of partisanship, Republican control over the redistricting process is associated with a 12 percentage point increase in the Republican seat share among upper chambers, an 8 percentage point increase among lower chambers, and an 8 percentage point increase among Congressional delegations. While the coefficients are negative for Democratic control, as one might expect, the effect is only statistically significant for upper chambers. The coefficients for the VRA variable suggest that controlling for partisan control of the redistricting process and the spatial efficiency of partisanship, states subjected to Section Five oversight had significantly higher Democratic vote shares. This suggests that for the VRA states, the creation of majority-minority districts led to more Democratic seats than in states with similar overall partisanship, political geography, and partisan redistricting control that were not subjected to Section Five oversight.

In short, it is clear that the partisanship of those drawing the lines is quite important as a

\textsuperscript{4}Source: http://redistricting.lls.edu/who.php

\textsuperscript{5}States covered by Section Five included Alabama, Arizona, Georgia, Louisiana, Mississippi, South Carolina, Texas, and Virginia.
Table 1: Simple OLS models predicting average Republican seat share, 2012-2016

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Upper Chambers</th>
<th>Lower Chambers</th>
<th>Cong. Delegation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.  SE</td>
<td>Coef.  SE</td>
<td>Coef.  SE</td>
</tr>
<tr>
<td>Share of all R living in R districts</td>
<td>0.44 (0.22) **</td>
<td>0.43 (0.19) **</td>
<td>1.11 (0.40) ***</td>
</tr>
<tr>
<td>Ave R share</td>
<td>0.69 (0.46)</td>
<td>0.91 (0.43) **</td>
<td>1.98 (0.72) ***</td>
</tr>
<tr>
<td>Interaction</td>
<td>-0.31 (0.46)</td>
<td>-0.40 (0.42)</td>
<td>-1.40 (0.82) *</td>
</tr>
<tr>
<td>Dem. control</td>
<td>-0.13 (0.04) ***</td>
<td>-0.05 (0.03)</td>
<td>-0.06 (0.06)</td>
</tr>
<tr>
<td>Rep. control</td>
<td>0.12 (0.03) ***</td>
<td>0.08 (0.03) ***</td>
<td>0.08 (0.04) *</td>
</tr>
<tr>
<td>VRA Section 5</td>
<td>-0.10 (0.04) ***</td>
<td>-0.06 (0.03) *</td>
<td>-0.10 (0.05) **</td>
</tr>
<tr>
<td>Constant</td>
<td>0.02 (0.18)</td>
<td>-0.07 (0.16)</td>
<td>-0.62 (0.31)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.83</td>
<td>0.85</td>
<td>0.84</td>
</tr>
<tr>
<td>Observations</td>
<td>49</td>
<td>48</td>
<td>36</td>
</tr>
</tbody>
</table>

*** = p < .01, ** = p < .05, * = p < .1
Dependent variable: Republican seat share from 2012 to 2016
Congressional models include only states with more than two seats.

predictor of seat shares, even when we control for the interaction of statewide partisanship, political geography, and the potential influence of Section Five of the Voting Rights Act. However, this analysis allows us to go much further and address a question of crucial importance in the courts: to what extent are individual states outliers? Given what we now know about the translation of partisanship and political geography to seats among U.S. states, for each state we can ask: for a state with its overall partisanship and political geography, what is the expected seat share? We can then contrast this seat share with the observed seat share.

In order to get a clearer sense of the role of partisan gerrymandering in legislative representation and evaluate specific cases, we estimate the basic political geography model— including only overall statewide Republican partisanship, the share of Republicans residing
in Republican neighborhoods, and their multiplicative interaction—and focus on only those states whose districts were drawn through a bipartisan process, the intervention of courts, or an independent commission. In the Figures below, the predictions from these models are presented with small black data markers, including predictions for those for out-of-sample states where districts were drawn by partisans. For those states with partisan redistricting, actual seat shares obtained by the parties since the 2010 round of redistricting are indicated with red (Republican) and blue (Democratic) data markers. Given a state’s combination of overall partisanship and political geography, the predicted seat shares from these models can be viewed as benchmark seat shares that would be expected under a districting process that is not overtly partisan. Of course it is quite possible that courts and independent commissions sometimes act with partisan motives, but the predicted seats based on the experience of these states present an especially useful baseline. In the appendix, we consider models in which all states—even those where districts were clearly drawn with partisan motives—are included in the estimation of the baseline.

---

6This was the case for 23 upper chambers, 21 lower chambers, and 14 Congressional delegations where there are more than two districts. See the appendix for details.
Figure 12: Predicted seats from spatial efficiency model (small black markers) and actual seats in legislatures with districts drawn by Republicans (red markers) and Democrats (blue markers), state lower chambers.
Figure 13: Predicted seats from spatial efficiency model (small black markers) and actual seats in legislatures with districts drawn by Republicans (red markers) and Democrats (blue markers), state upper chambers.
As in Figure 11, which focused on a longer time period and drew upon all states, the black data markers in these graphs reveal that even when the predictions are based only on states without partisan redistricting, Republicans can often expect and advantage due to their superior geography. The most striking feature of these graphs is the distance between the red and black data markers. These graphs make it very clear that the large Republican advantage in many states—especially those in the middle of the distribution—cannot be explained by the superior spatial efficiency of support for Republicans. The red data markers indicate that
the seat advantage for Republicans goes well beyond that which can be predicted purely from their political geography. For the most part, the states where Republican districting failed to improve on the predictions of the spatial efficiency model were those whose districting plans were required under Section Five of the Voting Rights Act to be approved by the Department of Justice. This is consistent with the finding of Chen and Cottrell (2016), who discover that Congressional delegations of VRA pre-clearance states are somewhat more Democratic than a series of simulated districting plans that are blind to party and race.

It is especially worthwhile to examine the cases where plaintiffs have filed lawsuits related to partisan gerrymandering. For example, the Wisconsin lower chamber at issue in *Gill v. Whitford* is a clear outlier in Figure 15. As discussed above, the spatial distribution of Democrats in Wisconsin is not especially inefficient, and accordingly, the small black data marker corresponding to the Republican seat prediction for Wisconsin is lower than for other states with similar Republican statewide support. Yet the observed Republican seat share is higher than the model prediction by over 12 percentage points. In this case, courts can make a strong inference that much of the disproportionate representation of Republicans is explained by efforts to “pack” and “crack” Democrats. The same can be said about the North Carolina Congressional map (see Figure 17). These graphs also suggest that plaintiffs would be able to characterize some of the districting plans in Michigan, Florida, Indiana, Pennsylvania, and Ohio as extreme partisan outliers. In each of these cases, the cross-state model captures the fact that because Republicans are distributed more efficiently than Democrats, they can expect seat shares beyond the cube rule, and beyond proportional representation. Nevertheless, we can also see that Republicans have obtained seats far beyond what would be expected based on the state’s overall partisanship and the spatial
arrangement of Republicans.

The Democrats have controlled the redistricting process in far fewer chambers. Above all, Figure 17 allows for an interesting perspective on the Maryland Congressional map that has been challenged in *Benisek v. Lamone*. Because Republicans are well distributed throughout suburban Maryland and Democrats are highly concentrated in immediate Washington suburbs and Baltimore, the spatial efficiency model predicts a seat share for the Republicans that goes well beyond the cube rule. By drawing a districting plan that produces 7 Democrats out of 8 seats, the Democrats have brought their representation in line with the expectation of the cube rule. In recent years, Maryland is a 64 percent Democratic state. There is little question that Democratic map-drawers attempted to maximize Democratic representation. By doing so, they have brought their partisan representation in line with Republican seat shares obtained in Southern states that are 64 percent Republican.

In other words, Maryland Democrats appear to have successfully overcome their spatial inefficiency. They appear to have done something similar in both chambers of the state legislature. The same can be said for Democrats drawing state legislative districts in both chambers in Rhode Island, and those drawing the upper chamber districts in Illinois. Only in Massachusetts (all three chambers) and in the state legislature of West Virginia do Democrats in charge of the redistricting process manage to improve upon the seat share that would be predicted from the cube rule. For the most part, gerrymandering efforts of Democrats in recent years appear to have focused on neutralizing the superior spatial efficiency of Republicans.
Alternative Approaches

This approach is useful because it provides a way to distinguish asymmetries in representation that emerge from patterns of political geography, and those that emerge from gerrymandering. It does so by characterizing the extent to which partisans cluster into “neighborhoods” defined at varying spatial scales, and exploring the cross-state relationship between partisanship, spatial clustering, and legislative representation.

It is thus a complement to an existing nascent literature that attempts to disentangle political geography from gerrymandering via redistricting simulations. Recent papers by Chen and Rodden (2013), Chen and Rodden (2015), Chen and Cottrell (2016), and Cho and Liu (2016) start with precinct-level electoral data and simulate large numbers of alternative districting plans, contrasting the partisan outcomes of the simulated plans with the partisanship of a specific plan that has been called into question. For instance, Chen (2017) simulates many alternative redistricting plans for the Wisconsin lower chamber at issue in Gill v. Whitford, and finds that the partisanship of the enacted plan is an extreme outlier when contrasted with the distribution of simulated plans, leading to the inference that the partisanship of the enacted plan can be explained by partisan gerrymandering.

With the simulation approach, the relevant baseline comes not from a characterization of the parties’ relative geographic efficiency in comparative perspective, but from an examination of hypothetical alternative non-partisan ways to draw districts. Both approaches have costs and benefits. The simulation approach is appealing above all because it offers a very direct way to examine the relevant counter-factual in a lawsuit about partisan redistricting: the likely result of a non-partisan process given a state’s underlying political geography.
Critics of the simulation approach worry, however, about whether the universe of simulated plans resemble plans that would be considered by human map-drawers, whether we can characterize the objectively “correct” distribution of plans to draw from, and whether the simulated plans are sufficiently random draws from that distribution (Altman et al. 2017). There is also no agreement about the best way to characterize the partisanship of the alternative plans once they have been simulated. For instance, Buzas and Warrington (2017) are critical of the logit model used by Chen and Cottrell (2016) to predict wins and losses in simulated districts.

The nearest neighbor approach described in this paper circumvents these critiques. Rather than attempting to sample from the distribution of potential redistricting plans, we treat every voter in a state as the centroid of his or her own bespoke district, and then characterize the extent to which one party or the other has a more efficient spatial distribution of support at each relevant spatial scale. A strength of our approach is that it explicitly situates the political geography, statewide partisanship, and representation of each state legislative chamber and Congressional delegation in comparative perspective. The reliance on cross-state analysis for the production of a baseline against which to contrast specific plans, however, is also the primary weakness of our approach. Unlike the simulation approach, we do not have a truly “non-partisan” baseline. Rather, we rely on the political geography and representation outcomes experienced by other states, paying special attention to those where the redistricting process can most plausibly be characterized as non-partisan.

The simulation and nearest-neighbor approaches are quite distinct, but fortunately, they produce similar inferences about the states that are most likely to be contested in the courts. For instance, many of the same states identified as partisan outliers by Chen and Cottrell
(2016) are also clearly partisan outliers using the nearest neighbor approach: Pennsylvania,
Ohio, Michigan, Florida, North Carolina, Wisconsin, Virginia, and Maryland. Courts might
find it reassuring that these very different approaches yield similar inferences.

**Conclusion**

The Democrats have become an overwhelmingly urban political party. Their geographic
concentration has increased with each election since the New Deal (Rodden 2017), and this
phenomenon perhaps reached its apotheosis in the general election of 2016. Something
similar is happening in other countries, where parties of the left have come to dominate in
poor post-industrial neighborhoods and dense city centers that are connected to the global
economy, while parties of the right are more successful in rural and exurban areas.

Thus in the United States and beyond, due to an inefficient geographic distribution
of support, parties of the left can be systematically under-represented when single-member
districts are drawn, even if those districts are drawn without the intention to produce unequal
representation. However, it is also the case that in the United States, districts are often
drawn by incumbents bent on providing an advantage for their political party. As a result
of these two forces, the under-representation of the Democratic Party in Congress and state
legislatures relative to its overall support has been one of the most striking features of
contemporary American politics.

Opponents of partisan gerrymandering have attempted to convince the courts that the
practice runs afoul of state and federal law, but thus far, their efforts have been plagued by
the observational equivalence of these two explanations of Democratic under-representation.
The most obvious defense for supporters of partisan redistricting is to claim that since Democrats are inefficiently distributed in cities, large asymmetries in the transformation of votes to seats are to be expected in all states. Thus a crucial task for legal challenges to partisan gerrymandering is to clarify the conditions under which asymmetric Democratic clustering leads to under-representation, and to quantify the impact of this phenomenon on representation, making it possible to reveal the impact of partisan gerrymandering if there is one.

In a study of U.S. states, this paper has shown that the party of the urban left indeed suffers from a relatively inefficient distribution of support in many states, and this has serious implications for representation. However, the problem is not universal. With the right distribution of city sizes and locations and a fortuitous spatial scale for districting, the party of the urban left can sometimes avoid an inefficient spatial distribution of support. Nor does the spatial inefficiency of Democrats fully explain their under-representation. We have quantified this relative spatial inefficiency at various spatial scales using a simple “nearest neighbor” approach, which provides us with a way to account for spatial inefficiency in cross-state analyses of representation, thus revealing the impact of gerrymandering.

In general, it is clear that when Republican legislators are able to control the process of redistricting, they can increase their party’s representation well beyond that which would be predicted purely from their superior spatial distribution, or in some instances, create a spatial advantage that would otherwise not exist. On the other hand, when the Democrats control the redistricting process, they often endeavor to craft districts that assuage their spatial inefficiency. Our approach provides courts with a useful tool when faced with the task of evaluating whether, and to what extent, asymmetry in the transformation of votes
to seats can be explained by such efforts.
References


Appendix

A  Empirical models used to produce Figure 11

Table 2: Simple OLS models predicting long-run Republican seat share

<table>
<thead>
<tr>
<th></th>
<th>Upper Chambers</th>
<th>Lower Chambers</th>
<th>Cong. Delegation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>SE</td>
<td>Coef.</td>
</tr>
<tr>
<td>Share of all R living</td>
<td>0.47 (0.26)</td>
<td>*</td>
<td>0.43 (0.19)</td>
</tr>
<tr>
<td>in R districts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term R share</td>
<td>1.97 (0.51)</td>
<td>***</td>
<td>1.60 (0.42)</td>
</tr>
<tr>
<td>Interaction</td>
<td>-0.89 (0.56)</td>
<td></td>
<td>-0.70 (0.44)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.45 (0.19)</td>
<td></td>
<td>-0.45 (0.19)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.72</td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>Observations</td>
<td>49</td>
<td></td>
<td>48</td>
</tr>
</tbody>
</table>

*** = p < .01, ** = p < .05, * = p < .1
Dependent variable: Republican seat share from 2008 to 2016
Congressional models include only states with more than two seats.
B Empirical models used to produce Figures 15 - 17

Table 3: Simple OLS model predicting Republican seat share since 2012, only states with non-partisan or bipartisan redistricting

<table>
<thead>
<tr>
<th>Coef.</th>
<th>SE</th>
<th>Coef.</th>
<th>SE</th>
<th>Coef.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper Chamber</td>
<td>Lower Chamber</td>
<td>Cong. Delegation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of all R living in R districts</td>
<td>0.50 (0.35)</td>
<td>0.58 (0.28)</td>
<td>**</td>
<td>-0.25 (0.81)</td>
<td></td>
</tr>
<tr>
<td>Long-term R share</td>
<td>1.25 (0.79)</td>
<td>1.61 (0.69)</td>
<td>**</td>
<td>1.04 (1.11)</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>-0.62 (0.75)</td>
<td>-0.94 (0.65)</td>
<td>1.07 (1.61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.20 (0.28)</td>
<td>-0.33 (0.23)</td>
<td>-0.15 (0.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.75</td>
<td>0.82</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>21</td>
<td>20</td>
<td>13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** = p<.01, ** = p<.05, * = p<.1
Dependent variable: Republican seat share from 2012 to 2016
Congressional models include only states with more than two seats.

C Predictions based on all states

Figures 15 - 17 in the main text displayed predictions from models that only included states where districts had been drawn by legislatures or commissions under split partisan control, or by courts or independent commissions. It is also useful to examine predictions from models that include all states, even if districts were drawn by legislators under unified partisan control. Since over 40 percent of the districting plans under analysis were drawn under unified Republican control, this cannot be viewed as anything like a non-partisan baseline, but it is still useful for identifying partisan outliers. The details of the model are presented in the table below, and the subsequent figures are identical to Figures 15 - 17, except that the black data markers are in-sample predictions from models that include all states. As
one would anticipate given the results in the paper, the predicted seat shares are more Republican than in the models that only include states with non-partisan or bipartisan redistricting processes. Nevertheless, it is still possible to identify the same set of partisan outliers.

Table 4: Simple OLS model predicting Republican seat share since 2012, all states

<table>
<thead>
<tr>
<th></th>
<th>Upper Chambers</th>
<th>Lower Chambers</th>
<th>Cong. Delegation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef.</td>
<td>SE</td>
<td>Coef.</td>
<td>SE</td>
</tr>
<tr>
<td>Share of all R living in R districts</td>
<td>0.59 (0.27) **</td>
<td>0.55 (0.20) ***</td>
<td>1.25 (0.43) ***</td>
</tr>
<tr>
<td>Long-term R share</td>
<td>1.73 (0.52) ***</td>
<td>1.56 (0.43) ***</td>
<td>2.61 (0.75) ***</td>
</tr>
<tr>
<td>Interaction</td>
<td>-0.89 (0.56) ***</td>
<td>-0.82 (0.45) *</td>
<td>-1.84 (0.87) **</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.38 (0.20)</td>
<td>-0.32 (0.16)</td>
<td>-0.87 (0.32)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.72</td>
<td>0.79</td>
<td>0.81</td>
</tr>
<tr>
<td>Observations</td>
<td>49</td>
<td>48</td>
<td>36</td>
</tr>
</tbody>
</table>

***= p<.01, **= p<.05, *= p<.1
Dependent variable: Republican seat share from 2012 to 2016
Congressional models include only states with more than two seats.
Figure 15: Predicted seats from spatial efficiency model (small black markers) and actual seats in legislatures with districts drawn by Republicans (red markers) and Democrats (blue markers), state lower chambers.
Figure 16: Predicted seats from spatial efficiency model (small black markers) and actual seats in legislatures with districts drawn by Republicans (red markers) and Democrats (blue markers), state upper chambers
Figure 17: Predicted seats from spatial efficiency model (small black markers) and actual seats in legislatures with districts drawn by Republicans (red markers) and Democrats (blue markers), U.S. Congress