Supplemental Information for *Partisan Dislocation: A Precinct-Level Measure of Representation and Gerrymandering*

Daryl R. DeFord, Nicholas Eubank, Jonathan Rodden

A Sampling Variability

As noted in Section 3, our estimates of voter dislocation are subject to two forming of sampling variability: downsampling the number of voters, and then placement of these voters within each precinct.

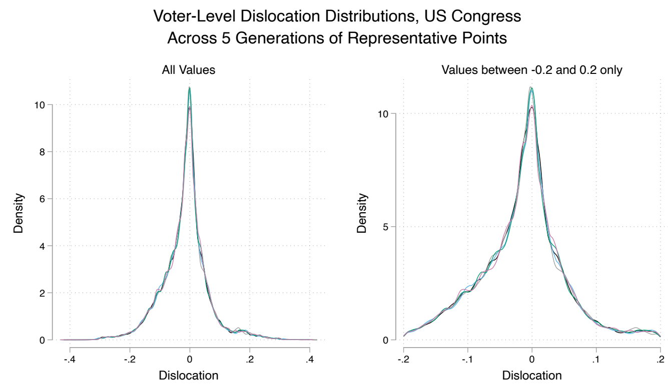
The first source of variance comes from our need to downsample the universe of all US voters for computational tractability. In particular, we create a set of “representative voters” in each precinct for each party by taking a binomial draw from the total number of actual voters for each party in each precinct. The binomial probability varies by state-chamber, but is equal to $prob_k = \frac{number of districts}{number of voters in state} * k$, where k=1,000 for state legislative districts and 5,000 for US Congressional districts. This probability generates k voters per district in expectation. This downsampling makes it computational feasible to calculate the partisan composition each representative voter’s k nearest neighbors. A larger k is used for US Congressional districts as they are much larger with respect to individual precincts, resulting in lower binomial draw probabilities for each precinct, thus increasing sampling variance.

The second source of variance comes from distributing points uniformly within each precinct. Thankfully, US precincts are generally quite geographically compact, limiting the amount of variation introduced by this process.

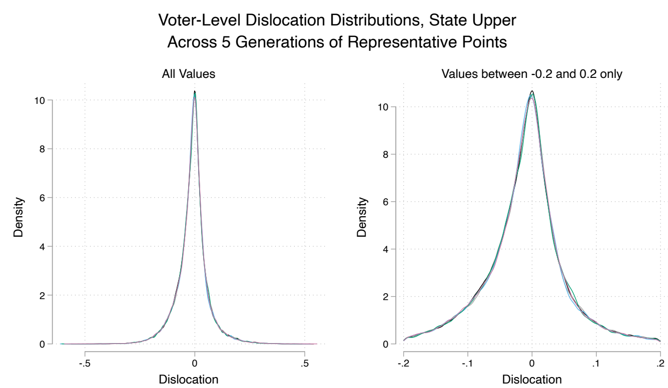
To evaluate the impact of these sources of variability, Figure 13 below plots the distribution of (representative) precinct-level dislocation scores across five rounds of representative-voter point generation. As the Figures show, variation across each round is extremely small, especially within respect to cross-voter simulation: between-round standard deviations constitute only 0.101 %, 0.103 %, and 0.104 % of total variation for these five rounds for state lower, state upper, and US House chambers respectively.

Figure 14 presents analogous diagnostic distribution at the level of legislative districts (plotting the distribution district-level AAPD scores). Again, between-round standard deviations constitute only 1.10 %, 1.40 %, and 2.98 % of total variation for these five rounds for state lower, state upper, and US representative chambers respectively.

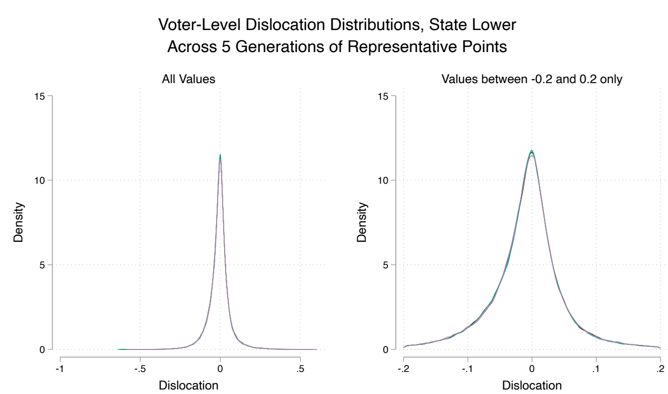
Figure 13



Within Simulation Std. Dev.: 0.0776, Between Simulation Std. Dev.: 0.0001
Between As Pct of Total Std. Dev.: 0.104%
Kernel densities plotted from 10% sample; variance decomposition from full sample.



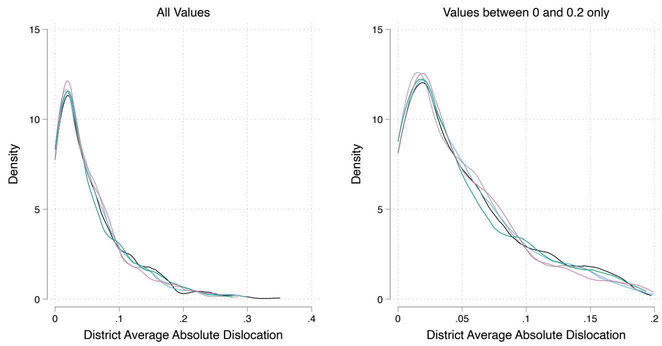
Within Simulation Std. Dev.: 0.0709, Between Simulation Std. Dev.: 0.0001
Between As Pct of Total Std. Dev.: 0.103%
Kernel densities plotted from 10% sample; variance decomposition from full sample.



Within Simulation Std. Dev.: 0.0659, Between Simulation Std. Dev.: 0.0001
Between As Pct of Total Std. Dev.: 0.101%
Kernel densities plotted from 10% sample; variance decomposition from full sample.

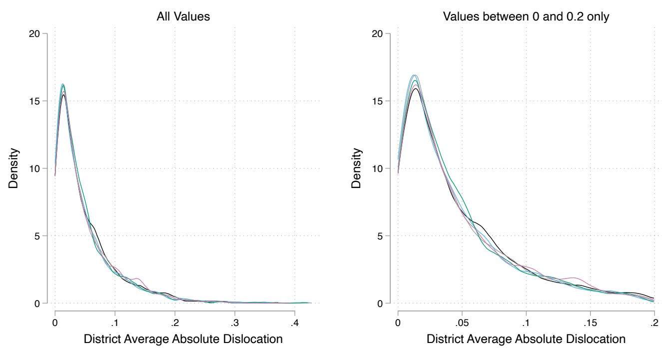
Figure 14

District-Level Absolute Dislocation Distributions, US Congress
Across 5 Generations of Representative Points



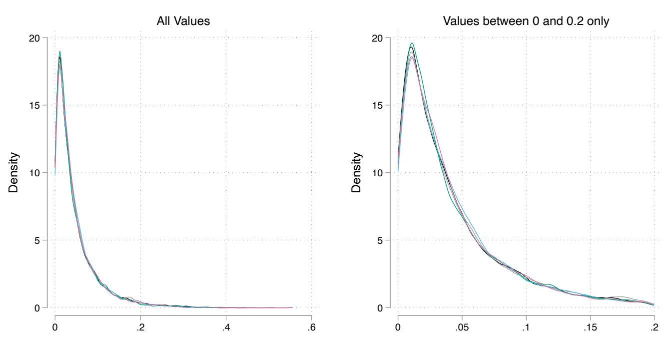
Within Simulation Std. Dev.: 0.0549, Between Simulation Std. Dev.: 0.0016
Between As Pct of Total Std. Dev.: 2.98%

District-Level Absolute Dislocation Distributions, State Upper
Across 5 Generations of Representative Points



Within Simulation Std. Dev.: 0.0517, Between Simulation Std. Dev.: 0.0007
Between As Pct of Total Std. Dev.: 1.40%

District-Level Absolute Dislocation Distributions, State Lower
Across 5 Generations of Representative Points



Within Simulation Std. Dev.: 0.0515, Between Simulation Std. Dev.: 0.0006
Between As Pct of Total Std. Dev.: 1.10%

B Additional Partisan Dislocation Maps

Figure 15: Partisan Dislocation in Texas US House Districts

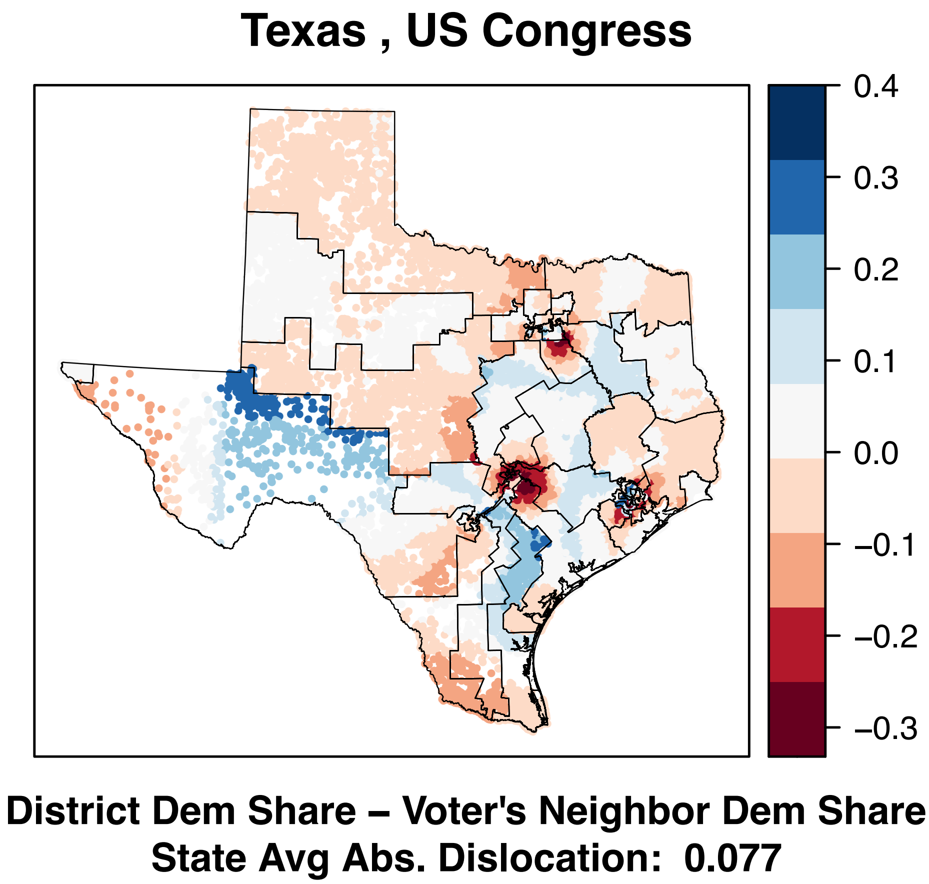


Figure 16: Partisan Dislocation in Louisiana US House Districts

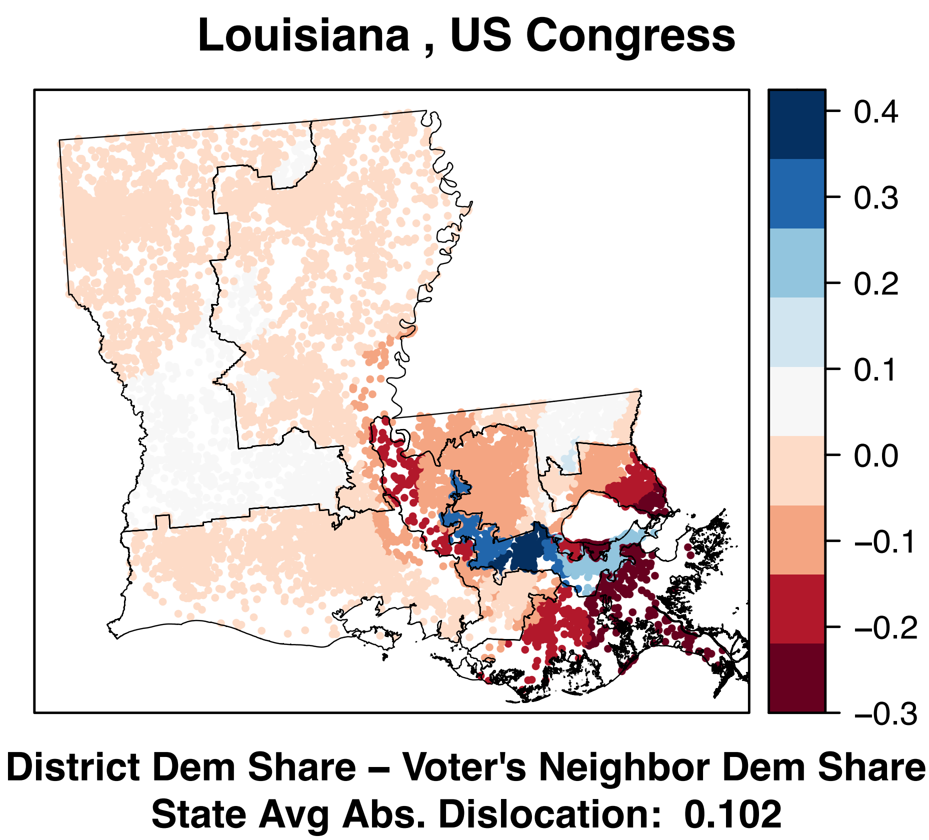


Figure 17: Partisan Dislocation in North Carolina US House Districts

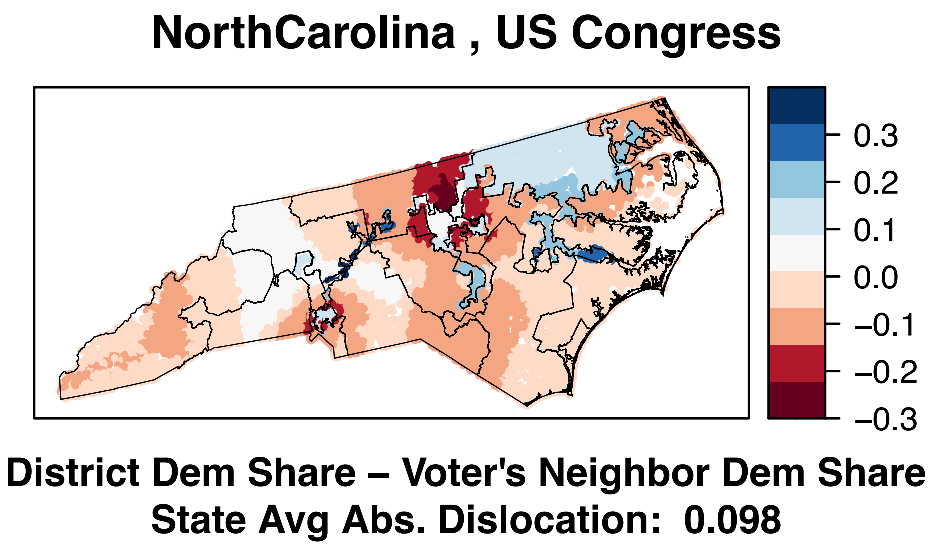
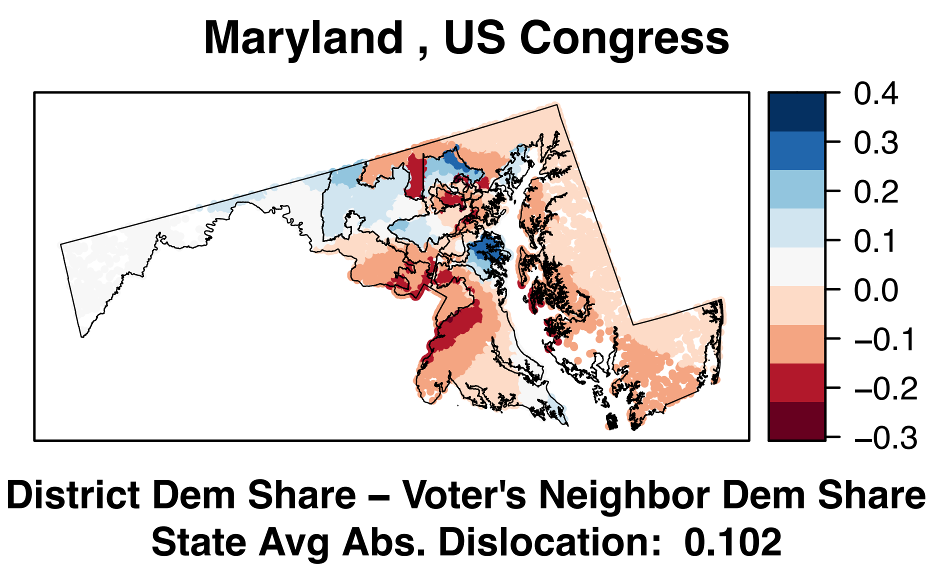


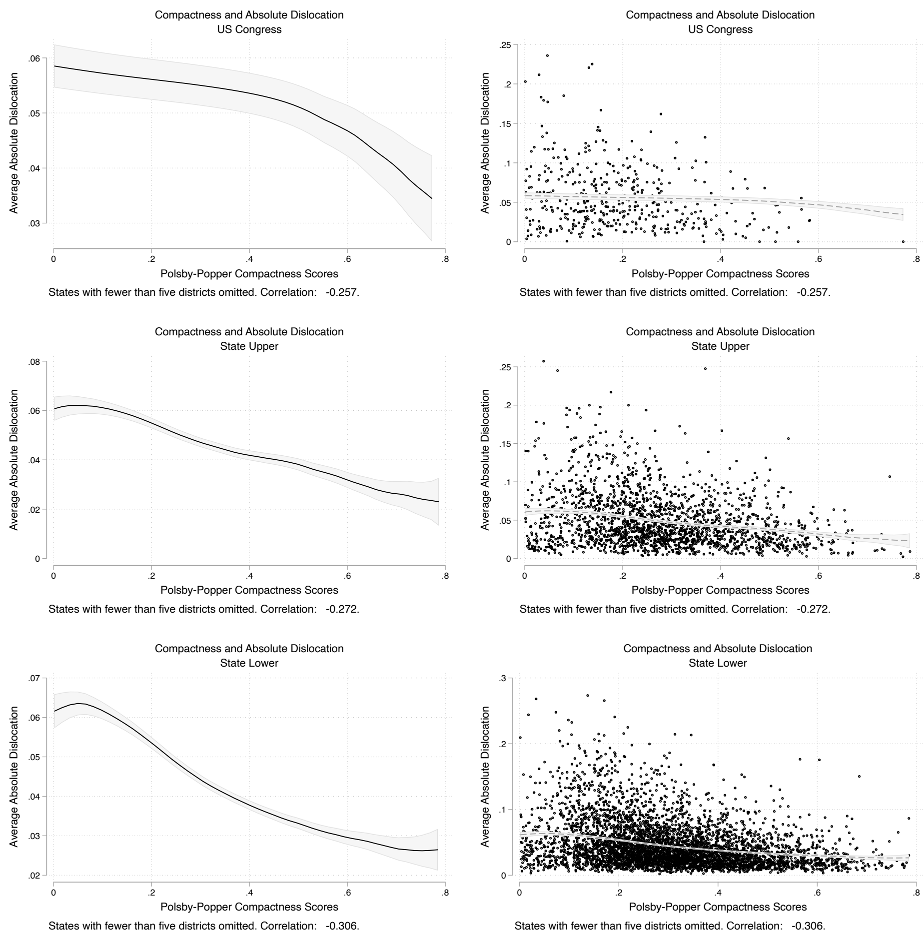
Figure 18: Partisan Dislocation in Maryland US House Districts



C Partisan Dislocation and Compactness

As Partisan Dislocation contrasts the partisan composition of a voter's actual district to what would be the composition of a perfectly compact (circular, modulo boundary reflections) district centered on the voter, once might worry that dislocation simply measures deviations from compactness. As shown in Figure 19 below, while it is the case that dislocation and compactness are related (as we would expect, given the types of deliberately gerrymandered districts dislocation aims to identify) the relationship between the two factors is weak: the correlation is only around ~ -0.275 at all district levels.

Figure 19: District AAPD and District Compactness
With and Without Scatter Overlay



D Simulated Districts

Markov chain based approaches have become a common tool for generating large collections of districting plans. For this analysis, we used the ReCom chain introduced in (DeFord, Duchin and Solomon 2019), which modifies the plan at each step by selecting a pair of adjacent districts, forming a uniform spanning tree on the nodes assigned to those districts, and then selecting a uniformly chosen edge to cut that leaves the remaining parts population balanced to within 1% of ideal. We select the districts to merge proportional to the number of edges on their boundary, in order to promote compactness.

Dual graphs for each state were constructed directly from the precinct shapefiles used in the main analysis. We connected islands and other disconnected regions automatically, finding the nearest precincts in the main body of the state. Florida and California required extra processing, as the shapefiles contained empty polygons that spanned large regions of the state. These outliers were removed and the resulting dual graphs were reconnected.

Initial seeds for the ensembles were constructed using a recursive spanning tree method that generates a single district at a time by drawing a spanning tree for the remaining portion of the dual graph and separating a single edge whose smaller part has population within 1% of ideal. Once the initial population balanced seeds were constructed, an optimization version of the ReCom chain was used to generate starting plans that complied with our chosen VRA bounds.

D.1 Voting Rights Act

In order to model potential impacts of including Voting Rights Act districts in the ensembles, we count the number of districts in each state's 2014 plans whose adult voting age population is at least 40%, 45%, or 50% Black or Hispanic using data from the 2010 census. We then ensure that all simulated district plans have at least the same number of districts that clear these bars. Although the Voting Rights Act does not necessarily support specific numerical percentages by matching the values observed in the enacted plans we are attempting to generate ensembles that represent similar constraints. Results presented in the paper use a 45% threshold, but our results are similar using 40% or 50%.[12] Note that as currently jurisprudence considers the proportion of voting age population in a district that are Black, or the proportion of the voting age population in a district that are Hispanic, but not the proportion of the voting age population that is either Black or Hispanic (so called "coalition" districts), we also use this operationalization.

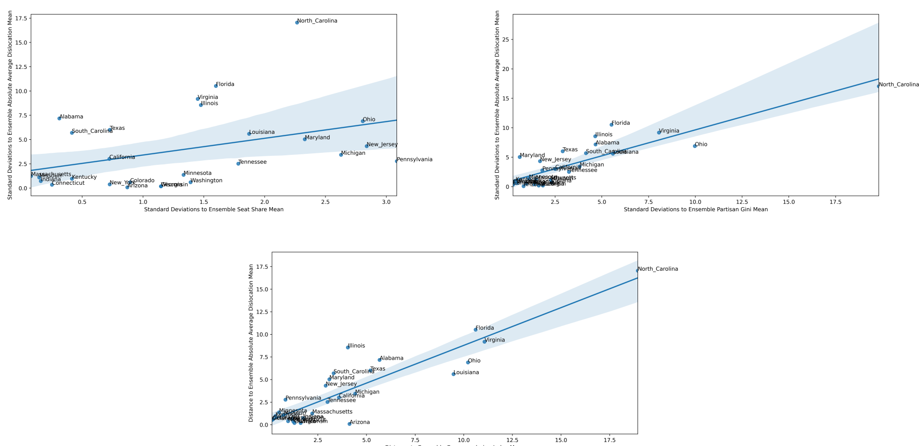
The only exceptions to this procedure were North Carolina and Florida. The map that was in place in North Carolina in 2014 was ruled unconstitutional as a racially

[12]Despite the term "majority-minority," it is rarely the case that the majority of voting age populations in majority-minority districts are actually minority. Exact thresholds vary across states and court cases, however.

packed gerrymander and matching the percentages from that plan would have encoded this packing. Instead, the three ensembles kept two districts over 40%, one district over 40% and one district over 35%, and one district over 40%, respectively. The plots in the main text use the middle ensemble, which is very similar to the approach used in (Herschlag et al. 2018) and related expert testimony in court. In Florida, the state of the precinct data discussed above made it difficult to match the values observed in the enacted plan for Black percentage districts. Thus, districts we used we used bounds of two districts over 40% and one over 35%, one over 45% and one over 35%, and one over 50%, respectively, while computing the Hispanic district bounds as in the other states.

E Simulation-Based Metrics and Simulation-Normalized AAPD

Figure 20



Notes: The above figures plot normalized AAPD scores for states' enacted 2014 US Congressional district plans against simulation-based measures of gerrymandering. Both AAPD and other metrics are normalized by calculating the difference between enacted plan scores and the average score across all ensemble plans (in standard deviations of simulated district plans). Figures include only results for states with five or more districts. As detailed in Appendix D, simulated district plans are subject to compactness and population balance constraints, and all plans have the same number of districts that are more than 45% minority (Black or Hispanic) as enacted plans. Results are similar using either 40% or 50% thresholds for minority share.