ABSTRACT: When one of the major parties in the United States wins a substantially larger share of the seats than its vote share would seem to warrant, the conventional explanation lies in overt partisan or racial gerrymandering. Yet this paper uses a unique data set from Florida to demonstrate a common mechanism through which substantial partisan bias can emerge purely from residential patterns. When partisan preferences are spatially dependent and partisanship is highly correlated with population density, any districting scheme that generates relatively compact, contiguous districts will tend to produce bias against the urban party. We apply automated districting algorithms driven solely by compactness and contiguity parameters, building winner-take-all districts out of the precinct-level results of the tied Florida presidential election of 2000. The simulation results demonstrate that with 50 percent of the votes statewide, the Republicans can expect to win around 59 percent of the seats without any “intentional” gerrymandering. This result occurs because urban districts tend to be homogeneous and Democratic while suburban and rural districts tend to be moderately Republican. We show that these results generalize across many industrialized US states with significant urban concentrations.
In Florida’s tied presidential election of November 2000, Bush voters outnumbered Gore voters in 60 percent of Florida’s 40 state senate districts, and Republican legislators have consistently controlled around 60 percent of these seats over the last decade. More broadly, while recent statewide races for governor, senator, and U.S. president have been extremely close in pivotal states like Florida, Michigan, Ohio, and Pennsylvania, Republicans have consistently dominated both chambers of the legislatures. This disconnect between statewide partisanship and election outcomes has been noticed by academics (King and Gelman 1991, Hirsch 2003, McDonald 2005) as well as pundits and party strategists.

Conventional wisdom holds that by “packing” left-wing voters into a few heavily Democratic districts, intentional efforts at racial and partisan gerrymandering allow Republicans to win a supermajority of the remaining legislative seats. Cartographers with partisan motives can draw district lines so as to “pack” and “crack” the support of their opponents. Such partisan gerrymandering is the focus of a large theory literature (e.g., Gilligan and Matsusaka 1999; Gul and Pesendorfer 2009) as well as empirical studies (e.g., Abramowitz 1983; Cain 1985; Cox and Katz 2002; and Herron and Wiseman 2008). Likewise, a common argument emerging from formal theory (e.g. Shotts 2001, 2003) and empirical studies (e.g., Brace, et al. 1987, Hill 1995, Lublin 1997, Cameron, Epstein and O’Halloran 1996, Grigg and Katz 2005) is that efforts at enhanced minority representation tend to “pack” Democrats into relatively few districts.

While these sources of electoral bias often receive the lion’s share of attention in popular and academic discussions, electoral bias can also emerge from so-called “unintentional gerrymandering,” whereby one party’s support is more geographically clustered than that of its opponent due to a combination of residential choices and natural and human geography (Dixon 1968, Johnston and Hughes 1978). In the United States, a potentially important source of electoral bias comes from the fact that outside of the Deep South, Democrats have been highly concentrated in densely populated
areas since the industrial revolution and Great Migration (Erikson 1972, 2002; Jacobsen 2003, McDonald 2009a, 2009b).

A large reform movement in the United States is predicated on the notion that most observed electoral bias stems from intentional gerrymandering. Districting reformers in many states have advanced various statutory and constitutional proposals to prohibit partisan gerrymandering and enforce more neutral, objective criteria and procedures in the redistricting process. In Florida, for example, in response to a striking pattern of pro-Republican electoral bias, a coalition of left-wing interest groups invested significant energy and resources into passing Amendments 5 and 6, which voters approved in November 2010. These ballot initiatives mandate that newly drawn congressional and state legislative districts be compact and contiguous in shape, and the initiatives prohibit redistricting plans drawn with the intent to favor either political party.

These reforms will have little impact on electoral bias if it turns out that much of the bias observed in recent years actually stems from residential patterns rather than intentional racial or partisan gerrymandering. Thus an important challenge with far-reaching policy implications is to clarify the logic of unintentional gerrymandering in the United States, and to devise and empirical strategy for disentangling its impact from that of intentional partisan and racial gerrymandering.

This paper takes up that challenge in three steps. In the first section, we describe a common American residential pattern in which Democrats are highly clustered in dense central city areas, as well as in smaller 19th century railroad agglomerations, while Republicans are scattered more evenly through the suburban, exurban, and rural periphery. We demonstrate this pattern with an in-depth case study of Florida, but also show that it holds up in many other industrialized states. Precincts in which Democrats typically form majorities tend to be more homogeneous and extreme than Republican-leaning precincts. More importantly, when these Democratic precincts are combined with neighboring precincts to form districts, the nearest neighbors of extremely Democratic precincts are more likely to be similarly extreme than is true for Republican precincts.
The second section then isolates and measures the electoral bias caused by this asymmetric residential concentration of voters, as distinct from that caused by the intentions of cartographers. It does so by conducting repeated computer simulations of the legislative districting process. Using precincts as building blocks, our computer simulations draw districts randomly and without regard to any partisan, racial, or other political criteria. Instead, our simulation algorithm uses only the traditional districting criteria of population equality, contiguity, and compactness.

Since the early 1960s, scholars have suggested automated districting as a solution to the problem of partisan gerrymandering (e.g., Vickrey 1961, Weaver and Hess 1963, Nagel 1965). More recently, scholars have used hypothetical districting experiments to examine partisan polarization (McCarty, Poole, and Rosenthal 2009), partisan representation (Altman 1998), and the impact of various districting criteria (McDonald 2009). These previous studies have often used automated redistricting in order to obtain a baseline against which to detect the intentions of those drawing the lines. Cirincione, Darling, and O’Rourke (2000) use a simulated districting algorithm to detect whether South Carolina's congressional districting plan constituted a significant racial gerrymander, while Altman and McDonald (2004) propose an enhanced method of this algorithm for detecting partisan gerrymandering. In a similar spirit, we use simulations to examine the electoral effects of hypothetical districting plans drawn without regard to race or partisanship.

We focus in detail on simulations using precinct-level vote counts from the November 2000 presidential race in Florida. Because this contest produced a virtually tied outcome between George Bush and Al Gore, this election presents a unique opportunity to compare the partisanship of our simulated legislative districts against the baseline of the 50-50 partisan split of the statewide presidential vote. Our simulations reveal that, although Bush and Gore split the statewide vote evenly, Bush received a majority in 58-61% of the simulated districts, a finding that suggests significant Republican electoral bias in Florida, even with our seemingly apolitical, random districting procedure.
In the third section, we conduct similar simulations using a larger group of states. Upon conducting thousands of random simulations across several states, we find that the Republican share of legislative seats in our simulated districting plans strongly predicts the actual Republican seat share observed in the real life legislatures. Using this method, we are able to identify various industrialized states, such as Florida and New York, which exhibit relatively high underlying pro-Republican bias. We demonstrate that the highest levels of such bias occur in states where Democratic voters are relatively more concentrated in urban areas than Republican voters.

The contributions of the paper are threefold. First, we add depth and detail to a classic political geography literature in which electoral bias can emerge purely from residential patterns in a two-party system because one party’s support is “naturally” more packed within districts (e.g. Gudgin and Taylor 1979, Johnston 1976). Second, we demonstrate an empirical technique for isolating this form of electoral bias and distinguishing it from that created by intentional gerrymandering. Finally, having done so, we are able to extract policy implications related to redistricting reform: in urbanized U.S. states, the abolition of overt partisan or racial gerrymandering is not likely to eliminate electoral bias.

1. Political Geography and Electoral Bias

1.1. The Geographic Distribution of Democrats and Republicans

Electoral maps from recent U.S. presidential elections communicate rather clearly that in much of the United States, support for Democrats is highly clustered in densely populated city centers, declines gradually as one traverses the suburbs and exurbs, and levels off in largely Republican rural areas. An additional observation is that in the rural periphery, there are scattered pockets of strong support for Democrats in smaller agglomerations associated with 19th century industrial activity along railroad lines, canals, and rivers, as well as in college towns.

[FIGURE 1 ABOUT HERE]
In order to display the relationship between population density and voting behavior, we match precinct-level results from the 2000 presidential election to precinct boundary files produced by the census department. Using GIS software, we generate block group estimates of election results, which we plot against population density data from the census in Figure 1.\(^1\) The relationship between population density and Democratic voting is striking in general, but there is some cross-state heterogeneity. The relationship is most pronounced in the most industrialized and urbanized states, including Florida, but it is less pronounced or lacking altogether in less industrialized Southern states with large rural African American populations and in relatively sparse Western states.

[FIGURES 2 AND 3 ABOUT HERE]

Note that the densely populated urban districts in the lower right corners of the scatter plots in Figure 1 are not randomly distributed in space; many of them are in close proximity to one another. Focusing on Florida, Figure 2 provides a map of results of the 2000 presidential election, where the centroid of each block group is represented as a dot the goes from blue to red according to the Bush vote share. This map demonstrates that support for Democrats (blue dots) was highly concentrated in downtown Miami and the other coastal cities to its immediate North, as well as downtown Orlando, Tampa, St. Petersburg, Daytona, Gainesville, Jacksonville, Tallahassee, and Pensacola, as well as a few other smaller cities. The suburbs of these cities, along with rural Florida, while largely leaning Republican, are relatively more heterogeneous in voter partisanship.

Figure 3 provides another way to visualize the spatial arrangement of partisanship in Florida. This figure displays the distance in kilometers between the center of Miami’s central business district and the location of every census block group in Florida. Figure 3 displays this distance on the horizontal axis, and the vertical axis displays the block group's Bush vote share. Block groups toward

\(^1\) The coverage of states was limited by our ability to match the precinct-level election results with GIS boundary files.
the right of this plot are further away from Miami, and the extreme right side of the plot depicts block
groups in the Florida panhandle. The lower left corner of the plot displays the large number of
overwhelmingly Democratic precincts in downtown Miami, Ft. Lauderdale, and Palm Beach. To the
north of these urban cores are more heterogeneous suburban neighborhoods where the Bush vote
share, on average, only slightly exceeds 50 percent.

The tips of each of the other “stalactites” in Figure 3 are city centers where Al Gore's vote
share in November 2000 often exceeds 90%. The counties in which Florida’s largest cities are
located are each marked with a separate color, and the identity of the largest city in the county is
indicated in the legend. In each case, as one moves outward from the city center, the Bush vote
increases, and each city is surrounded first by a very mixed area, second by a suburban periphery that
produced solid but not overwhelming support for Bush, and then finally by a rather heterogeneous
but moderately Republican periphery. This pattern can also appears clearly in the insets for Orlando,
Jacksonville, and Tampa in Figure 2. Graphs and maps like Figure 2 and 3 look extremely similar in
all of the other urbanized states of Figure 1 that are characterized by high correlations between
population density and voting.

Figures 2 and 3 illustrate two important patterns with consequences for districting. First,
Democrats are far more clustered within homogeneous precincts than are Republicans. For
example, while Bush received over 80 percent of the vote in only 80 precincts, Gore received over 80
percent in almost 800 precincts. Second, the clustering of Democratic block groups in Figure 2 and
the stalactite shape of cities and their surroundings in Figure 3 both illustrate that Democratic
precincts tend to be closer to one another in space than Republican precincts. That is, the nearest
neighbors of predominantly Democratic precincts are more likely to be predominantly Democratic
than is the case for Republican precincts.

Some simple spatial statistics allow us to move beyond mere visualization of this
relationship. First, we can identify the nearest neighbor of every precinct, defined as the precinct with
the most proximate centroid, and ask whether that neighbor has the same partisan disposition. For any reasonable cut-off used to differentiate “Democratic” and “Republican” precincts (e.g. lower than 40th vs. higher than 60th percentile values of Bush share, 30th vs. 70th, etc.), we find that indeed, the nearest neighbors of Democratic precincts are significantly more likely to be Democratic than is the case for Republicans, whose neighbors are more heterogeneous.

Alternatively, rather than forcing precinct partisanship to be binary, it is useful to examine the extent to which each precinct’s election results are correlated with those of its neighbors, and ask whether the extent of this “spatial association” is higher in Democratic than in Republican districts. Luc Anselin’s (1995) local indicators of spatial association (LISA) is well suited to this task. For each precinct, \( i \), the local index of spatial association is given by:

\[
I_i = \frac{Z_i}{m_2} \sum_j W_{ij} Z_j
\]

where:

\[
m_2 = \frac{\sum Z_i^2}{N}
\]

and \( Z_i \) is the deviation of Bush share with respect to the mean across all precincts, \( N \) is the number of precincts, and \( W_{ij} \) is a matrix of weights with ones in position \( i,j \) whenever precinct \( i \) is a neighbor of precinct \( j \), and zero otherwise. We define “neighbors” as precincts that share any part of any boundary.\(^2\)

[FIGURE 4 ABOUT HERE]

Overall, \( I \) is much higher for Democratic precincts than for Republican precincts. For instance, the average local spatial association for the most Democratic quartile of precincts is just

\[ \]

\(^2\) The results are very similar if we use other types of contiguity or distance-based spatial weights matrices.
over 2, while the index averages 1.16 for the most Republican quartile. To display these patterns more comprehensively, Figure 4 displays $I_i$ for each precinct using an extruded map, in which the height of each extrusion corresponds to the extent of spatial association, and the color moves from blue to red as the precinct's Bush vote share increases. Figure 4 illustrates clearly that the most Democratic precincts in Florida’s city centers are also those with the highest levels of local spatial association; that is, they are surrounded by other very Democratic precincts. While there are some Republican-leaning areas of high spatial association in little Havana, suburban Jacksonville, and the Panhandle, Republican precincts overall tend to be located in more heterogeneous neighborhoods.

1.2. Implications for Electoral Bias

Whether done by incumbent politicians, districting boards, or computer algorithms, the process of building electoral districts involves stringing together contiguous census blocks. Our automated simulation algorithm takes seriously the rhetoric of reform advocates by implementing a districting process in which these census blocks are assembled without political or racial manipulation. To illustrate, consider the process of randomly selecting one of the dots in Figure 1 and randomly connecting it with surrounding dots until enough dots have been selected to form a state legislative district or Congressional district.

This process is likely to undermine the representation of Democrats for three reasons. First, suppose that the initial seed is a precinct in one of the Figure 3 “stalactites” representing Florida’s large cities, such as Miami, Jacksonville, or Tampa. Then such a city is sufficiently large that this process will likely combine extremely Democratic districts with other extremely Democratic districts, thereby forming a district that is “packed” with Democrats. Second, outside of little Havana, it is difficult to find a Florida precinct that, when randomly chosen as the initial seed, would produce
an analogously extreme-Republican district. In addition to being more internally heterogeneous, Republican precincts tend to be located in heterogeneous suburban and rural areas of the state. For instance, suppose the initially chosen precinct is a rural and extremely pro-Republican. If one simply strings together neighboring precincts until reaching the population threshold for a district, this will usually require the inclusion of some rather heterogeneous precincts, often including pockets of Democrats in small cities and on the fringes of larger cities.

A third reason concerns the locations of small Democratic-leaning towns throughout Florida. While some medium-density pro-Democratic districts might conceivably come together to form Democratic districts along the Eastern Coast, there are also small and isolated pockets of strong Democratic support in the medium-density 19th century railroad towns and manufacturing centers, such as Ocala or Pensacola, and the college towns of Tallahassee and Gainesville. When the size of districts is large relative to these small clusters of Democrats, these towns are likely to be subsumed into predominantly rural districts with moderate Republican majorities, hence "wasting” Democratic votes.

In short, the complex process of migration, sorting, and residential segregation that generated the spatial distribution of partisanship in Florida has left the Democrats with a more geographically concentrated support base than Republicans. When compact, contiguous districts are imposed onto this geography, the result will be a skew in the distribution of partisanship across districts. This skew is at the heart of the electoral bias we document in this manuscript.

2. Automated Districting and Electoral Bias

Studies of electoral bias typically flow from the normative premise that a party with 50 percent of the votes should receive 50 percent of the seats. Empirical studies use either aggregate data over several elections, or transformations of district-level data from individual elections, to examine the seat share that would be obtained by the parties under a hypothetical scenario of a tied
election. Our goal is different. Rather than examining the bias associated with existing districting plans, which were undoubtedly influenced by efforts at partisan and racial gerrymandering, we seek to estimate the electoral bias that would emerge under hypothetical districting plans that are not intentionally gerrymandered.

Rather than using information from existing districts to simulate hypothetical tied elections, we use information from the precincts of a real-life tied election, and we simulate a variety of districting plans. We perform a large number of automated, computer-based simulations of legislative districting plans. Our computer simulations construct these districting plans in a random, partisan-blind manner, using only the traditional districting criteria of equal apportionment and geographic contiguity and compactness of single-member legislative districts. For each of these simulated districting plans, we calculate the Bush-Gore vote share of each simulated single-member district, and we use this vote share to predict whether the district would have been a Democratic or Republican seat.

Because of the virtual 50-50 Bush-Gore tie in Florida, an unbiased partisan division of Florida’s legislative seats would result in approximately 50% of the seats being Republican, defined as any seat having a pro-Bush majority. In other words, we are using the distribution of Bush-Gore (Republican-Democrat) vote shares across the simulated Florida districts as a measure of electoral bias.

2.1 The Automated Districting Algorithm

As of the November 2000 election, Florida consisted of 6,045 voting precincts. These precincts are the smallest geographic unit at which election results are publicly announced, so we use the precinct as the building block for our simulations. Hence, a complete districting plan consists of assigning each one of Florida’s precincts to a single legislative district. Florida voters cast 5.96
million Presidential election ballots in 2000, so the average precinct cast a total of 986 presidential votes.

We perform our automated simulations using the legislative districting algorithm presented by Cirincione, Darling, and O’Rourke (2000). These authors performed computer simulations of South Carolina’s congressional districting to show that the state’s actual redistricting plan exhibited significant racial gerrymandering. More importantly, for our purposes, Cirincione et al. (2000) show that their districting algorithm guarantees equal apportionment of population across all legislative districts while substantially achieving geographic contiguity and compactness for nearly all simulated districts. Furthermore, these simulated districts are drawn without regard to either voter partisanship or any demographic information other than simple population counts. Hence, the simulation algorithm is designed to be a partisan-neutral and race-blind districting process, using only traditional geographic criteria.

We implement’s Cirincione et al.’s (2000) automated districting algorithm as follows: Suppose we wish to divide Florida into \( m \) number of single-member legislative districts, where \( m \geq 2 \). First, we select one precinct at random and assign it to the first district. Next, we randomly select and add one of the precincts that borders the initially-chosen precinct. We continue building up this first district by adding more bordering precincts until the emerging district contains \( 1/m^{th} \) of the state’s total population. Before we add each additional precinct, however, we first construct the smallest bounding box\(^3\) that encloses all of the existing precincts of the emerging district. When randomly selecting the next precinct for the district, we first randomly choose among those bordering precincts

\(^3\) Specifically, the bounding box is defined by the four directional (ie, east, north, etc.) extremes among the centroids of the precincts already assigned to the district.
that are already located within the bounding box. Only if the bounding box contains no unassigned precincts do we randomly select among bordering precincts located outside of the box.

Once the first district is fully apportioned, we begin construction of the second district by randomly selecting a precinct among those bordering the first district. The identical process begins anew, except that precincts assigned to the first district cannot be assigned to any further districts. We repeat this process until all $m$ districts have been fully constructed.

Our use of precincts as the building blocks of districting plans introduces the possibility of slightly over or under-apportioned districts, and we address this problem by introducing a simple assumption allowing our simulation algorithm to split precincts. Suppose that an emerging district is currently just below the target population size – that is, it contains just under $1/m$th of the state’s total population. But the addition of one new precinct would increase the district’s population well over the target size. To remedy this problem, we split up the new precinct by assigning just enough randomly selected voters from the precinct to our emerging district. The remaining unassigned voters are grouped together as a precinct to be assigned to a later district. Hence, in implementing this remedy, we are effectively assuming that all voters within a precinct are geographically contiguous with one another. This remedy also allows us to simulate districting plans that contain more districts than the total number of precincts in Florida.

Once we have divided all of Florida up into $m$ districts, the districting simulation is complete. After completing this districting simulation, we aggregate the precinct-level Bush-Gore vote counts within each district, and determine whether each of the $m$ districts is a Republican (pro-Bush) or a Democratic (pro-Gore) seat.

We repeat a simulation of this sort for many different hypothetical legislative sizes, ranging from a legislature of two districts to a legislature of 100,000 districts. For each legislative size, we repeat the simulation procedure 200 times, constructing an independent districting plan each time. For example, we conduct 200 independent simulations dividing Florida into 100 districts; hence, this
set of simulations constructs a total of 20,000 districts, of which 11,506 (57.9%) are Republican seats.

To evaluate the accuracy of our simulation procedure, we conduct the same set of Florida districting simulations using the Better Automated Redistricting software created by Altman and McDonald (2009), which includes an implementation of the Cirincione et al (2000) algorithm. Using the Altman and McDonald software for districting plans in which Florida precincts were combined into a reasonable (2 to 200) number of districts, we obtained results that were virtually identical to those reported below.

2.2 Simulation Results

Our simulations reveal significant pro-Republican bias in the partisan distribution of seats in any realistically sized legislature; that is, significantly over one-half of the legislative seats have Republican majorities. Figure 5 summarizes the distribution of seat shares produced under our simulations. In this figure, the horizontal axis represents the number of single-member districts in each simulated plan. The vertical axis reports the average percentage of these districts that have Republican majorities. For each different hypothetical legislative size, the dot represents the mean, district-level Bush vote share across the simulated districts, and the vertical line represents a 95% confidence interval. The Figure illustrates, for example, that when we conducted 200 independent simulations of dividing Florida into 100 districts, Republicans won an average of 57.9% of the seats, with a confidence interval of 57.2 to 58.6%. Overall, this plot illustrates the significant pro-Republican bias that results from the districting of the legislature based solely on the traditional principles of geographic contiguity, compactness, and equal apportionment.
Why does this significant pro-Republican bias arise in our districting simulations? Figure 6 illustrates the distribution of district-level Bush vote shares that emerges when we repeatedly simulate dividing Florida into 10 districts. This histogram reveals that Republicans win well over one-half of the seats because of the pattern we described earlier: Democratic voters tend to be clustered in heavily left-leaning precincts, so the Democratic party’s electoral base is concentrated in a relatively smaller number of urban-based districts. The Republicans’ electoral base, by contrast, is geographically spread throughout the moderately right-leaning hinterlands. As a result, for most reasonable legislative sizes, the distribution of seats across the state consists of a large number of moderately Republican districts in the rural and suburban areas and a relatively smaller number of more extreme Democratic, urban districts. Too many left-wing voters are wasted in urban, landslide Democratic districts, so the overall seat share across the state favors the Republicans.

The plot in Figure 5 details how this pro-Republican bias increases as the legislature grows in size from two to eight districts. A legislature consisting of only two single-member districts will always have exactly one Democratic and one Republican seat, a result that follows naturally from Florida’s 50-50 Bush-Gore vote share. But as the legislature grows in size, the partisan division of legislative seats begins to favor the Republicans. When the simulated legislature has eleven seats, Republicans win an average of nearly 66% of the districts.

As the size of the legislature increases beyond eleven seats, some of the medium-density Democratic clusters in suburbs and small towns that had previously been subsumed in their surrounding Republican peripheries begin to win their own seats, and thus the Republican seat share slowly declines. However, a striking result is that the Republicans always continue to control over one-half of the total seats. In fact, this pro-Republican bias never fully disappears until the size of the simulated legislature becomes unrealistically large. As the hypothetical legislature grows in size to several million seats, we approach the equivalent of a direct democracy in which each voter represents only himself or herself in the legislature, and the partisan seat share is identical to the
underlying population’s overall partisanship by definition. Our simulation results in Figure 5 reflect this approach toward direct democracy as the hypothetical legislature becomes extremely large: As the simulated legislature grows to several thousand districts, the pro-Republican bias begins to disappear, and the Republican share of total legislative seats approaches 50%.

Nevertheless, for any districting plan of realistic size, the pro-Republican bias exhibited in our simulations is significant. Florida’s state Senate and House chambers consist of 40 and 120 single-member districts, respectively, and the Congressional delegation is divided among 25 districts. Our simulations demonstrate that for these legislative sizes, Republicans should control an average of 58-61% of the seats statewide. The confidence intervals for these estimated average seat shares rule out the null hypothesis of no electoral bias.

2.3 Political geography and the simulation results

Why do Republicans win such a disproportionate share of these 25 districts? The simulations results can help shed further light on the mechanism through which the political geography of partisanship generates electoral bias. In Figure 7, we present the results of 200 independent random simulations in which Florida was divided into 25 districts.

[FIGURE 7 HERE]

Each plotted point in Figure 7 represents one of Florida’s 6,045 precincts, and we plot high, medium, and low density precincts separately, referring to them loosely as urban, suburban/town, and rural. For each plotted point, the horizontal axis measures the partisanship of the precinct, as measured by Bush-Gore vote share in November 2000. The vertical axis measures the average partisanship of the 200 simulated districts to which the precinct was assigned during our simulations.

Overall, given the patterns of spatial association reported above, we are not surprised to see a generally positive correlation between the partisanship of a precinct and the partisanship of the legislative district to which the precinct was assigned. In other words, pro-Bush precincts are
typically assigned to pro-Bush districts. In particular, the top and middle plots reveal that outside of
dense city centers, pro-Bush precincts were almost always assigned to majority-Bush districts. Note
that the lower-right quadrants of the plots—where pro-Republican precincts are assigned to majority
Democratic districts—are essentially empty.

In contrast, majority-Gore precincts outside of dense urban neighborhoods very often ended
up in the upper-left off-diagonal portion of these plots. In other words, while there are quite a few
rural, small-town, and suburban precincts where majorities favored the Democratic candidate, they
tend to be subsumed in moderately Republican districts. As described above, there are isolated
pockets of support for Democrats in college towns as well as blue-collar and/or African-American
enclaves in the suburbs of big cities and in small towns with a history of railroad industrialization.
However, these are generally surrounded by Republican majorities, and as a result, the Democrats are
poorly situated to win districts outside of the urban core.

By contrast, note that the bottom plot in Figure 7 illustrates that pro-Gore precincts in urban
areas were generally assigned to overwhelmingly Democratic districts during our simulations. There
is a large cluster of observations at the bottom of the lower left quadrant of the bottom graph, where
Democratic precincts end up in extremely Democratic districts, and few corresponding observations
in the extreme upper right of any of the plots. Taken together, these plots show that because of their
geographic support distribution, Democrats not only waste more votes in the districts they lose, but
they also rack up more surplus votes in the districts they win. These two phenomena explain the
rather extreme pro-Republican bias revealed by our simulations.

2.4 Simulations Using Alternative Florida Elections

A possible concern with our simulations is that, for a variety of reasons, Bush-Gore vote
shares from November 2000 may not be an accurate measurement of voter preferences among
Florida’s voting precincts. One reason for this suspicion is that the two parties may have employed
geographically asymmetric campaign strategies in 2000; for example, perhaps the Democrats targeted urban voters, while the Republicans targeted the hinterlands. Another reason for suspicion is that in November 2000, various non-presidential elections, such as local and Congressional races, may have affected voter turnout differently in Republican and Democratic regions of Florida. Moreover, we wish to make inferences about the roots of electoral bias in state legislative elections, and it is possible that presidential vote shares are of limited value if the state party system is sufficiently distinctive from the national party system.

To address these and other concerns about the possible uniqueness of the 2000 election, we show that our simulations produce a similar pro-Republican bias when we use alternative election results from different years and offices to measure the partisanship of simulated districts. Specifically, we re-conduct our legislative districting simulations using election results from the following Florida statewide races: 1) The 1992 Presidential election between Democrat Bill Clinton and Republican George Bush; 2) The 1994 Gubernatorial election between Democrat Chiles and Republican Jeb Bush; 3) the 1998 Gubernatorial race between Democrat MacKay and Republican Jeb Bush; and 4) the 2000 U.S. Senate race between Democrat Nelson and Republican McCollum. We choose these four races because in each election year from 1992 to 2000, these are the four races that produced the closest to a 50-50 split of the statewide two-party vote share. It is noteworthy that three of these are for statewide offices, and two are gubernatorial elections. Using each of these four sets of election results, we conduct a new set of 200 random districting simulations for each of a wide range of legislative sizes.

Overall, these new simulations, displayed in Appendix A, reveal a pattern of pro-Republican bias that is comparable to the electoral bias we find in Figure 5. In each election, for any reasonable
legislature size, the Republicans win significantly more than 50% of the simulated legislative seats, even though the underlying two-party split in each of the four elections is close to 50-50.4

2.5. Observed Electoral Bias in Florida

Next, we ask how the bias estimates from the simulations compare with electoral bias observed with actual districting plans. The most direct comparison is to simply aggregate precinct-level results of presidential elections to the level of state and Congressional legislative districts. The 2000 presidential election provides a nice analytical opportunity since we can overlay GIS boundary files from the 1992 districting plan (drawn primarily by Democrats) that was in effect in November 2000 and sum up the raw precinct-level Bush and Gore votes within those districts, and then do the same for the new districts drawn up in 2002 (by Republicans). In spite of the tied election, Bush had a majority in 13 of Florida’s 23 pre-redistricting Congressional seats, or around 57 percent. But in the new districts drawn up in 2002, Bush had majorities in 17 of 25, or 68 percent. Above, we used the exact same data to simulate a large number of plans with 23 and 25 districts, and the average among our simulations for this legislature size falls somewhere between that created by the Republican and Democratic districting plans.

In fact, the Republicans seem to have achieved a result that is at or slightly beyond the upper limit of the pro-Republican bias achieved by our party-blind algorithm. This dramatic pro-Republican bias was not a fluke associated with the use of 2000 data. If we apply a uniform swing to the district-level aggregates for more recent presidential elections in order to examine a hypothetical

4 Note that some of the differences in estimated bias across elections can be explained by deviations in the overall two-party vote from 50 percent. For example, the estimated bias is unusually large in 1998 in large part because Jeb Bush won by a comfortable margin.
statewide tied vote, we see that Republicans formed majorities in 68 percent of the Congressional districts in 2004, and 64 percent in 2008.

When we aggregate the 2000 presidential results to the level of the 2002 state legislative districts, Bush forms majorities in 61 percent of the lower chamber districts, and 58 percent of the upper chamber districts. These plans exhibit pro-Republican bias that is roughly in line with the simulation results. As with the simulation results, we can get very similar bias estimates if we aggregate votes in statewide rather than presidential races. Moreover, roughly similar estimates of electoral bias can be obtained by using the approach of Gelman, King, and Thomas (2008) to estimate electoral bias directly from district-level results of legislative elections between 1992 and 2008.

3. Is Florida an Outlier?

The most striking result thus far is the rather consistent size of the pro-Republican bias in Florida measured through each of these rather different techniques. Moreover, it appears that much of this bias would have occurred with a simple districting scheme that is blind to race or partisanship. This finding raises at least two broad questions. First, to what extent does an urban concentration of Democrats generate a similar political geography of electoral bias in other states? Second, to what extent does the electoral bias that would be generated by our automated districting algorithm track electoral bias observed in actual districting plans of other states?

In order to provide the necessary cross-state perspective, we have linked November 2000 precinct-level data reported by county governments with corresponding GIS boundary files provided by the US Census Bureau. The use of completely different precinct identifiers in the two data sets makes this a difficult challenge. While improved coordination between the census department and state election officials will soon allow for a more complete data set for more recent elections, thus far we have only been able to complete the matches for November 2000 election results in 19 states. We
have applied exactly the same automated districting algorithm introduced above and produced graphs like those in Figure 5.

The only difference is that because elections in other states were not tied, before performing the simulations we applied a uniform swing to the precinct-level results in order to examine the seat share in a “hypothetical” tied election. We then calculate the average bias estimates across all simulations corresponding to the number of districts in each state’s lower chamber, its upper chamber, and its U.S. Congressional delegation. A useful feature of the 2000 presidential election is the fact that it was very close in a number of states, so that the uniform swing used to achieve a hypothetical tie is not a far stretch of the imagination. However, in consistently lopsided states like Massachusetts or Oklahoma, close statewide elections do not occur frequently.

Figure 1 revealed that the extent to which Democrats are spatially concentrated in urban areas varies considerably across states. We capture this heterogeneity in a simple way by using block-group-level data and regressing, state by state, the Democratic vote share in the 2000 presidential election on logged population density, weighting by the block group’s population. The coefficient from this regression is displayed on the horizontal axis of Figure 8a. The vertical axis displays the average estimated Republican vote share obtained from our simulations of the state’s Congressional and state legislative districts. Observations above .5 indicate that on average, the districting algorithm produced districts that would turn tied elections into Republican legislative majorities.

[FIGURE 8 ABOUT HERE]

Figure 8 suggests that Florida is not an outlier. The correlation between population density and Democratic voting is even higher in several other states, and in most of them, the simulations consistently produced even higher levels of pro-Republican bias than in Florida. Average bias in favor of Republicans is substantial—surpassing five percent of legislative seats—in around half the states for which simulations were possible, and in no state did the simulations produce substantial bias in favor of the Democrats. While more refined analysis of all 50 states is needed once the data
are available, it appears that in some of the largest and most urbanized U.S. states, even without overt racial or partisan gerrymandering, the Democrats are at a disadvantage in translating votes to seats simply because their voters are inefficiently clustered in urban areas. According to the simulations, the Democrats do not suffer from this problem in Western and Southern states, where their voters are more efficiently spread out in space.

Next, we compare the simulated bias we calculate against the actual electoral bias observed in these states. First, these comparisons illustrate the substantial ability of our simulations to predict electoral bias throughout the states. Second, comparing actual bias against simulated bias allows us to identify outlier states, where electoral bias may result not only from voters’ residential geography, but also from intentional or overt gerrymandering.

To perform these comparisons, we aggregate the precinct-level presidential results on which we base our simulations to the level of actual legislative districts. Gerald Wright has aggregated such data for the 2000 presidential election using the districting schemes in place in 2000 and after the 2002 redistricting. Using his data, we apply the uniform swing to the district-level Bush-Gore votes in order to examine hypothetical tied elections, and for both chambers of the state legislature and the state Congressional delegation, we calculate the share of seats that would be won by Republicans. Figure 9 plots this observed electoral bias against the average Republican seat share emerging in our simulations for the corresponding legislature size. Each plot thus contains two data points for each state, with state houses (lower chamber) plotted in blue and state senates (upper chamber) in red.

Three patterns emerge from the Figure 9 plots. First, the positive correlation between the simulation estimates and those based on actual districts provides a useful validation of the simulation results. In general, the large industrialized states where the simulations produced large pro-Republican bias, such as New York and Texas, are the same states where the actual districting plans produced similar bias. Hence, these results suggest the strong ability of our simulations to predict the direction and extent of electoral bias across states. The Figure 9 plots include a 45-degree line, such
that any observation above (below) the line indicates that the observed pro-Republican bias from the existing plan exceeds (falls short of) the bias found in our race- and partisan-blind simulations. Most of the districting plans are clustered fairly close to this 45 degree line, suggesting that in most states, electoral bias would not necessarily disappear in the absence of intentional partisan and racial gerrymandering.

Second, our comparison of observed bias with simulated bias identifies some important outliers. In several Deep South states with VRA coverage, the observed pro-Democratic bias is often much larger than the simulated bias. In these states, efforts to comply with the VRA have produced notoriously non-compact majority-minority districts. The South Carolina and Mississippi state legislatures are especially interesting in this regard. The simulations suggest that, probably because of their dispersed, rural African American populations and lack of urban concentration of Democrats, a simple districting plan based on compactness and contiguity would produce only modest pro-Republican electoral bias. But in the actual legislative districting plans adopted in these two states, the observed Republican electoral bias is quite pronounced, suggesting that the creation of majority-minority districts account for this disparity. Consistent with the conventional wisdom in the racial gerrymandering literature, these plots demonstrate why, as in Florida, Republicans have a strategic incentive to advocate for minority representation in southern states.

Other outliers in Figure 9 are states with notorious partisan gerrymanders. For instance, the Democrats have controlled the districting process in the New York Assembly, while the Republicans have maintained greater influence in the Senate. In Figure 9, though still modestly biased against the Democrats, state House plans are below the 45-degree line, indicating that Democrats have done somewhat better for themselves than would be achieved under party-blind districting. The Senate plans, though they demonstrate a 10-point bias that has helped the Republicans maintain Senate control in a very Democratic state, are no different than what would be predicted from the automated simulations. Given the pronounced concentration of Democrats in New York’s 19th century cities, the
task of “packing” and “cracking” of Democratic appears to occur naturally through residential concentration, even without any intentional Republican gerrymandering.

Third, the graphs also display interesting differences before and after the 2002 redistricting. Georgia produced a famously pro-Democratic plan in its state Senate in 2002, and the graphs demonstrate a sudden drop for Georgia from above to below the 45-degree line. As discussed above, the shift from a Democratic to a Republican redistricting plan in 2002 moved the Florida Congressional districting plan in the opposite direction.

Overall, our cross-state comparisons demonstrate the usefulness of automated districting simulations as a baseline against which to compare historical plans as well as the plans that are proposed and eventually adopted in the current and future rounds of redistricting. Even from this modest sample of states, an important pattern emerges: Substantial observed electoral bias is not necessarily evidence of intentional racial or partisan gerrymandering.

5. Conclusion

This manuscript has demonstrated that in contemporary Florida and other urbanized states, partisan voters are arranged in geographic space in such a way that traditional districting principles of contiguity and compactness will generate substantial electoral bias in favor of the Republican Party. This result is driven largely by the partisan asymmetry in voters’ residential patterns: Democrats live disproportionately in dense, homogeneous neighborhoods in large cities that aggregate into landslide Democratic districts, or they are clustered in minor industrial agglomerations that are small relative to the surrounding moderate Republican periphery. Republicans, on the other hand, live in more sparsely populated suburban and rural neighborhoods that aggregate into districts that are geographically larger and more politically heterogeneous but moderately Republican. This phenomenon appears to explain a large part of the pro-Republican bias observed in recent legislative elections in Florida as well as other urbanized states.
Our findings complement and enhance the voluminous literature on the political effects of gerrymandering. A common theme throughout this literature is that gerrymandering can cause electoral bias by generating an asymmetry in the partisan distribution of voters across districts. Our manuscript draws upon this basic intuition, but we demonstrate that a skewed inter-district distribution of partisanship can arise from voters' residential patterns. Specifically, we show that the "packing" of Democratic voters into urban districts is not merely an elite-driven gerrymandering strategy, but also a mass-driven geographic consequence of urban settlement patterns. Thus our findings dovetail with arguments in the classic political geography literature, which argued that in spite of non-partisan districting procedures, anti-Labor bias emerged in postwar Australia, New Zealand, and the UK because of the concentration of workers in industrial neighborhoods (e.g. Gudgin and Taylor 1979).

Additionally, our findings may confound the traditionally hypothesized relationship between gerrymandering and the partisan control of legislatures. Past scholars have taken sharp positions in favor (e.g. Crespin et al. 2007) and against (Abromowitz, Alexander, and Gunning 2006, Mann 2007, McCarty, Poole, and Rosenthal 2009) the hypothesis that gerrymandering affects polarization in the House of Representatives, and scholars have also examined the impact of gerrymandering on the incumbency advantage (Friedman and Holden 2009). Other studies have analyzed the effect of racial gerrymandering (e.g., Hill 1995, Shotts 2001, 2003) and respect for municipal boundaries (e.g., McDonald 2009b) on electoral bias.

Our findings caution that the relationships between intentional gerrymandering and electoral bias are not necessarily identical across different states. Rather, the nexus between districting strategies and partisan control of legislatures is confounded by the electoral bias that emerges from underlying residential patterns in each state. Since geographic patterns of Democratic voter concentration vary widely across states, each state has a different baseline partisan seat distribution that would emerge under a districting process without overt gerrymandering. Hence, our work
suggests the possibility that each state's unique voter geography may either open up or restrict opportunities for redistricters wishing to implement politically motivated gerrymandering strategies. Simulation results like those presented in this paper might provide a useful baseline for future empirical studies.

Our simulation results also offer insight into the likely effect of various redistricting reforms, such as Amendments 5 and 6 in Florida, that mandate the seemingly objective districting criteria of compactness, contiguity, and respect for municipal boundaries. The simulation method used in this paper mimics the type of districting process mandated by such reforms. Our results suggest that in Florida, New York, Pennsylvania, and other heavily urbanized states, such reforms are likely to lock in a powerful source of pro-Republican electoral bias that emanates from the distinct voter geography of these states.

Although presidential and statewide elections have been quite close over the last decade, the Republicans have consistently controlled between 60 and 70 percent of the seats in Florida's state legislature and Congressional delegation. Beyond the electoral bias in the transformation of votes to seats that we illustrate in this paper, Ansolabehere, Leblanc, and Snyder (2005) describe another, more subtle impact of the asymmetric distribution of partisans across districts. It is conceivable that because of the extent to which liberals are packed into urban districts, the Democratic platform, or at least its perception by Florida votes, is driven by its legislative incumbents—a small group of leftists from Miami-Dade and Broward counties who never face Republican challengers—which in turn makes it difficult for the party to compete in the crucial moderate districts. This hypothesis may help to explain why the Democrats consistently receive higher vote shares in presidential than in state races.

It is striking that political geography can turn a party like the Florida Democrats, with a persistent edge in statewide registration and presidential voting, into something approaching a permanent minority in legislative races. One might imagine that a future Supreme Court would
entertain the notion that this situation reaches the rather high bar for justiciability of partisan
gerrymandering laid out in *Davis v. Bandemer* (1986), where a gerrymander must be shown to have
essentially locked a party out of power in a way that frustrates “the will of the majority.” The recent
opinions of the pivotal justices, however, suggest a claimant would apparently need to demonstrate
that an “egregious” gerrymander is intentional. But this paper shows rather clearly that such bias can
emerge without any malicious intent.
References


Figure 1: Population Density and Republican Presidential Vote Share, Census Block Groups
Figure 2: 2000 Bush Vote Share in Florida
Figure 3: The spatial arrangement of partisanship in Florida
Figure 4: 2000 Bush Vote Share:

Colors correspond to Bush vote share, heights correspond to Local Index of Spatial Association
Figure 5: Results of Districting Simulations Using 2000 Bush-Gore Vote Counts

Average Republican Seat Share in Simulated Districting Plans

![Graph showing the average Republican seat share in simulated districting plans. The x-axis represents the simulated legislative size (number of districts), ranging from 1 to 100,000. The y-axis represents the average Republican seat share, ranging from 45% to 70%. The graph plots the results of districting simulations, with a peak around 100 districts, showing a decrease in Republican seat share as the number of districts increases.]
Figure 6: The Partisanship of Districts Created by Random Simulations

Histogram of District-Level Bush Vote Share
(1,000 Independent 10-District Simulations)
Figure 7: The Partisanship of Precincts’ Assigned Districts

Rural Precincts:
(Under 0.3 Voters per Acre)

Suburban Precincts:
(0.3 to 1.5 Voters per Acre)

Urban Precincts:
(Over 1.5 Voters per Acre)
Figure 8: Simulated Electoral Bias and the Urban Concentration of Democrats

Simulated Electoral Bias and the Urban Concentration of Democrats

Note: The solid line represents the least-squares regression fit. The horizontal axis is measured as the estimated coefficient of population density when county-level Gore (November 2000) vote share is regressed onto county-level population density within each state.
Figure 9: Electoral Bias in Simulated vs. Observed Districting Plans

In both plots, the horizontal axis plots estimates of the share of seats in the legislature that would have Republican majorities from districting simulations under the hypothetical scenario of a tied statewide 2000 presidential vote. Also using 2000 presidential results, the vertical axis plots the percent of seats that would be won by Republicans after applying the uniform swing to votes aggregated to the level of actual districting plans. Each measure is displayed separately for the upper and lower chambers of each state’s legislature.
Appendix A: Districting Simulations using Alternative Election Results

2000 US Senate Race: Nelson (D) - McCollum (R) Votes

1998 Gubernatorial Race: MacKay (D) - Bush (R) Votes