A Geographic Explanation for Partisan Representation:
How Residential Patterns of Partisans Shape Electoral Outcomes

David Cottrell *

Abstract

Partisan favor in the votes-seats relationship is often attributed to gerrymandering, where parties improve their seat share by manipulating district boundaries in order to bias the distribution of votes in their favor. However, recent research by Chen and Rodden (2013), which builds on a thesis developed by Erikson (1972), provides support for the claim that such bias is less a function of gerrymandering than it is a function of the underlying geographic distribution of partisan voters. They find that the tendency for Democrats to cluster in dense urban areas - as we see in states like Florida - gives Republicans a natural seat advantage in legislatures, even in the presence of redistricting. Still, the link between partisan residential patterns and the votes-to-seats relationship remains undeveloped, in part because we are unable to observe the full spectrum of partisan votes at various levels of partisan clustering. In this paper, however, I attempt to develop this link through simulations. By simulating congressional elections at all levels of partisanship, I can measure the hypothetical electoral outcome that results from adjusting the degree and type of geographic clustering using an agent-based model of segregation. These simulations allow me to observe the potential electoral effect of partisan clustering on the entire vote-seats curve. As a result, I show that residential clustering of partisans tends to flatten the votes-seats curve, making it disadvantageous to the majority party, whichever party that may be. However, when Democrats cluster more than Republicans, the disadvantage is greater for Democratic majorities than for Republican majorities.

*Postdoctoral Fellow for the Program in Quantitative Social Science, Dartmouth College
1 Introduction

Although Democrats won the majority of the vote in the 2012 House elections, Republicans won the majority of the seats - a result that reveals the inherent disconnect between votes and seats in single member district electoral systems. The election is a reminder that even in the most representative chamber of Congress, there is the potential for unrepresentative electoral outcomes. Of course, with Democrats losing a majority of the seats, many have claimed that the system is not only unrepresentative, but also biased to favor Republicans. A common argument in explaining the source of this bias is that it is due to gerrymandering. Republican-controlled state legislatures manipulate district boundaries to more efficiently convert votes into seats ([Wang, 2012]). However, while recent research has found the effect of gerrymandering to be significant, its contribution to pro-Republican bias in districting plans is likely marginal in the aggregate ([Goedert, 2014; Chen and Cottrell, 2015]).

Therefore, to account for the bias beyond that which we can attribute to gerrymandering, scholars have returned to an old thesis that suggests that Democrats are the victims of a “natural” disadvantage that results from their geographic residential patterns. With the industrial revolution drawing populations into dense metropolitan areas and postwar suburbanization leading to residential sorting along socio-economic and racial lines, there has been an increase in the geographic division between Democrats and Republicans ([Nall, 2015]). This division is marked by a pattern where Democrats have clustered into densely populated urban areas, while Republicans have dispersed along peripheral suburban and rural communities ([Rodden, 2010]). As a result, Democratic voters tend to pack into urban districts producing a suboptimal distribution of Democratic votes, which allows Republicans to win more seats per vote than they would have in the absence of such clustering. And even as redistricting corrects for the shifts in the underlying population through reapportioning the districts, this geographic bias seems to persist. This is a phenomenon that [Erikson, 1972] noticed years ago when he suggested that pro-Republican bias was an unintended “accident” of geography. Even in the face of gerrymandering, as the thesis goes, the Democratic ten-
dency to reside in dense urban areas has the effect of reducing the Democrats’ share of the legislative seats.

This thesis has gained a lot of attention given the recent findings by Chen and Rodden (2013). By redrawing the district maps of state legislatures using a computer to remove intentional gerrymandering, the authors show that as long as districts are designed according to neutral rules - that they are compact, contiguous, and equally apportioned - an increase in the concentration of Democratic voters will naturally lead to an election outcome favoring Republicans. Specifically, Chen and Rodden conclude that in states where “Democrats are inefficiently concentrated in large cities and smaller industrial agglomerations...they can expect to win fewer than 50% of the seats when they win 50% of the votes” (239). Therefore, Democrats can expect to be disadvantaged by geography so long as there is a 50-50 split in the vote.

However, while geographic clustering might reduce Democratic seat share at a hypothetical 50% vote share, it is unclear how such clustering impacts Democrats at alternative shares of the two-party vote. Not only are we unable to observe electoral outcomes under the full spectrum of the potential two-party vote, we are also unable to observe each of those outcomes under various degrees of clustering. It is simply difficult to say whether a party was advantaged by partisan geography because we simply do not know what the counterfactual would look like under other degrees and types of clustering. Therefore, I investigate how the geographic clustering of partisans affects electoral outcomes in hypothetical situations where Democrats and Republicans split the vote across all margins of victory.

In this paper, I develop hypotheses for how partisan clustering affects representation. By simulating Congressional elections in a hypothetical ten-district state, I am able to isolate the effect of partisan geography on the votes-seats relationship across the full spectrum of the partisan vote. Using a simple agent-based model of segregation, I approximate real-world partisan residential patterns by adjusting the degree to which Democrats and Republicans cluster. I then use a districting algorithm to randomly assign these voters to compact,
contiguous, and equally apportioned districts, where they are then translated into seats.

As a result, I show how the residential clustering of Democratic voters affects the votes-seats curve. I find that Democratic clustering does not always disadvantage Democrats. Instead, in states where they hold a certain minority of the vote, Democrats can improve their seat share through clustering. While it is common to suggest that Democrats have lost seats due to geography - and there is no doubt that they have - the party has also won seats due to geography. In states where Republicans hold more than a marginal majority of the vote, clustering can have the effect of creating a Democratic district where one might not otherwise have existed. Hence, Republicans are disadvantaged by dense urban areas that create strong Democratic districts. However, in states where Democrats hold the majority of the vote, clustering has the opposite effect. Instead, the clustering of Democratic voters reduces the number Democratic seats that would have been won under a less clustered environment.

In either case, geographic clustering flattens the votes-seats curve, making it harder for both parties to convert additional votes into additional seats. And as the votes-seats curve flattens, minority parties will gain seats as the majority party loses seats. Therefore, the majority party - be it Democrat or Republican - is disadvantaged by Democratic clustering while the minority party is advantaged. However, if the clustering is asymmetric - where Democrats cluster more than Republicans - then the votes-seats curve will bend asymmetrically. As a result, the Democratic majority will lose more seats from the clustering of partisans than the a Republican majority.

2 How Votes Translate into Seats: Review of the Votes-Seats curve

Scholars of U.S. electoral systems have long been interested in how a party’s aggregate vote share in legislative elections translates into its seat share in the legislature. Unlike
proportional rule systems where a party’s seat share is approximately equal to its vote share, the votes-seats relationship in single-member district systems is not necessarily proportional, nor is it necessarily linear. In fact, to early observers of elections in both Great Britain and the United States the relationship appeared exponential in nature. Some even went so far as to suggest that the ratio of seats was determined by simply cubing the ratio of votes (Kendall and Stuart, 1950; March, 1957). Today, we understand that the true relationship between votes and seats is more complex. A party’s aggregate seat share is not just a function of its vote share, but it is also a function of how its votes are distributed across the many districts.

This phenomenon is interesting to political scientists because it means that a party need not change its aggregate vote share in order to change its seat share. Instead, a simple redistribution of its votes is often sufficient to drastically change the number of seats it wins in the legislature. For example, take the vote-seats relationship of a hypothetical ten-district legislature where the districts are equally apportioned. A party with just over 30% of the vote has the potential to win a majority of the seats. And a party with 50% of the vote can win as many as 100% or as few 10% of the seats, depending on how those votes are allocated across the districts.

Figure 1 displays this phenomenon. It gives the range of potential outcomes for an election in an equally apportioned ten-district legislature with a sufficiently large number of voters. Any votes-seats combination can be reached between the two stair-step lines. Which combination depends only on how those votes are distributed. Therefore, in determining the votes-seats relationship, it can be as important to explain the variation of a party’s votes across the districts as it is to explain its vote in the aggregate.

2.1 Partisan Districting

What, then, explains the variation of a party’s vote across districts? One common explanation is that it is the result of districting. In the United States, districting is simply the process of dividing the land among the districts and assigning the voters to the district in
Figure 1: The figure above displays the upper and lower bounds for the potential outcomes on the votes-seats curve in a hypothetical ten district legislature where the population is sufficiently large and the districts are equally apportioned. The figure conveys the point that there exists a distribution of the two-party vote across the districts that can produce any number of seats between the stair-step lines. For example, a party with just over 50% of the two-party vote can win as many as 100% or as few 10% of the seats, depending on how those votes are allocated across the districts.

which they reside. Therefore, the composition of a district is dependent on how the district is located among the voters. One need only to change the district’s location to alter its composition. And by changing the location of the district relative to the location of partisan voters, one could potentially redistribute partisans across the districts in a way that could change the votes-seats curve. As a result, those who control the districting process are able to potentially control the partisan results of a legislative election.

As one might expect, this is particularly troubling for democratic representation in the United States. With state governments in control of the districting process, there is always
the potential for partisan influence on district design. For this reason, parties have long been accused of using the districting process to bias legislative elections in their favor. And, of course, this suspicion continues to this day.

Yet, in an effort to move beyond the suspicion, political scientists have made a number of attempts to precisely identify gerrymandering and to measure exactly how and to what extent it affects representation. It was Tufte (1973) who began the practice of using variation in the bias and responsiveness of the votes-seats curve (where “bias” is the deviation from the 50%-50% intercept and “responsiveness” is the slope of the curve) as a way to measure how gerrymandering influenced representation. Scholars have since refined these measurements of bias and responsiveness (King and Browning, 1987; Campagna and Grofman, 1990; Gelman and King, 1994) and have analyzed its variation at both the national and state level, drawing mixed conclusions with respect to its effect. While some scholars have found that partisan gerrymandering leads to significant shifts in the votes-seats curve (Abramowitz, 1983; Squire, 1985; Erikson, 1972; Niemi and Winsky, 1992; Born, 1985) others have found that gerrymandering is not very influential at all (Ferejohn, 1977; Glazer, Grofman and Robbins, 1987; Squire, 1985; Abramowitz, Alexander and Gunning, 2006). However, to the extent that gerrymandering does play a role in determining the votes-seats curve, recent research has established that the effect is at least conditional on institutional constraints (Gelman and King, 1994; McDonald, 2004; Campagna and Grofman, 1990). Nonetheless, it remains difficult to isolate the effect of gerrymandering on state and federal elections.

---

1The term “Gerrymander” itself was coined in 1812 when Elbridge Gerry was famously accused of advantaging Republicans through the approval of a salamander-like district in Massachusetts.

2For example, Florida’s Supreme Court is currently adjudicating claims that the Republicans intentionally gerrymandered districts in the state legislature.

3Tufte measured the curve’s deviation from the 50%-50% intercept, which he referred to as the curve’s bias, and he measured its slope, which he referred to as the curve’s responsiveness, and he suggested that the bias and the declining responsiveness of the curve are partly the effect of gerrymandering. To be precise, he defined an unbiased votes-seats curve as one where 50% of the two-party vote would win 50% of the seats. Bias would occur when more or less than 50% of the vote is needed for a party to win 50% of the seats. He also defined responsiveness to refer to how seat share responds to a change in vote share. Therefore, it is a measurement of the change in seat share for a unit change in vote share.
Part of the difficulty in isolating the effect of gerrymandering is being able to attribute bias and responsiveness to partisan motivated gerrymandering. This is especially the case when the same bias and responsiveness in the votes-seats curve can result from non-political forces. For example, scholars noticed that prior to 1966 the votes-seats curve exhibited substantial Republican bias in the division of seats in Congress that seemed to run counter to the Democratic state legislatures. (Tufte, 1973). However, this bias was independent of intentional gerrymandering and instead likely arose from a combination of malapportionment and an increasing partisan divide between rural and urban voters (Erikson, 1972; Cox and Katz, 2002). Since the courts had not yet interpreted the Constitution as requiring states to apportion their districts according to roughly equal populations, district populations varied considerably in size. States rarely redrew their district boundaries to account for shifts in the population and, as a result, urbanization packed a disproportionate number of voters into urban districts. This effectively diluted the urban vote. And because these voters were largely Democrat, the diluted Democratic vote put them at a natural disadvantage in terms of legislative seats. Therefore, the source of the pro-Republican bias was not gerrymandering, but instead it was a combination of urbanization, geographic polarization, and a general failure for districts to correct for over-population. Republicans simply gained an edge through the natural migrations of the underlying population.

However, some of this edge vanished after 1965, when Westberry vs. Sanders (376, U.S. 1) required states to apportion their districts equally with respect to population. Coined as the “reapportionment revolution,” the next half of the decade would see numerous revisions to district boundaries in order to correct for the imbalance in population. This reapportionment naturally reduced the dilution of the Democratic vote that was caused by overcrowded urban districts. And since most of the unified state legislatures and courts responsible for drawing districts at the time were under the control of Democrats, the Democrats took advantage of the opportunity to draw pro-Democratic bias into their plans (Cox and Katz, 2002). Yet
even in the face of reapportionment and pro-Democratic gerrymandering, political scientists still recognized that the geographic concentration of Democrats remained a disadvantageous force against the Democratic party (Erikson 1972).

This link between the residential patterns of Democrats and pro-Republican bias is still recognized today (McDonald 2009, 2010; Erikson 2002; Jacobson 2003). Yet the link between the geographic clustering of Democrats and the entire votes-seats curve has not yet been thoroughly examined. While Chen and Rodden (2013) are the first to rigorously show that in a hypothetical situation where the partisan vote is split 50-50, urbanization tends to distribute votes across districts in such a way that produces a Republican seat advantage. However, they are also unable to make claims about how such clustering impacts representation across the entire votes-seats curve and not just under a 50-50 split of the vote. This is because they are limited in their ability to make inferences about the effect of geographic clustering at any other partisan share of the two-party vote. There simply aren’t enough election results in their data to make strong enough claims about the rest of the curve. However, by simulating elections one can begin to draw inferences about how the underlying geography of partisan voters influences the full votes-seats relationship.

In the next section, I attempt to do just this. I design a model of a hypothetical legislative election where individual voters - either Democrat or Republican - are arranged across two-dimensional space in a particular residential pattern. They are then assigned to randomly drawn, equally apportioned, roughly compact and contiguous districts where their votes are then aggregated and translated into seats. The model repeats this process for every potential partisan split of the vote while varying the geographic patterns of the voters. By manipulating the degree to which Democrats and Republicans cluster geographically, I am able to fully assess the electoral effect of geographic sorting. My findings suggests that Democratic clustering is damaging to Democrats only when they hold voting majorities. Otherwise, contrary to the implications of Chen and Rodden’s theory, clustering can lead to advantageous outcomes for the Democratic party.
3 Simulating Elections

In this section, I develop a computational model of a legislative elections where I am able to vary the residential patterns of partisan voters. To model partisan residential patterns I simply leverage an agent-based model of segregation. Since the objective is to study the effect of various degrees of clustering on electoral outcomes, I require a model that allows me to manipulate the spatial distributions of partisan voters where Democrats and Republicans can locate across geographic space in a manner that ranges from uniform to clustered. And, moreover, I require the model to be able cluster Democrats asymmetrically with respect to Republicans in order to approximate the real-world residential patterns we see in the United States.

To do this, I employ Shelling’s agent-based model of racial segregation \cite{Schelling1971, Chen2012}. Originally, this model was intended to explain the widely observed residential segregation of whites and blacks. Schelling uses it to show that relatively low degrees of intolerance toward living next to others of a difference race can transform an unsegregated neighborhood into a segregated one. While Schelling’s interest in this agent-based model was to explain racial segregation with micro-motivations, my interest in this model is primarily that it starts with an unsegregated population and ends with a segregated one, where - until convergence - each iteration is slightly more segregated than the previous iteration. By simply swapping whites and blacks with Democrats and Republicans, the model becomes one of partisan clustering.

Moreover, since there is a random component to the way partisans are initially distributed and there is a random component to the way partisans distribute at each iteration, the model allows for multiple geographic distributions to emerge. Therefore, by running the segregation model repeatedly, I can produce a randomly sampled set of clustered partisan environments that are distinct from their initial uniform distribution. Then to observe how this clustering translates into electoral results, I simply devise a districting algorithm that assigns the voters to a set of compact and contiguous districts of equal size. Then I aggregate the votes in each
district by counting the ratio of Democrats to Republicans and assigning district seats to the party that wins the majority of the vote in the district. This gives me the partisan seat share that results from the hypothetical legislative election, which can then be compared against the partisan vote share of the population. After repeating the “election” with different types of clustering and at different splits of the two-party vote, I can begin to make inferences about the vote-seats curve at different levels of partisan clustering.

It is important to note that the purpose of the model is not necessarily to capture the real-world residential motivations of partisans. The goal is to achieve a partisan segregated outcome that abstractly reflects real-world segregation. By randomly moving from a uniform outcome to a segregated outcome, I am able to loosely generate expectations about the effect of clustering across a spatial grid.

3.1 Designing the Spatial Patterns

The segregation model begins with 6,400 voters distributed uniformly across a 100 X 100 unit grid. The voters, therefore, occupy 64% of the units on the grid while 36% of the units are left as empty space. Once the voters have been uniformly distributed, each is randomly assigned to a party from an allotment of partisanship that has been predetermined. For example, Figure 2 displays the initial stage of a model where the aggregate partisan split between Republicans and Democrats is 50%-50%. At the initial stage, where partisanship is uniformly distributed, the grid will be then subdivided into a set of 10 districts of approximately equal size, each containing 640 voters. The votes of each district is aggregated and the number of Democratic seats are recorded.

Then Schelling’s segregation algorithm is initiated. The way the algorithm works is

---

4 Just as we should be skeptical that Schelling’s tipping model completely captures the micro-motives that lead to the kind of racial segregation that we see today (Bruch and Mare 2006), we should definitely be skeptical that it captures the micro-motives behind partisan segregation. I only use the model to achieve a segregated outcome, so I am able to remain agnostic to the micro-motives that lead to it. However, future iterations of this model may refine the micro-motives so that they better map onto real-world motivations such that real-world forms of clustering are subsequently generated.

5 Of course, an average Congressional district contains more than a thousand times this number. However, computation is made easier with fewer voters.
that the voters are asked whether they are satisfied with their current location. Voters are determined to be satisfied if they are in a neighborhood with voters that share their partisanship. For this model, voters are satisfied if more than 14 of their 24 nearest neighbors are of the same party. If they are unsatisfied, they are moved to a randomly chosen blank space. This is repeated for all voters until every unsatisfied voter has been relocated. Once all relocations are made, the algorithm begins a new iteration where voter satisfaction is queried and unsatisfied voters are once again relocated. I repeat this step ten times, regardless of convergence.

The effect of this process is that partisans segregate into geographic communities. Therefore, Democrats and Republicans have relocated to be nearer to each other. In this case,
Figure 3: Four Examples of Symmetric Partisan Clustering where Democrats hold 50% of the Vote

since Democrats and Republicans are given the same motivations, they separate from each other at equal rates. And as a result, their clustering tends to be symmetric. Figure 3 displays four examples of the distribution of partisans that result from running the segregation model at a 50-50 split of the aggregate partisan vote.

While the Schelling model accomplishes the goal of geographically concentrating Democratic voters, it also has the effect of concentrating Republican voters. Therefore, the segregation it creates is symmetric in nature. An argument could be made that this type of segregation is what we see with suburbanization in the United States. As urban centers become more homogeneously Democrat, the peripheries of these centers become more ho-
mogeneously Republican. One can easily see this effect in San Antonio or Houston, where Democrats and Republicans segregate almost equally along the suburban and urban divide. Such partisan sorting may lead to the type of symmetry that we see in Figure 3.

However, the more common argument is that the geographic concentration of partisans in the United States is often asymmetric. Democrats concentrate in urban centers while Republicans spread thinly across suburban, exurban, and rural communities. Therefore, to capture this effect, I make a simple modification to the Schelling model. Instead of relocating any voter that is unsatisfied with their location, I only relocate Democrats. In other words, Republicans remain at their initial location while Democrats gravitate toward one another. I allow the model to run for ten iterations, regardless of whether each Democrat is satisfied. As a result, Democrats tend to settle in clusters while a thinly distributed Republican party holds the periphery. Figure 4 displays four examples of the distributions generated by the model at a 50-50 split of the aggregate partisan vote.

I generate 30 of these grids at every percentage of the two-party vote from 30% to 70% Democrat. And I do this for each of the three partisan distributions that I mentioned above - uniform, symmetric, and asymmetric. These grids will act as the underlying geographic distribution of partisan voters over which the hypothetical elections will be held. However, the next challenge is assigning voters to districts in order to transform their votes into seats.

3.2 Designing the Districts

Since the objective of this hypothetical election is to observe how the spatial patterns of partisans impacts real-world elections, it is important to assign voters to equally assigned districts in a way that approximates the real-world districting process. For example, one important component of U.S. legislative elections is that districts are mapped to their own unique geographic location. Voters are then assigned to the district according to whether they reside in that location. Because districts are designed to be contiguous, no two districts will share the same geographic location and no geographic location will be without a district.
Figure 4: Four Examples of Asymmetric Partisan Clustering where Democrats hold 50% of the Vote

In addition to contiguity, districts tend to be designed to be relatively compact. This simply means that districts cover voters who live near each other rather than voters who are distant from each other. Although this is not an official constraint, it is a principle that is generally followed. It is the reason why a voter in Southern California will likely never share a district with a voter in Northern California. Rather, districts are drawn to contain voters that share the same proximate location. This is a pattern that can be observed across the states. Voters will be more likely to share a district with another voter if that voter is in their “neighborhood.”

In order to estimate the likely partisan outcome that would result from the hypothetical
ten-district elections generated by the segregation models, I would need to measure the outcomes of all potential voter-district assignments. However, determining the distribution of all potential assignments is a complex problem. Since there are innumerable ways to partition the voters on the grid into ten compact, contiguous and equally apportioned districts, I will instead have to randomly sample from the distribution of potential districts.

To do this, I leverage an algorithm that randomly searches the grid for a set of districts that fit the criteria of being compact, contiguous, and roughly equal in population. Specifically, the algorithm is a variant of a weighted k-means strategy and can be easily described in the following steps (see Figure 5 for a visual of the procedure):

**Step 1** Randomly generate a set of 10 seeds using k-means++. Therefore, the first seed is assigned to the location of any voter on the grid with uniform probability. Then the second seed is randomly assigned to the location of any voter on grid with a probability proportional to the distance from that voter to first seed. Then the third seed is randomly assigned to the location of any voter on grid with a probability proportional to the distance from that voter the nearest seed to that voter. This process repeats until all seeds have been set. The seeds will be the centroids of the ten districts. See Figure 5 (A).

**Step 2** Assign all voters to their nearest centroid, thereby partitioning them into districts with straight-line boundaries that divide the space according to its nearest centroid.

**Step 3** For each district, count the number of voters that have been assigned to it. Stop if all districts have a population that is within 2.5% of the mean population. Otherwise, continue.

**Step 4** Slightly shift the location of each district boundary in the direction that reduces the variance of the population across the districts. Hence, small districts expand to increase in size and large districts contract to decrease in size. Repeat this step, making

---

6The algorithm leverages a similar approach used in Fryer Jr and Holden (2011).
small shifts in the location of the boundaries, until all districts are within 2.5% of the mean population. Then stop. See Figure 5 (B).

**Step 5** The voters are assigned to their nearest centroids and the partisan vote is calculated. See Figure 5 (C).

While this districting algorithm is an attempt to replicate real-world map-making by assigning voters to districts in adherence to geographic standards, they are not, however, designed with partisan intent. Of course, partisan intent in the districting process is a reality. Real-world map-makers can choose the most favorable map from the set of potential maps. Or they can draw bias into their districts by ignoring compactness and drawing oddly-shaped districts where boundaries weave through the landscape to produce partisan favor. Yet there are a number of reasons to think that gerrymandering has its limitations and that legislators are bound by geographic constraints.

For example, although we would expect gerrymandered districts to take on wild shapes, most U.S. Congressional districts are designed with some reasonable compactness. Districts do not usually contain voters that live far from each other, which is to suggest that despite the opportunities for gaining advantage by drawing such non-compact, elongated, and far-reaching districts, map-makers tend not to do so. They seem to defer to a standard of maintaining some geographic compactness, meaning that the Congressional districts that contain voters in Los Angeles tend to contain other voters in Los Angeles. And while compact districts can still be designed to carry partisan favor, being constrained by the standard - at least in part - makes gerrymandering more difficult. Moreover, Chen and Cottrell (2015) find that most of the variation in partisanship across Congressional delegations can be explained by districts that have been drawn without partisan intent. Therefore the unbiased districting procedure used in my simulations may convert votes to seats similar to that of a real-world districting plan.
Figure 5: Once the locations of partisans have been established on the grid, an algorithm is used to randomly generate ten compact and contiguous districts of roughly equal size. It accomplishes this by randomly distributing ten centroids across the grid(a), where voters are assigned to the nearest centroid. Then, to correct for malapportionment, the gradient-step method iteratively moves the centroids in a direction that makes small improvements in apportionment(b). It stops when a roughly equal number of voters are assigned to a set of centroids(c).
3.3 Producing the Votes-Seats Curve

After running the segregation models and assigning voters to districts for each of the three types of clustering at every split of the two party vote, I am able to determine the outcome of each hypothetical election. This produces a Democratic seat share that is associated with every percentage of the Democratic vote. A LOESS smoother averages all the iterations of the model across every split in vote share. This generates the votes-seats curve for each of the three types of clustering.

Figure 6 plots these curves against each other. In each plot, the x-axis represents the Democrat’s share of the two-party vote and the y-axis represents the resulting Democratic share of the seats. While the first plot allows the x-axis to range from 30% to 70% vote share, the second plot limits the range from 45% to 55% in order to focus on less significant margins of victory. Each curve displays the relationship between vote share and seat share that occurs when drawing compact and contiguous districts over an electorate that is sorted according to the three different partisan spatial patterns. And by comparing the three curves we are able to observe how partisan geography can potentially influences U.S. representation. This is especially important in light of the fact that partisanship is becoming more and more segregated (Nall, 2015). Depending on the manner with which these partisans distribute - symmetrically or asymmetrically - these votes-seats curves will help us to understand its potential effect on partisan representation.

3.4 How partisan clustering affects the votes-seats curve

In Figure 6 each curve reflects the outcomes of the model under the three versions of clustering, where partisans are uniformly distributed (in blue), symmetrically distributed (in purple), and asymmetrically distributed (in pink). If voters are uniformly distributed, as displayed by the blue line, a “z-shape” curve will link votes-to-seats. This is consistent with winner-take-all representation, where small margins of victory allow the winning party to take all of the seats. This is because when the 6,400 voters are randomly distributed across
Figure 6: The votes-seats curve that results from 30 simulations performed at each split of the two-party vote for the three types of Democratic clustering. Each curve is smoothed using a LOESS smoother and 95% confidence intervals are included.

space and assigned to districts, the partisanship of the 640 members of each district will be very similar to the partisanship of the state. Therefore, when the partisanship of the state is just above 50% Republican, the partisanship of each of the districts will be just above 50%
Republican. In an election where the partisans are distributed uniformly, a simple margin of victory produces strong gains in seats for the majority party.\footnote{There is some bend in the curve due to noise in the partisan variation across districts, which would likely go away as the number of voters increases. However, we can think of the variation due to the low number of voters as the variation that occurs due to minimal partisan clustering. In which case, each voter on the grid would represent clusters of many voters as opposed to just one.}

The curve generated from symmetric clustering - the purple line - is a symmetric curve much like the curve generated by the uniform distribution. The symmetry simply means that Democrats and Republicans win the same number of seats when holding the same majority of the vote. Therefore, Democrats are no more well-off than Republicans would be when holding the same vote share. However, when compared to the curve produced by the uniform distribution of partisans, the curve produced by the symmetric distribution of partisans is much flatter. Although they both cross the 50%-50% intercept, the curve is closer to proportional representation than winner-take-all. Therefore, a party that wins 55% of the vote wins fewer seats as the underlying distribution of partisans begin to cluster. This means that geographic polarization has the effect of making the majority party worse off, regardless of which party it is.

A similar effect can be observed under asymmetric Democratic clustering, although there are some important differences. When Democrats tend to cluster more than Republicans (asymmetric Democratic clustering), as we often see with urbanization, it produces an asymmetric votes-seats curve. Therefore, when compared to the curve produced by the uniform distribution, an asymmetric votes-seats curve produces different gains and losses with respect to the parties. First, as Chen and Rodden (2013) find, when the two parties approximately split the vote, the Democrats win fewer seats than they would have under a less-urban environment. In which case, Republicans win more seats due to geography alone. This is the Republican advantage that is commonly associated with geography. However, analyzing the effect of Democratic clustering at a hypothetical 50% vote share and inferring that geography is only a problem for Democrats misses the full picture of how geography impacts the votes-seats curve. The model improves upon our understanding of the advantages and
disadvantages that both parties experience in the face of partisan clustering.

Moreover, it is important to know how parties are represented beyond a 50% votes share. Parties are often rewarded for obtaining a supermajority of the seats. For example, in U.S. legislatures, certain supermajorities can override a veto. This presents a major partisan advantage if the opposite party was able to win the executive seat. Moreover, supermajorities allow for parties to overcome potential defectors when passing legislation. The larger the majority, the more the party can afford to lose potential swing voters.

To achieve supermajorities, parties will need to achieve larger majorities of the vote. So understanding what happens to seat share as vote share increases or decreases across the full spectrum of potential partisan splits is crucial to understanding a party’s success in the legislature. Therefore, the benefit of the computational model is that it generates a number hypotheses about what exactly happens to a party’s electoral success when it gains or loses votes under various degrees and types of partisan clustering.

3.5 Hypotheses

The model tells us about what would happen to the electoral success of a party as partisan clustering increases or decreases, as the type of partisan clustering changes, and as a party’s vote share moves from a majority to a minority. Moreover, it gives insight about how each of these variables interact with each other to produce different electoral results. Therefore, the expectations generated by the model are numerous. However, for simplicity I’ve reduced the expectation to the following five general hypotheses about how Democratic clustering affects the votes-seats curve:

3.5.1 Clustering Reduces Majority Party Seat Share

H1: The geographic clustering of partisans has the effect of reducing the seat share for the party that wins the majority of the vote.

Formally, if a party’s vote share, $V_P$, is linked to its seat share, $S_P$, through the function
$f_T$, where $T \in \{\text{unif, symm, asym}\}$ indexes the function generated by each of three types of clustering, then for both parties, $P \in \{\text{Dem, Rep}\}$, it is the case that $\int_{\frac{1}{5}}^{1} f_{\text{symm}}(V_P) - f_{\text{unif}}(V_P) < 0$ and $\int_{\frac{1}{5}}^{1} f_{\text{asym}}(V_P) - f_{\text{unif}}(V_P) < 0$.

This means that partisan clustering - regardless of the type - flattens the votes-seats curve. This means that the expected number of seats generated by a majority share of the votes is reduced by the clustering of partisans. Both Republican and Democratic majorities would weakly prefer the less clustered environment than the clustered environment, as long as they both hold non-marginal majorities of the vote. For example, from the second plot in Figure 6, we can move along the x-axis to observe the outcome of a hypothetical election where a party receives 55% of the vote both in the absence of geographic clustering (the blue line) and in the presence of geographic clustering (purple and pink lines). Comparing the two, we see that when the Democratic party holds 55% of the vote, moving from a non-clustered environment to a clustered environment will reduce the seat share of Democrats. Similarly, if the Republican party holds 55% of the vote, clustering will reduce the seat share of Republicans. Formally, the model implies that $f_{\text{symm}}(.55) - f_{\text{unif}}(.55) < 0$ and $f_{\text{asym}}(.55) - f_{\text{unif}}(.55) < 0$.

We can see this effect in the plots of Figure 7. The plots capture the effect of partisan concentration on majority party seat share for both Republicans and Democrats when they hold 55% of the vote. The downward sloping lines show the change in seat share when the underlying geographic distribution of partisans moves from a uniform to a clustered environment. While the first plot captures the effect of symmetric clustering, the second plot captures the effect of asymmetric clustering. And in both cases, clustering reduces the seat share of the majority party.

Observing the blue line in both plots, it is clear that the effect of clustering on the majority party is that it reduces the number of seats for Democratic majorities. Democratic votes get packed into too few districts, thus having the effect of wasting votes in districts.

---

8This also assumes that the parties prefer a greater share of the seats and does not consider a preference for incumbency protection
Figure 7: The figure above shows how the geographic clustering of Democrats changes the seat share of a party with 55% of the vote. Plot A gives the effect of symmetric clustering, while plot B gives the effect asymmetric clustering. The red line signifies the effect when Republicans hold the majority, while the blue line shows the effect when Democrats hold the majority. The implication of both figures is that clustering disadvantages both parties. However, in the presences of asymmetric clustering.
that are already overwhelmingly Democratic. However, contrary to Chen and Rodden, as the red line implies, when the Republican party has a similar 55% vote share, they too receive a disadvantage from clustering. In which case, clustering has the effect of improving Democratic seat share against the Republican majority.

However, as the second plot of Figure 7 makes clear, asymmetric Democratic clustering affects the representation of Republicans and Democrats differently. While both parties lose seats due to clustering when they hold the majority of the vote, Democrats lose more seats than Republicans. Therefore, asymmetric clustering has asymmetric partisan consequences. This leads to the second hypothesis.

**H2:** Asymmetric clustering - where Democrats cluster more than Republicans - has the effect of reducing the seat share for Democratic voting majorities more than it does for Republican voting majorities.

Formally, under asymmetric clustering, the seat loss for Democratic majorities is greater than the seat loss for Republican majorities, \( \int_{1/2}^{1} f_{asym}(V_{Dem}) - f_{unif}(V_{Dem}) < \int_{1/2}^{1} f_{asym}(V_{Rep}) - f_{unif}(V_{Rep}) \). Asymmetric Democratic clustering has the effect of bending the votes-seats curve such that the return in seats for any given share of the vote will differ depending on the party that receives that share of the vote. In particular, when Democrats hold a majority of the vote, asymmetric Democratic clustering will reduce the number of Democratic seats more than it would reduce the number of Republican seats for Republicans holding the same majority. This effect is displayed in the second plot of Figure 7. If the majority party held 55% of the vote, asymmetric Democratic clustering would reduce the seat share for Democratic majorities more than Republican majorities. Therefore, the model implies \( f_{asym}(0.55_{Dem}) - f_{unif}(0.55_{Dem}) < f_{asym}(0.55_{Rep}) - f_{unif}(0.55_{Rep}) \). Therefore, asymmetric Democratic clustering will penalize Democratic majorities more than Republican majorities.
3.5.2 Clustering Reduces Electoral Responsiveness

As Democrats continue to locate in dense urban areas, not only will the majority vote share translate into fewer seats, but there will also be a reduction the number of seats gained per additional vote. In other words, the slope of the curve will flatten, reducing what Tufte (1973) called the “responsiveness” of the electoral system. A flatter curve means that a party’s seat share is less “responsive” to a change in vote share. We see that under a clustered environment, a gain in votes translates into fewer additional seats than in a non-clustered environment. For example, in New York, the density with which metropolitan New York City Democrats cluster tend to waste valuable Democratic votes that could be used to win districts elsewhere. Moreover, as Democrats continue to improve their vote share in a state like New York, their return in seats for every additional vote is much lower if the votes continue to be urban votes. As long as their votes join the cluster of urban voters, their seat gain will be small compared to a less clustered environment. This leads to the third hypothesis.

H3: Clustering reduces the electoral “responsiveness” for both parties by reducing the slope of the curve. Therefore, as clustering increases, a unit increase in vote share returns fewer additional seats.

We can see from the model that a party gains fewer seats under a clustered environment than a non-clustered environment when it moves from a 50% vote share to a 55% votes share. Formally, this can be expressed as

\[ f_{symm}(.55P) - f_{symm}(.50P) < f_{unif}(.55P) - f_{unif}(.50P) \]

\[ f_{asym}(.50P) - f_{asym}(.55P) < f_{unif}(.55P) - f_{unif}(.50P) \]

Moreover, if the type of clustering is asymmetric, there are different consequences for the two parties. In particular, the seat share achieved by Democratic majorities will be
less responsive to an increase in vote share than Republicans in the same position. This asymmetric effect is stated in the fourth hypothesis.

**H4:** Asymmetric Democratic clustering reduces the electoral “responsiveness” of Democratic majorities more than it does for Republican majorities.

In an asymmetrically clustered environment, Democrats have less to gain from increasing their vote share than Republicans do in similar positions. We can see from the model that a party gains fewer seats under a clustered environment than a non-clustered environment when it moves from a 50% vote share to a 55% vote share. Formally, this can be expressed as

\[ f_{asym}(0.55_{Rep}) - f_{asym}(0.50_{Rep}) > f_{asym}(0.50_{Dem}) - f_{asym}(0.55_{Dem}) \]

The asymmetry of the votes-seats curve produced by asymmetric Democratic clustering reduces the slope of the curve for Democratic majorities more than it does for Republican majorities. Therefore, in states where Democrats hold a majority of the vote (and given that there are still seats to be won), asymmetric Democratic clustering has the effect of reducing the efficiency with which additional votes translate into additional seats. As a result, the asymmetric residential patterns create conditions where Democratic majorities have less of an incentive to win additional votes per seat than Republican majorities.

### 3.5.3 Asymmetric Clustering Produces Pro-Republican Bias in Close Elections

Finally, as Chen and Rodden (2013) find, the results of the model suggest that in close elections, the asymmetric clustering of Democrats improves Republican seat share over the non-clustered counterfactual. This allows Republicans to achieve a majority of the seats without a majority of the vote. Therefore, this gives the fifth hypothesis generated by the model.
H5: In close elections, asymmetric clustering reduces Democratic seat share, allowing Republicans to gain a majority of the seats without a majority of the votes.

The final hypothesis suggests that asymmetric clustering produces electoral bias toward Republicans. In Figure 6 we can see this bias by observing how the pink line passes below the 50%-50% intercept. Therefore, in states where Democrats are asymmetrically clustered, Republicans can achieve a majority of the seats without a majority of the vote. Formally, \( f_{\text{asym}}(.50_{\text{Rep}}) > .50 \).

3.6 Policy Implications of the Model

Given that map-makers are constrained by geographic principles like contiguity and compactness, the model reveals a number of predictions for how residential patterns of partisan voters affect partisan representation in the United States. These predictions provide important policy implications for both the Democratic and Republican Parties. Because of this, parties must pay attention to how their partisans are located spatially. And they must pay attention to how partisan platforms can affect the spatial distribution of their voters, especially in light of the continuing spatial divide between Republicans and Democrats.

Where Democrats have clustered asymmetrically, they can find themselves with a number of challenges. First, Democrats will have trouble winning the majority of seats with marginal majorities of the vote. Unlike the other forms of clustering, the asymmetric clustering causes the votes-seats curve to dip below the 50%-50% intercept, biasing the system in favor of Republicans. Therefore, in competitive elections, the asymmetry of partisans causes Democrats to inefficiently waste votes by packing most of their partisans into a few districts and spreading the rest too thinly across the other districts. The result is that Republicans gain a natural seat advantage in the legislature.

We see this in states like Florida and Pennsylvania, where toss-up elections will naturally produce Republican congressional delegations. Because districts in Miami and Philadelphia produce overwhelmingly Democratic majorities while districts in the periphery produce
marginal Republican majorities, Democrats lose seats that they might not otherwise have lost. For example, although Obama won 54% of the vote in Pennsylvania in 2008, a majority (10) of the 19 districts supported McCain over Obama. The asymmetry in the distribution of votes across the districts was clear. All but two of the McCain-majority districts supported McCain with less that 55% of the vote while only two of the Obama-majority districts supported Obama with less than 55% of vote. In McCain’s two most supportive districts, he won 56% and 63% respectively. While in Obama won nearly 90% of the vote in both of his most supportive districts located in the heart of Philadelphia. The asymmetric distribution of the votes was at the root of Pennsylvania’s Republican delegation.

A second consequence of asymmetric clustering is that, like symmetric clustering, it reduces the majority party’s seat share as well as the rate at which the majority party can pick up seats with additional votes. However, unlike symmetric clustering, it has the effect of disadvantaging Democratic majorities more than it does Republican majorities. Therefore, where Democrats receive voting majorities, asymmetric residential patterns will be detrimental to the party’s legislative seats share. Democratic majorities will be worse off than Republicans in the same position and they also have far less to gain by increasing their popular support.

Therefore, in Democratic states Democrats have the incentive to reduce the spatial clustering of their votes. The overwhelming urban support for Democrats does little to advance Democratic seat share in these situations. Instead, Democrats would prefer to swap urban votes for non-urban votes. By improving their support in the periphery, they are more likely to win back the marginally Republican seats in the state. As a result, there is much to gain from de-clustering.

On the other hand, when Democrats find themselves in a state where they hold the minority of the vote, they can make partisan gains from doing the opposite. In these cases, Democrats can use the spatial clustering of partisanship to their advantage. When Republicans hold a strong majority, clustering tends to help save Democrats from steep losses to
their seat share. Therefore, to improve Democratic seat share, Democrats can target policy that favors metropolitan areas in Republican states. This can have the effect of producing seat gain where they might not otherwise receive it. By packing Democrats into a single location, Republicans are simply unable to break the Democratic majorities without explicitly drawing oddly-shaped, non-compact districts. Strong metropolitan support, therefore, advances Democratic seat share when statewide support is generally low. We see this effect in Texas, for example, where Democrats not only pick up seats from densely packed Democratic voters in metropolitan cities like Dallas, Houston, and San Antonio, but the strongly Democratic Texas South, where the Hispanic vote is difficult to dilute even without VRA protections. The clustering of the Democratic vote in Texas is potentially explains why Democrats receive at least a third of the seats in Texas despite a significant minority of the vote.

Moreover, the model reveals another potential solution for Democrats to improve their representation in these Republican states. The solution is to induce symmetry within the spatial distribution of the partisan vote. While asymmetric clustering is an improvement for Democrats compared to a non-clustered environment, it does not deliver the same return in seats as symmetric clustering does. In other words, the non-urban Democrats that join the Republicans on the periphery - which is the cause of the asymmetric clustering - have the effect of improving the ability for Democrats to win seats in states where Republicans hold a majority. But compared to more symmetric clustering, asymmetry is less efficient at doing so. This is because the Democrats that live outside of urban areas do not have the numbers to win non-urban districts. Therefore, these votes would be more useful to Democrats if they were densely concentrated in metropolitan cities, where they might be able to pick up another seat through packing. So, in states where Republicans are in the majority, packing has beneficial consequences for Democrats. They would be better off creating policy that further polarizes voters along geographic dimensions in these states. So, just as Democrats lose seats due to the geographically polarized environments when they are in Democratic
states, they gain seats in the same way when they are in Republican states.

An additional consequence of asymmetric clustering is that it has an asymmetric effect on the slope of the votes-seats curve. Because the slope of the curve is lower when Democrats hold the majority, it is much more difficult to gain additional seats from additional votes. On the other hand, when Democrats hold the minority of the vote, there is the opposite effect. The slope of the curve is much greater, which means an additional vote translates into a greater number of additional seats. Therefore, there is greater incentive to target supporters in states with Republican majorities. By increasing vote share in these states, Democrats can make great improvements to their seat share in the legislature.

In either case, the model generates the following expectation: Even in the face of redistricting, partisan clustering affects how partisans are distributed across legislative districts in such a way that any party holding a non-marginal majority of the vote will likely lose legislative seats. This includes the Republican party.

4 Conclusion

Obtaining a greater share of the vote is a major goal for parties. Yet the efficiency with which votes translate into seats is perhaps of greater concern. This is why the votes-seats curve is important to scholars. Since it is theoretically possible to achieve a wide ranging share of the seats given any share of the vote, parties will always be interested in improving their votes-seats relationship. Changing the curve in a way that plays to a particular party’s advantage means altering the playing field on which parties compete for power. The consequences can be major. Therefore, it is in the best interest for parties to attempt of the change this curve if they are able to.

Because of this incentive, when partisan favor is observed in an election, many point the finger at gerrymandering. By manipulating district boundaries, parties are effectively able to change the votes-seats relationship. And since political parties are often unchecked in the districting process, the accusations of gerrymandering are common. Simply observing a
votes-seats curve that appears favorable to one party over another triggers the assumption that the partisan favor was intentionally designed by the boundary-makers.

However, recent research by Chen and Rodden (2013) suggests that partisan favor might not be the work of parties at all. Instead, the favorable votes-seats curve might be attributable to the geographic distribution of partisan voters. They argue that boundary-makers are constrained by geographic principles like compactness and contiguity when constructing districts. And because they adhere to these principles, they allow for the geographic distribution of voters to influence the votes-seats relationship. They find that the tendency for Democrats to cluster their vote in dense urban areas, gives Republicans a natural seat advantage in legislatures, even in the presence of redistricting. In particular, they suggest that given a 50-50 split of the vote, geographic clustering would tip the scales to favor Republicans.

Yet, despite their findings, it remained unclear how the clustering of Democrats influenced the full votes-seats relationship. This is because we are unable to observe elections at every split of the vote under various forms of partisan clustering. However, by simulating elections using an agent-based model to segregate partisan voters, I was able to develop expectations about how different types of geographic clustering affects the votes-seats curve across the full spectrum of votes. This provides a set of predictions about the consequences of partisan residential patterns that have been previously unexplored.

For example, assuming that districts are drawn to uphold the geographic principles of contiguity and compactness, the model finds that when Democrats cluster asymmetrically with respect to Republicans, the votes-seats curve bends in an asymmetric fashion. The asymmetry of the curve tends to disadvantage Democrats by reducing their seat share for most splits in the two-party vote. However, when Republicans hold a significant majority of the vote, Democratic clustering can have the opposite effect. Instead of disadvantage Democrats, clustering can disadvantage Republicans. In states with strong Republican support, the geographic concentration of Democrats forces Democratic seats that might not
otherwise have existed. In states like Tennessee, for example, it is difficult to design a set of nine compact and contiguous districts without drawing Democratic districts in Memphis and Nashville. The very nature of the residential patterns of partisans in these two metropolitan cities gives Democrats a boost in their seat share, effectively reducing Republican representation.

The findings of the model provide an important connection between partisan geography and representation. Simply because legislative districts are linked to geographic jurisdictions, there are major representational consequences to something that is so seemingly inconsequential to politics as residential location. As partisanship divides along the urban-rural line, for example, the partisan composition of the legislature can change without changing the overall partisan vote. Representation is therefore conditional on the geographic manner in which partisans differentiate themselves rather than the extent of partisan support. Because of this, scholars must pay closer attention to how partisan divisions translate into geographic divisions.
References


McDonald, Michael. 2010. *Midwest Mapping Project*. George Mason University, Department of Public and International Affairs.


