

**Expert Report of Wendy K. Tam Cho**

**RE: League of Women Voters v. Wolf et al.**

December 4, 2017

I am a Full Professor with appointments in the Department of Political Science, the Department of Statistics, the Department of Asian American Studies, and the College of Law, a Senior Research Scientist at the National Center for Supercomputing Applications, a Guggenheim Fellow (2016), a faculty member in the Illinois Informatics Institute, and an affiliate of the Cline Center for Democracy, the CyberGIS Center for Advanced Digital and Spatial Studies, the Computational Science and Engineering Program, and the Program on Law, Behavior, and Social Science, all at the University of Illinois at Urbana-Champaign.

I have published scholarly research in the fields of political science, law, operations research, computer science, high performance computing, geography, statistics, economics, and racial and ethnic politics. My research has been supported by multiple research grants from various National Science Foundation (NSF) programs, including political science, statistics, and engineering, as well as multiple computing allocation grants on the Blue Waters Supercomputer, the fastest research supercomputer in the world, with 724,480 processor cores, and peak performance of more than 13 quadrillion calculations per second.

I have been a member of a number of advisory boards, including the Committee of Visitors for the National Science Foundation's Social, Behavior, and Economic Sciences Division; PI4, an NSF funded program to broaden the research background and career prospects of mathematics graduate students; and President Obama's Commission on Election Administration; as well as a member of seven different NSF Review Panels spanning directorates in political science, statistics, big data, and engineering. I was elected to the Executive Council of the American Political Science Association, served as editor of the journal, *Political Analysis*, and am or was a member of the editorial board for nine different scholarly journals. I have served as a reviewer for over 80 different academic journals, agencies, foundations, or presses, spanning a dozen different academic disciplines.

I have had a particular interest in redistricting for over 30 years. Recently, I was awarded a research grant from the National Science Foundation for the development of computational tools for redistricting analysis. I was also recently awarded 6.4 million normalized computing hours

on the Blue Waters Supercomputer to support my computational research on redistricting. I understand and have written about redistricting from a variety of perspectives. My redistricting research has been published in many different academic fields including operations research (Liu, Cho and Wang, 2016; King et al., 2012), high performance computing (Cho and Liu, 2017, 2016a, 2015), engineering (Liu, Cho and Wang, 2015), law (Cho, 2017; Cain et al., 2017; Cho and Yoon, 2001, 2005), and political science (Cho and Liu, 2016b). Some of my redistricting research is aptly described as technical in nature, while other work is pointedly substantive. In 2016, I won the first place prize in Common Cause’s “Gerrymander Standard” writing competition, which was judged by law school deans, law professors, and lawyers. My redistricting research has attracted media attention from popular outlets (e.g., *Vox*, *Salon*, *Chicago Inno*, *Reason*, *The Washington Post*), super-computing outlets (e.g., *Cray Inc.*, *Top 500*, *Communications of the ACM*), and outlets aimed at the science and mathematics communities (e.g., *Quanta Magazine*, *Science Node*, *WIRED*, *Nature*). I regularly teach courses in Constitutional Law and in Election Law. A complete list of my credentials is contained in my curriculum vitae, which is supplied along with this report.

My hourly consulting rate is \$450/hour. My compensation for work expended in connection with this matter is in no way contingent on the opinions I express in this matter. All of my opinions expressed herein are expressed to reasonable degree of professional certainty.

**RE: League of Women Voters v. Wolf et al.**

I have been asked to comment on the expert reports of Wesley Pegden and Jowei Chen.

**Comments on the Pegden Expert Report**

**Description of Pegden’s Report**

Pegden analyzes whether the current Pennsylvania map is an outlier with respect to partisan bias. His finding is that the current map “is indeed a gross outlier with respect to partisan bias, among the set of all possible districtings of Pennsylvania” (emphasis added).

For his analysis, he devises a Markov chain to traverse the state space of possible redistricting plans, runs that chain for up to  $2^{40}$  (approximately 1 trillion or  $10^{12}$ ) steps, and records the maps that satisfy his criteria for a feasible map. While he reports that his algorithm took a trillion steps, it is unclear how many of those steps resulted in a feasible map.

He runs his Markov chain 8 different times, each time beginning with the current map, but modifying the criteria he uses to define a feasible map. In each of these runs, a measure of compactness (either total perimeter or Polsby-Popper) is incorporated and population equality (at either the 1% or 2% level) is enforced. In all but 2 of the runs, he preserves counties. In half of the runs, he holds District 2 constant to preserve it as a Voting Rights district. In all 8 runs, he finds that the current map is “dramatically gerrymandered” because it is an “extreme outlier among the set of possible alternatives.”

### Markov Chains and Markov Chain Monte Carlo

To begin, I will set some groundwork. Pegden has published an article in the *Proceedings of the National Academy of Sciences* (Chikina, Frieze and Pegden, 2017). In that paper, he and his colleagues describe the significance of their work.

Markov chains are simple mathematical objects that can be used to generate random samples from a probability space by taking a random walk on elements of the space. Unfortunately, in applications, it is often unknown how long a chain must be run to generate good samples, and in practice, the time required is often simply too long. This difficulty can preclude the possibility of using Markov chains to make rigorous statistical claims in many cases. We develop a rigorous statistical test for Markov chains which can avoid this problem, and apply it to the problem of detecting bias in Congressional redistricting (p. 2860).

For the purposes of this report, it is enough to understand a Markov chain as described by Pegden and his colleagues. To rephrase and to put in the specific context and language of redistricting, their Markov chain explores the space of possible redistricting maps. The beginning of the chain is anchored at the current Pennsylvania map. The current map or “state” is referred to as  $\sigma_0$ . The algorithm moves from  $\sigma_0$  through a series of  $k$  maps/states to create his Markov chain,  $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_k$ . The way the chain arrives at map  $\sigma_{n+1}$  from the previous map  $\sigma_n$  is by shifting a boundary voting tabulation district (VTD) from one district to its neighboring district.

He runs his chain for  $2^{40}$  steps. If the proposed VTD shift results in a valid map, he places that map into his bag of alternatives. Some shifts do not produce valid maps. He does not report how many steps produce valid maps and how many do not. He needs to report the number of valid maps, not the number of steps. It is inconsequential how many times he tried and failed to find a feasible map. What is important is how many feasible maps he found.

Table 1: Selected Stirling Numbers of the Second Kind

$k$	$n$								
	...	5	6	...	10	...	15	...	55
2	...	15	31	...	511	...	16,383	...	$1.8 \times 10^{16}$
3	...	25	90	...	9,330	...	2,375,101	...	$2.9 \times 10^{25}$
4	...	10	65	...	34,105	...	42,355,950	...	$5.4 \times 10^{31}$
5	...	1	15	...	42,525	...	210,766,920	...	$2.3 \times 10^{36}$
6	-	-	1	...	22,827	...	420,693,273	...	$8.7 \times 10^{39}$
...	-	-	-	...	...	...	...	...	...
55	-	-	-	-	-	-	-	-	1

Markov chains, when they have certain properties (irreducible, aperiodic, and positive recurrent), can be devised as part of a statistical technique, referred to as Markov Chain Monte Carlo (MCMC) to identify the features of an unknown distribution. In the context of redistricting, this is significant because the characteristics of possible redistricting maps in Pennsylvania are unknown. However, if we were to know the partisan metrics of all possible redistricting maps, then we could make statements about whether the partisan metrics of the current map might be an outlier in some defined sense. Devising such an MCMC technique, while theoretically possible, is not practically obtainable because the number of possible redistricting maps is astronomically large so that the amount of computing time required for MCMC to estimate the characteristics of redistricting maps is, for all practical purposes, infinite.

To get a sense for how large this problem is, note that drawing electoral maps amounts to arranging a finite number of indivisible geographic units into a smaller number of larger areas/districts. Since every unit must belong to exactly one district, a map is a partition of the set of all units into a pre-established number of non-empty districts. The redistricting problem is an application of the set-partitioning problem that is known to be *NP*-complete and computationally challenging (Garey and Johnson, 1979). Without any constraints on the process, the total number of possible maps when drawing  $k$  districts using  $n$  units is a Stirling number of the second kind,  $S(n, k)$  (Keane, 1975), defined, combinatorially, as the number of partitions of an  $n$ -element set into  $k$  blocks. The Stirling number of the second kind can be computed recursively as  $S(n, k) = k S(n - 1, k) + S(n - 1, k - 1)$ , which is valid when  $n \geq 1$  and  $1 \leq k \leq n$ . Table 1 shows  $S(n, k)$  for a selection of small values of  $n$  and  $k$ , to provide a sense of magnitude. Even



with a modest number of units, the scale of the unconstrained map-making problem is awesome. If one wanted to divide  $n = 55$  units into  $k = 6$  districts, the number of possibilities is  $8.7 \times 10^{39}$ , a formidable number. There have been fewer than  $10^{18}$  seconds since the beginning of the universe. Of course, as constraints such as contiguity, equal population, and the traditional districting principles are applied, this number declines significantly. We do not have a way to precisely count the number of constrained maps, but the “smaller number” of constrained maps for the state of Pennsylvania is still far far in excess of numbers we think of as large, like, say, a centillion ( $10^{303}$ ). This is why Pegden says that “it is unknown how long a chain must be run to generate good samples, and in practice, the time required is often simply too long.”

The length of time a Markov chain must run to generate good samples is referred to as the mixing time. The Pegden technique does not require a Markov chain to mix. That is, Pegden does not obtain a good sample of the possible redistricting maps. Instead, he devises a reversible Markov chain that begins at the current map, steps away from the current map by randomly shifting one VTD at a time, does this for a large number of steps, observes how many maps encountered on the Markov chain have better metrics than the current map, and then makes a statement that the current map is what he calls an  $\varepsilon$ -outlier that is significant at the  $p = \sqrt{2\varepsilon}$  level.

At issue here is how such a test might be operationalized and applied to the redistricting problem and whether Pegden’s particular implementation and operationalization warrants the conclusions that he draws.

### **The Set of All Possible Redistricting Maps**

Pegden makes this “extreme outlier among the set of possible alternatives” claim despite not examining the set of all possible redistricting maps in the state of Pennsylvania. In a series of claims through his report, his wording on this point is unambiguous, over-reaching, and incorrect. Examples of this language are provided below. The emphasis in each of these claims is mine.

- “I find that the present Congressional districting of Pennsylvania is indeed a gross outlier with respect to partisan bias, among the set of *all possible districtings of Pennsylvania*.” (p. 1)
- “Quantitatively, the [CFP] theorem tells us that more than 99.99% of *the possible Congressional districtings of Pennsylvania* would pass our gerrymandering test, showing in a mathematically rigorous way that the present districting was an extremely careful choice made to maximize partisan advantage.” (p. 2)

- “We will see, in fact, that my analysis shows that the current Congressional districting of Pennsylvania is more unusual than the *vast majority of districtings* with respect to partisan bias.” (p. 2)
- “when I report that Pennsylvania’s 2011 Congressional districting is gerrymandered, I mean not only that there is a partisan advantage for Republicans and that districtings with less partisan bias were available to mapmakers, but indeed that *among the entire set of available districtings of Pennsylvania*, the districting chosen by the mapmakers was an extreme outlier with respect to partisan bias, in a statistically rigorous way.”
- “Our finding is that Pennsylvania is dramatically gerrymandered; its current Congressional districting is an *extreme outlier among the set of possible alternatives*, in a way that it is insensitive to how precisely I define the set of alternatives.” (p. 8)

Pegden is certainly aware that he has not examined all possible redistrictings. In footnote 5, he states that “the number of districtings in the comparison bag can be astronomical; larger than the number of elementary particles in the known universe, for example, so we cannot simply look at them one by one for a comparison.” Indeed, the number not only “can” be astronomically large. The number of possible redistrictings for any state that has more than one district *is* astronomically large.

It is possible to make such a statistical claim with analysis from a method that produces a large representative sample that he could employ in lieu of the set of all possible redistricting maps. He does not, however, create such a representative sample. On this task of drawing an efficient random sample of the set of all possible redistrictings, which is a smaller, but by no means a straightforward or simple task, Pegden states that “there is no general purpose algorithm known which can accomplish this task” (p. 4). In specific reference to the ability of the Markov Chain Monte Carlo (MCMC) method to accomplish this task, he states in his published work (p. 2862), “Indeed, no work has even established that the Markov chains are irreducible... even if valid districting was only required to consist of contiguous districts of roughly equal populations. Additionally, indeed, for very restrictive notions of what constitutes valid districting, irreducibility certainly fails.” Pegden does not attempt to design an MCMC that would accomplish the task of producing a representative sample of all possible redistrictings.

In the absence of either examining the entire set of possible redistrictings or a large representative sample of the set of all possible redistrictings, Pegden is not able to make a credible unqualified claim that a map is an “extreme outlier among the set of possible alternatives.” Note, however, that the Pegden T3 test (emphasis added),

(T3) The overwhelming majority of *all alternative districtings* of the state exhibit (T1), (T2) less than the districting in question (p. 2).

is predicated on a comparison with *all alternative districtings*. He claims to apply this test and draw a conclusion without having examined either all possible redistrictings or a representative sample of all possible redistrictings or by exploring more than a minuscule portion of the set of all possible redistricting maps.

### **The Pegden Algorithm**

In the introduction of his report, he states that he “published a paper which gave a new statistical test to demonstrate that a configuration is unusual from among a set of candidate configurations” (p. 1). Pegden follows this description with the unqualified claim that his “test can be used to demonstrate that a Congressional districting is gerrymandered.” Herein lies our fundamental point of disagreement. This leap cannot be made. While he has a statistical test that provides a *p*-value to indicate how unusual a configuration is from a set of candidate configurations, this is not equivalent to and does not imply that he has developed a general purpose gerrymandering detection tool. The disconnect is between the math and the reality of redistricting.

The title of Pegden’s paper is “Assessing Significance in a Markov Chain without Mixing.” To translate to layman’s language, the clear implication from the title is that even without producing a representative sample, one can determine if a particular configuration is unusual. Pegden did publish a paper that proposes a statistical test to demonstrate that a configuration is unusual from among a set of candidate configurations. The key here, however, is that his “set of candidate configurations” is not all possible redistrictings or a representative set of all possible redistrictings. It is, instead, a set of “local redistrictings.” Because his candidate configurations consists only of “local redistrictings,” he can, at best, only make the claim that the current map is unusual among the set of “local redistrictings.” He can make a claim that the current map is highly unusual for this set and even attach a number to that claim, but that claim and that number apply only to claims about the local redistricting that are “similar” to the current redistricting and not to all possible redistricting maps in a state.<sup>1</sup>

---

<sup>1</sup>It is also worth noting that Pegden does not attempt, in either his report or his published work, to make a rigorous connection between his proposed method and either the case law that surrounds partisan gerrymander claims or the literature in political science. Rather, in Section 3 of Pegden’s report, he describes his own “conservative notion of gerrymandering.” This definition is not rooted in and does not make reference to a legal understanding of gerrymandering. It is, rather, how Pegden would choose to define gerrymandering. He does not connect his “T2,” that “[s]mall random

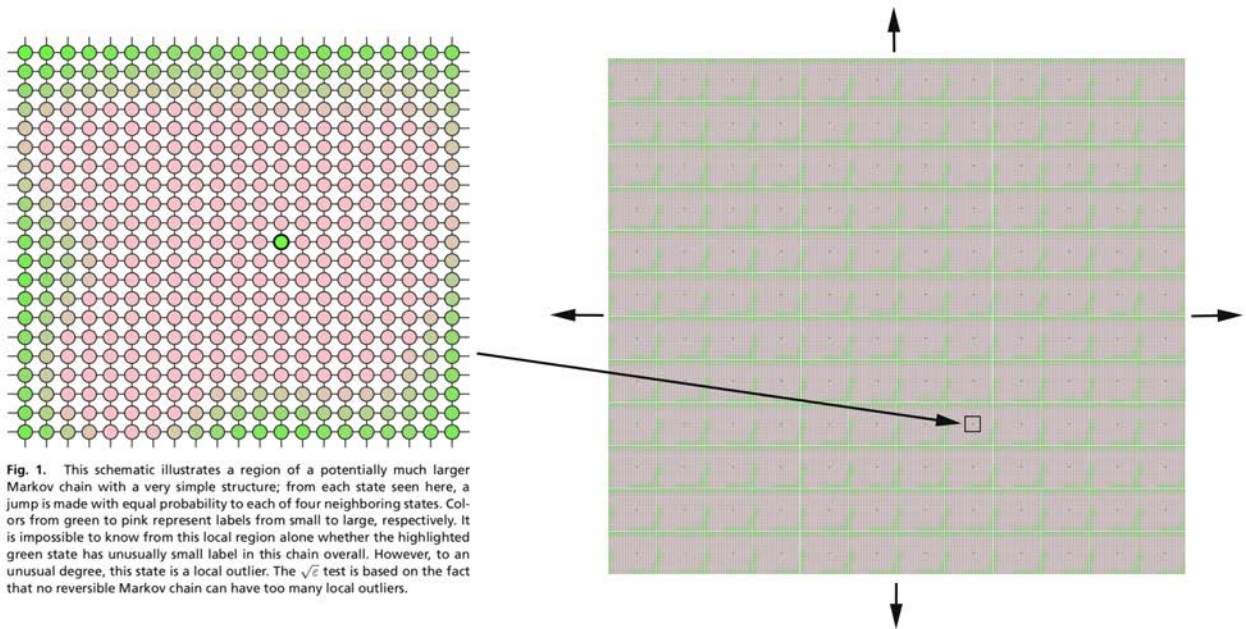


Figure 1: Local Outliers

We can visualize this idea in Figure 1. The picture on the left and its accompanying caption is from Pegden’s article. The green circle that has the bold black outline, which we will call  $\sigma_0$ , is a local outlier because the pink states around it are all bigger on some metric. He states that it is impossible to know from the local region alone whether  $\sigma_0$  is unusually small. However, to an unusual degree,  $\sigma_0$  is a local outlier. Pegden’s  $\sqrt{\varepsilon}$  is based on the fact that no reversible Markov chain can have too many local outliers. While this may be true, the Markov chain also explores only a tiny portion of the entire space of redistricting maps, which is visualized on the right. The arrows on the outside of the figure indicate that the space of maps goes on for quite some time. It can simultaneously be true that a state is a local outlier, but not be an outlier at all in the global space. It is also true that given how astronomically large the state space is for redistricting, a Markov chain of length one trillion explores only a minuscule portion of the entire space of redistricting maps.

Note as well that the space of all possible redistricting maps is highly idiosyncratic. The space is also notoriously difficult to traverse (Liu, Cho and Wang, 2016). The shifting of one VTD indeed results in a different map. However, this new map is essentially identical both to the map from

---

changes to the districting rapidly decrease the partisan bias of the districting, demonstrating that the districting was carefully crafted,” to any Supreme Court ruling on partisan gerrymandering or to any political science research.

which it was derived as well as to a large number of other proximate maps that differ by only one VTD assignment. Moving around this type of space with a one-shift algorithm does not allow one to visit much of the overall space even when this algorithm is run for what sounds like a large number of steps (like 1 trillion).

### **The Bag of Alternatives**

On p. 5 of the Pegden report, he states that “[t]he theorem from [CFP] says that among all possible districtings *in the bag of alternatives...*” (emphasis added). Notice here that “all possible districtings” is modified with “in the bag of alternatives.” This is a significant and critical modification. The bag of alternatives does not have all possible redistrictings. If all possible redistrictings were in the bag of alternatives, which they are not, then we would be able to make claims about the current map with respect to all possible redistrictings. Further, we are all in agreement that the computation needed to create a bag of alternatives with all possible redistrictings is unobtainable within our current computing capacity. Instead, we can only make claims about the current map in comparison to the set of redistricting maps that are represented in the bag of alternatives.

A key to proper (and not overbroad interpretation) of Pegden’s results is to understand what he places in his “bag of alternatives.” To be clear, what is *not* in his bag of alternatives is the set of all possible redistrictings in the state of Pennsylvania or a set that is representative of the set of all possible redistrictings. The comparison is *not* to the set of all possible redistrictings.

On p. 3, Pegden lays out how he determines what goes into his bag of alternatives. The “bag of alternatives” will not magically be composed of all possible redistrictings. How he defines this set of maps and how he identifies this set of maps determines what is in the bag. Here, he says that he has “a model for what would constitute a valid Congressional districting of Pennsylvania,” and that “[s]pecifying constraints such as these determines a ‘bag of districtings’ which are candidate districtings of the state.” His list has 5 elements.

1. The districting consists of 18 contiguous districts.
2. The districting has equipopulous districts.
3. The districting has reasonably shaped (“compact”) districts.
4. The districting does not divide any counties not divided by the current map of Pennsylvania.
5. The districting includes the current District 2, a Majority-Minority district, intact, in case it was drawn to comply with the Voting Rights Act.

This list, as it should be, is derived from legal requirements and the traditional districting principles. However, not all of the traditional districting principles are included. For instance, in Pegden’s candidate map set, cities are not preserved. He does not give a reason for why his candidate maps do not preserve cities. At the same time, he appears to be aware that the preservation of cities affects what type of maps are possible and that the partisan metrics of these maps that preserve cities are different from the partisan metrics of maps that do not preserve cities. On p. 5, he states that

... it is possible for political geography to make a state more favorable to one party or the other. (For example, Democrats, clustered in cities, could conceivably “waste” more votes even for districtings drawn without bias.) This means that in principle, if one only looks at election outcomes under the districting in question without considering how alternative districtings behave, political geography might conceivably give a false impression that a districting was drawn with bias, whereas really it was not.

Importantly, in the current Pennsylvania map, 97.3% of the municipalities are preserved. Such an outcome is not likely to have occurred by chance. It would be fair to say that the current map was drawn with the legal criteria of preserving municipalities in mind. Since keeping cities together (i.e. political geography) “may give a false impression that a districting was drawn with bias, whereas really it was not,” it would not be proper to compare the current map to a set of alternative maps or a “bag of alternatives” where no attempt is made to preserve cities. Given that Pegden is aware of this issue, it is odd that he does not incorporate this traditional districting principle into his algorithm. It is also then not proper for him to then make the broad claim that “it is mathematically impossible for a state’s political geography to inherently produce partisan bias that evaporates quickly when small random changes are made to the state’s districting,” (p .2) when he, himself, singled out preserving cities as “political geography” and then failed to include it in his measure of political geography.

Pegden also does not include incumbent protection in the list of criteria that he considered in creating his “bag of alternative” maps. In the current map, 17 incumbents are not paired with any other incumbent. Pennsylvania had 19 districts in the previous decade and lost one during reapportionment so that they now have only 18 districts. Hence, two incumbents must necessarily be paired. Given that they lost a seat, the reality is that all 18 of the districts are incumbent protection districts. Incumbent protection has been mentioned by the Court as one of the traditional districting principles (See, e.g. *Shaw v. Hunt*, *Easley v. Cromatie*, or *Karcher v. Daggett*) and discussed in the

political science literature as a common consideration in the redistricting process (Mann and Cain, 2005; Bullock, 2010). Given that incumbent protection was a factor in the drawing of the current plan, it must also be one of the factors that determines what goes into the bag of alternatives. It is not. Note that just as preserving cities would affect partisan metrics, considering incumbent protection is also likely to affect the partisan metrics of the bag of alternatives since protecting an incumbent amounts to drawing that incumbent into a district where he is likely to be re-elected.

Pegden states in his published work that “[t]he rigor of the approach thus depends on the availability of a precise definition of what constitutes valid districting; in principle and in practice, the best choice of definition is a legal one” (p. 2862). Pegden does not expend sufficient effort toward understanding what a valid redistricting would be in the state of Pennsylvania. For him to draw any legally valid conclusions from his analysis, his bag of alternatives must include maps that factor in all of the same legal criteria that led to the current map. Pegden’s candidate maps account for some of these factors but omits others. This omission affects his results and subsequent conclusions.

He provides a justification for how he creates his bag of alternatives by saying that the current map is considered “reasonable” and that his choices are based on the metrics of the current map.

It is important to note that, for all of these choices I consider for how to define the bag of districtings, my parameters are chosen so that the 2011 districting meets all of corresponding requirements under consideration. In particular, my goal is not to compare the current districting to other “better” districtings which satisfy stricter requirements on the shapes of the districts, etc. Instead, my test assumes the geometric properties of the current districting are reasonable, and compares the districting to the other possible districtings of Pennsylvania with the same properties (p. 3).

However, he does not require his bag of alternatives to meet all of the same criteria (preserving cities and incumbent protection), and on other criteria, such as population equality, he allows his candidate maps to systematically be worse than the candidate map. This decision biases what appears in the candidate set of comparison maps.

The population deviation of the current map is essentially 0%, within a 1 person deviation. However, rather than require population equality in his candidate maps, Pegden uses either a 1% a 2% population deviation threshold. He justifies his use of 2% with three arguments. The first is that 2% is small in comparison to the error in the Census. While this may be true, this argument

was made and rejected by the Supreme Court (see *Karcher v. Daggett*).<sup>2</sup> Second, Pegden claims that the even if he were to use equal population, they would still exhibit less partisan bias than the current map. This is a conjecture and highly sensitive to the partisan metric he employs. Certainly it is not at all obvious that all partisan metrics decrease by a factor of 2 or more or that all sequences of shifts have this result. His third point is that the threshold does not affect the outcome. Instead of producing candidate maps with a 0% threshold to justify this claim, he states that he “should already see signs of trouble when using a 1% threshold, which is not the case.” This statement is a broad and sweeping claim that is not backed up with empirical evidence. He simply asserts the fact, which is non-obvious. Partisan bias is not a proxy for population deviation. The two do not move in lock step with one another.

It is true that given Pegden’s algorithm, setting the population threshold at 0% would require him to redefine his algorithm since then every step away from the current map would violate population equality. This does not mean that there are not candidate maps with 0% population deviation. It simply means that via his current algorithm, he cannot identify them. His current algorithm would always get stuck at his Step 2 where he randomly selects a “census tract” on the boundary of 2 districts and shifts it if the shift results in a districting that still satisfies the constraints on the bag of districtings.<sup>3</sup> His decision to use a 1% or 2% population deviation makes it easier for him to devise and implement an algorithm, but that is an algorithmic decision, not a decision based on the legal realities of the redistricting problem and the properties of the current Pennsylvania map. There are many ways to devise a Markov Chain. The way Pegden devises it makes it more algorithmically simple to identify maps but precludes the ability to identify 0% population deviation maps since virtually any VTD shift would violate population equality.

---

<sup>2</sup>In *Karcher v. Daggett* (462 US 725 (1983)), the Court states that Appellants

contend that the Feldman Plan should be regarded per se as the product of a good-faith effort to achieve population equality because the maximum population deviation among districts is smaller than the predictable undercount in available census data. *Kirkpatrick* squarely rejects a nearly identical argument “The whole thrust of the ‘as nearly as practicable’ approach is inconsistent with adoption of fixed numerical standards which excuse population variances without regard to the circumstances of each particular case.” Adopting any standard other than population equality, using the best census data available, would subtly erode the Constitution’s ideal of equal representation. . . . We thus reaffirm that there are no de minimis population variations, which could practicably be avoided, but which nonetheless meet the standard of Art. I, § 2, without justification.

<sup>3</sup>Almost certainly, he means that he is selecting a voter tabulation district (VTD). The geography in the shapefile he provided is the VTD. I am not aware that he uses data from the census tract level.



Pegden could prune his bag of alternatives of those maps that do not achieve population equality. He does not do this. Since he has a bag of maps and presumably knows what their population deviation is, this should be simple. He should also then be able to make a statement about whether 0%, 1%, and 2% maps have a fixed relationship with the level of partisan bias observed. He conjectured on this point, but there is no need to conjecture when the data are at hand. It would not allow him to make a general point about all such maps, but it would at least be more credible than simply making an unsubstantiated claim.

Pegden has made many decisions that affect what appears in his bag of alternatives. The bag he creates is not comparable to the current map since he 1) omits legal factors (preserving cities and incumbent protection) that were used to construct the current map and that affect the partisan metrics and 2) redefines other requirements (population equality) so that they are not comparable and worse than the requirements fulfilled by the current map.

### **Local Redistrictings**

It is important to note that even if all the legal criteria for the creation of the candidate maps were the same as the current map, Pegden's algorithm remains incapable of providing a comparison to the set of all possible redistrictings. The way he constructs his bag of alternatives is to begin with the current map and then to shift a boundary VTD. It is obvious that such a mechanism necessarily results in a new map that is essentially identical to the map before the shift. Even after aggregating a trillion such moves, one has explored only a minuscule portion of the set of all possible redistrictings. In Pegden's published article, he states on p. 2863 that "in Fig 2, we see that several districts still seem to have not left their general position from the initial districting even after  $2^{40}$  steps." At best, his bag of alternatives consists of "local redistrictings," certainly they do not represent an array of independent maps that would be representative of all possible redistrictings.

It would be simple for Pegden to provide a sense for how much the maps in his bag of alternatives differ from the current map. He could, for instance, easily find, for each map, how many VTDs were changed from the current map to create that map. He could then supply a histogram that shows the distribution of the number of VTDs that were changed to create each of the maps in his bag of alternative maps.

The current map is, at best, a “local outlier.” It is clearly not a global outlier or an “extreme outlier among the set of possible alternatives.” The legal significance of a local outlier is unclear. However, there is no need to explore this quandary because since Pegden did not produce the proper bag of alternatives, we cannot even make a claim about whether the current map is a local outlier.

### **Results from the Set of 8 Markov Chains**

Pegden reports the results from 8 different runs of his Markov chain. Run 1 and 2 do not preserve counties. Run 3 and 4 do not preserve the Voting Rights district. None of the results from Runs 1–4 should be considered because they leave out either traditional districting principles that should have been part of the definition for feasible maps or legal requirements for feasible maps.

The set of results from Runs 5–8 represent maps that have population deviations in excess of the current plan and so would not be comparable in that respect since Pegden relaxed the population equality constraint. It is noteworthy, however, that the general pattern is that when the constraints become tighter, his results, while remaining quite significant, are less significant. His results are also sensitive to the chosen metric. For instance, using total perimeter makes the results more significant than using Polsby-Popper, even though both are measures of compactness. These patterns suggest that making the population deviation more constraining would reduce the significance of his results even more. The effect is, of course, unknown without the proper analysis. However, since incumbent protection has a partisan element to it, it seems that accounting for this criterion would absorb some of the noted “partisan bias.” Preserving cities likely would absorb more of this “partisan effect.”

### **Measuring Partisan Bias**

Pegden’s use of terms like “partisan bias” imply a false precision. There is no legally accepted definition or measure of partisan bias. Pegden chooses to measure partisan bias with the mean-median difference. The mean-median difference is simply the difference between the average vote share and the median vote share of either party across the set of districts. He does not discuss the impact of this choice on his analysis, which is non-trivial. If he had used the number of seats with a Republican advantage, his algorithm would not likely have identified much change since it requires many VTD shifts to change the map in a substantive way if the measure is the number

of seats with a Republican advantage. It is clear, however, that given his algorithm, he needs a measure that changes even when the only change to the map is the shifting of a single VTD.

In his published work, Pegden refers to a label function, which in this case would be the partisan bias metric. On p. 2 of his “Supporting Information,” he writes,

When we choose which label function to use, we are making a choice based on what is likely to achieve good significance rather than what is valid statistical reasoning. (subject to the caveat discussed below). To choose a label function that was likely to allow good statistical power, we want to have a function that is

- i) likely very different for a gerrymandered districting compared with a typical districting and
- ii) sensitive enough that small changes in the districting might be detected in the label function

That is, he uses the mean-median difference because it changes for even a small change like the shifting of a single boundary VTD. He states that property ii) “discourages the use of ‘coarse-grained’ label functions, such as the number of seats of 18 that the Democrats would hold with the districting in question, because many swaps would be needed to shift a representative from one party to another.”

Note that the “discouragement” here has mathematical origins. Pegden chooses to use the mean-median difference for a mathematical reason, not because it is especially apt for this redistricting case. It is true that the mean-median difference will change for even small changes to a map, like shifting one VTD, but these changes, while resulting in different mathematical quantities, are not politically consequential or interesting. Collectively, many many small changes may aggregate so that they actually result in a substantively different map. Significant, substantive, and politically consequential changes occur only between maps that are sufficiently different from one another.

### **The Trillion Steps**

The algorithm takes a trillion steps. This sounds like a big number, but when one is exploring the space of redistricting maps, it is not a big number. It is, in fact, relative to the size of the solution space, quite a small number. Further, a trillion steps does not result in a trillion maps. It would be simple for Pegden to state how many maps are produced relative to the number of steps taken. This information would be both interesting and insightful about the algorithm’s behavior.

There is also a substantive point that needs to be made here about whether we care about the maps created via this process. Can we justify from a substantive understanding of the redistricting problem whether these maps should be in our comparison set? If the change is substantively meaningless, why is that map in the comparison set? In my opinion, all one-shift “new maps” should be thrown out of the comparison set or else some justification should be made for including them. This should apply to all maps that are substantively equivalent to the current map. How one defines “substantively equivalent” must be determined, but this is a substantive question that requires domain knowledge in the area of redistricting. Mathematical convenience should not be the guide.

It is also not clear that Pegden’s steps are crafted in a way that would allow him to traverse much of the space or find a large number of feasible maps that should be in the bag of alternatives. For instance, if he shifts a VTD and the result is an infeasible map, what should be the next step? Should he return to the previous map and try a different step or should he start from the infeasible map and attempt to find a feasible map? This obviously has an impact on what is identified by the algorithm. If he moves from the infeasible map, likely there will be a large number of other infeasible maps near that map which means many of his trillion steps will be wasted. However, without wasting steps, there are many maps that he would never identify. If he discards the infeasible map, then he also wastes many of his algorithmic steps on movement without identifying feasible maps. In either case, the number of identified feasible maps is likely to be much smaller than the number of algorithmic steps.

If his criteria for a feasible map had included a 0% population deviation, a trillion one-shift steps would have resulted in very few feasible maps. And, likely, all the ones it would have identified encompass only trivial changes from the current map. If his trillion steps identify almost a trillion maps, then this is an indication that many of his maps are substantively identical (despite being treated as mathematically distinct) and that his criteria for a feasible map is not very constraining.

In any case, it is unclear to me from the report how algorithmic steps are related to the number of feasible maps. Clarification on this point would help illuminate how the algorithm proceeds and also provide insight into what types of maps are in the bag of alternatives and how similar

these maps are to the current map. All of these considerations are important for understanding and interpreting Pegden's results.

### **Summary**

To be useful, mathematical rigor must meet the rigor of the law. Mathematical models must be formulated with a deep and nuanced understanding of the problem to which they are applied. Redistricting is a complex, intricate, large, and idiosyncratic problem. Pegden's formulation of the problem is troublesome for analyzing Pennsylvania's congressional redistricting because it does not adhere closely to the reality and complexities of the redistricting process.

In choosing how to construct his bag of alternatives, Pegden makes consequential decisions (e.g., how population deviation should be defined) for mathematical convenience rather than for rigorous adherence to the reality of redistricting and the case law that governs it. He further omits other legal criteria like the preservation of cities despite being aware of its potential influence in partisan metrics. Incumbent protection is not even mentioned.

Pegden's unqualified claims are overbroad and do not match the analysis that he performed.

## **Comments on the Chen Expert Report**

### **Description of Chen's report**

Chen analyzes Act 131 and concludes that it could not have been the product of something other than the intentional pursuit of partisan advantage. He bases this assessment on a comparison of the current map with 1,000 simulated maps. In his words, "[b]y generating a large number of drawn districting plans that closely follow and optimize on these traditional districting criteria, [he is] able to assess an enacted plan drawn by a state legislature and determine whether partisan goals motivated the legislature to deviate from these traditional districting criteria."

He measures "partisan goals" with two measures. The first is a count of the number of districts in a plan that have a Republican advantage. The second is the Mean-Median difference.

He defines traditional districting principles as "equalizing population, maximizing geographic compactness, and preserving county and municipal boundaries."

He provides 2 sets of 500 simulated maps. The first set "optimizes" on population equality, contiguity, avoiding county splits, avoiding municipality splits, and geographic compactness (operationalized via either the Polsby-Popper measure or the Reock measure). The second set uses these same criteria but adds incumbent protection.

He provides figures that indicate that the current map is far from his set of simulated maps, and so concludes that the current map is an “extreme statistical outlier.”

### **What is the Simulation Algorithm?**

Chen does not describe his algorithm in any detail in his report, but merely describes that he has “developed various computer simulation programming techniques that allow [him] to produce a large number of non-partisan districting plans that adhere to traditional districting criteria using US Census geographies as building blocks.” He claims that “[b]y randomly drawing districting plans with a process designed to optimize on traditional districting criteria, the computer simulation process thus gives us a precise indication of the range of districting plans that plausibly and likely emerge when map-drawers are not motivated primarily by partisan goals.” Given that the algorithmic details determine the output produced, omitting the details is not acceptable. It is not acceptable in academic work and not acceptable if one wants to present the output to compel a legal decision.<sup>4</sup>

Consider, for instance, that a number of different criteria are “optimized.” In operations research, we refer to this as a multi-objective optimization. There is not one way to perform a multi-objective optimization. There are many ways, and they do not all lead to the same output. In a multi-objective optimization, the various objectives are not all optimized with every algorithmic step. The movement of one voter tabulation district (VTD) from one district to another district, for instance, may simultaneously preserve a city but make population deviation worse. There are a large number of such conflicts between the objectives, but Chen does not describe how his algorithm would resolve such conflicts. There is not an obvious way to resolve such a conflict and information about the specific choices made in an algorithm are critical to interpreting the output produced as well as to determining whether the algorithm achieved its stated purpose.

There is no dispute in academia than when one creates an algorithm that produces outcomes upon which we make decisions, that the details of the algorithm are material. While precise code may not need to be disclosed, pseudo code or detailed algorithmic steps are minimal. The thresh-

---

<sup>4</sup>After his report was served in this case, Chen offered to make his code and maps available on a confidential basis to be used only in this case. However, the short amount of time that I would have been allowed to view the code would not have been sufficient for me to explore or vet it properly. Further, indeed, the point is not whether I would have been allowed some short amount of time to view the code, but whether the algorithm has been sufficiently scrutinized by the scientific community to allow others, including the courts, to have confidence in the process and results. Transparency is warranted, not simply to me in a short amount of time for one court case, but to the entire scientific and legal community. It should be subject to peer review and accepted in the scholarly community.

old is that a learned reader has sufficient information to be able to independently evaluate and implement said algorithm. It is not acceptable to present “a black box” that produces output.

Chen does not sufficiently describe or validate his algorithm in his academic work. He has a non-technical publication that describes the basic idea that inspires his algorithm (though he has obviously modified that general framework for his analysis of Pennsylvania, which is far more complex). He has not a single technical publication in a statistics, operations research, or computer science journal that rigorously explores the properties of his algorithm or how the algorithm might scale with problem size. He does not describe or validate his algorithm in his report here.

### **Generating a Random Set of Maps**

It is not simple or straightforward to devise an algorithm that produces a random sample of maps that Chen describes as the output from his algorithm. It is not clear that his algorithm produces a set of maps that is not biased in some systematic way. The number of legal maps that can be drawn for the state of Pennsylvania is astronomically large. By just examining the set of maps that Chen produces, there is no way to tell if his sample is a representative set. To examine the properties of an algorithm like his, it is instructive to use a smaller data set for which we know the answer.

As I have already discussed, I am unsure of the details of Chen’s algorithm for Pennsylvania. However, it is clear that he calls them “randomly drawn.” He also provides some guidance in his published article in the *Quarterly Journal of Political Science*. There, he describes a type of Monte Carlo simulation where a geographic units are merged until the number of desired districts is achieved. Neighboring units are then shifted until a population deviation threshold is achieved. Also, as I have already discussed, it is not straightforward how to modify or scale this algorithm when there are many constraints to consider.

We can bypass some of these uncertainties and gain some insight into the Chen method by examining a very simple example that has only one constraint. Consider the very small redistricting problem of partitioning a data set that consists of 25 precincts (from the state of Florida) into 3 contiguous districts. This data set is freely downloadable from the R “redist” package available at <https://cran.r-project.org/>. It was created by [Fifield et al. \(2017\)](#) for a small scale validation study to explore the properties of their MCMC redistricting algorithm. This data set is small enough that all possible redistricting maps with 3 districts can be fully enumerated. That

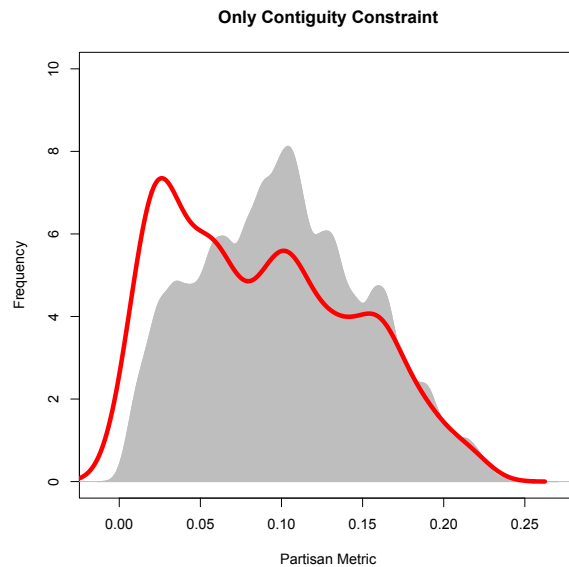


Figure 2: Toy redistricting problem to examine the behavior of “random” map creation algorithms.

is, we know the right answer for this problem. At the same time, the data create a large problem size since the number of ways to partition 25 districts into 3 districts without constraints is  $S(25, 3) = 141, 197, 991, 025$ . If we impose a contiguity constraint, the number of valid partitions reduces by several orders of magnitude to 117,688.

These data allow us to examine the behavior of an algorithm like the one Chen describes that use some random element to construct maps since we know the metrics for every possible map. These types of data sets are essential in designing algorithms for large problems such as redistricting. To be sure, if one cannot design an algorithm that is able to solve this small problem, then it would be ill advised to simply apply the same algorithm to the redistricting problem in Pennsylvania that is astronomically larger with far more complex constraints.

Figure 2 shows the result from an algorithm like Chen’s that uses a randomly element to choose and build districts. The gray area shows the distribution of a partisan metric for all of the possible contiguous maps in the data set.<sup>5</sup> The red line shows the density plot for 1,000 “randomly drawn” contiguous maps. Notice that the “randomly drawn” maps oversample from one part of the distribution while undersampling from other parts, leading to a systematically biased estimate on the partisan metric. In the data set, there are 117,688 possible maps. The size of our “random sample”

<sup>5</sup>The partisan metric is the “Republican dissimilarity index,” which is provided in the data set.



is 1,000. If the maps drawn were truly random, the red line should, in expectation, closely outline the gray area. It does not. Moreover, this systematic bias does not improve if we draw a larger sample. A sample of size 10,000 of these “random maps” exhibits the same bias. Identifying more maps does not produce a good estimate because the identification mechanism, while embodying random elements, does not identify random maps.

Note how simple and small our test data are. There are only 25 precincts in total. Pennsylvania has more than 9,000 VTDs. We partitioned into 3 districts. Pennsylvania has 18 congressional districts. We imposed only one constraint: contiguity. The imposition of any and every other constraint makes the problem significantly more difficult. Drawing a legal and valid map for Pennsylvania includes many constraints. In addition to contiguity, the list includes population equality and compliance with the Voting Rights Act, and may include compactness, respect for political subdivisions such as counties and municipalities, and incumbent protection. An algorithm that is unable to perform well for such a simplified problem is not likely to be useful for an actual redistricting problem.

### **Proper Comparison Set**

As we can see from our simple example, an algorithm that randomly selects units and builds districts does not necessarily result in a random set of maps. However, let us assume that Chen has executed an algorithm that produces a random set of feasible maps that is not systematically biased in some way. What types of simulated maps are needed to make an assessment that the current plan was motivated by partisan goals that deviate from the traditional districting criteria.

Chen provides two sets. The first set of 500 defines the traditional districting principles as population equality, contiguity, avoiding county splits, avoiding municipality splits, and geographic compactness. The second set adds in incumbent protection. Chen states that he has been informed by Petitioners’ counsel that the Pennsylvania Supreme Court has recognized the traditional districting principles of compactness, contiguity, population equality, and the preservation of counties and municipalities in the context of congressional districting. He further states that this list “aligns perfectly with and confirms [his] expert understanding of traditional districting criteria as commonly practiced in congressional districting across the US states.”

I presume Chen supplies two different set of simulation maps because he unambiguously states twice in the report (p. 3 and p. 24) that incumbent protection is not a traditional district-

ing principle. In my opinion, this statement is in error. He does not elaborate, so it is unclear how he would reconcile this position with the Court's statement in *Karcher v. Daggett* when they said "[a]ny number of consistently applied legislative policies might justify some variance, including, for instance, making districts compact, respecting municipal boundaries, preserving the cores of prior districts, and avoiding contests between incumbent Representatives," or when the Court stated in *Shaw v. Reno* that "Neither the *UJO* plurality nor the *Shaw* majority indicates that compactness, contiguity, and respect for political subdivisions are the only districting principles which can be considered 'sound,' and long-standing Supreme Court precedent makes clear that they are not. See . . . *Burns v. Richardson*, 384 U.S. 73, 89 n.16, 86 S.Ct. 1286, 1295 n.16, 16 L.Ed. 2d 376 (1966) (avoiding contests between incumbents)."

While some state constitutions prohibit the protection of incumbents, Pennsylvania is not one of those states. In any case, the proper simulation comparison set for the current map is the set of maps that is created using the same legal criteria that were used to create the current map.

In the current map, given that 17 incumbents are drawn into districts without being paired with another incumbent, it seems fair to infer that incumbent protection was one of the criteria in devising the current map. Further, since Pennsylvania lost one congressional seat in 2010, going from 19 districts to 18 districts, the plan essentially preserved 18 incumbent districts (both Republicans and Democrats). One incumbent is not protected because all 19 cannot be protected when there are no longer 19 districts.

By law, Section 2 of the VRA prohibits discrimination on the basis of race that might dilute the minority vote. In the current plan, there is arguably at least one black district (District 2 is 58.6% black) that needs to be preserved.

A proper simulation set, then, is comprised of maps that are produced respecting the traditional districting criteria, which reasonably include incumbent protection, and certainly include compliance with the Voting Rights Act. Incumbency protection is not unconstitutional and compliance with the VRA is required by law. Further, since both incumbent protection and VRA influenced how the current map was drawn, both should be part how the simulated districts are created.

Of Chen's 1,000 simulated maps, only a handful meet this requirement of a valid and legal set of comparison maps. Rather than drawing maps that satisfy the VRA, Chen simply says on p. 33 that "259 of these 1,000 simulated plans contain one Philadelphia-area district with a 56.8%

or higher African-American VAP.” If only 259 of his 1,000 maps are in compliance with the VRA, then the other 741 simulated maps must be thrown out of the comparison set. It makes no logical sense to compare the current map that satisfies this legal requirement with maps that do not satisfy this legal requirement. The non-compliant maps could not be legally valid maps because they are not drawn with legally mandated requirements. Note that Chen claims on p. 6 that “the computer algorithm generates complete and legally compliant districting plans based purely on traditional districting criteria.” Since these plans do not consider the requirements of the Voting Rights Act, they are not legally compliant districting plans.

If incumbency protection is a legally valid criterion, and the current map also sought to protect incumbents, then only the maps that preserve the VRA district and were from the second simulation set that was produced with the incumbency protection criterion should be in the valid set of comparison maps. Chen does not say how many maps in this second set satisfied the VRA, but we can find this information in his Figure 10. By my assessment, it appears that 54 maps from his total set of 1,000 simulated maps satisfy the traditional districting principles and are in compliance with the Voting Rights Act. Nearly a quarter of these 54 maps had 11 districts with a Republican advantage.

It is important to note here that as the number of constraints on the map creation process increases, the number Republican seats in Chen’s “random maps” also increases. In Simulation Set 1, the number of Republican seats was commonly 8–9. In Simulation Set 2, which was more constrained since incumbency protection was added, the number of Republican seats was commonly 9–11. When we include the requirements of the VRA, 10–11 Republican seats was common. On every criterion, when we compare the set of maps in Simulation Set 1 to those in Simulation Set 2, the maps from Simulation Set 2 are closer to the numbers from Act 131. If Chen were to map his Simulation Set 1 maps along with his Simulation Set 2 maps together in the same plots, we would see that the “cloud” of Simulation Set 2 maps is always closer to the enacted plan.

Consider also that when Chen simulated “incumbent protection,” he merely required that no two incumbents are paired in the same district. This is not how incumbency protection is traditionally understood. Whether one favors incumbent protection or not, incumbent protection is not simply not having another incumbent in the same district, it is the drawing of lines so that the incumbent retains his core constituency. This is clear from the political science literature ([Mann and Cain, 2005](#); [Bullock, 2010](#)). It is not anyone’s common understanding that incumbents are

protected when Republican incumbents are drawn into a newly majority Democratic districts and Democratic incumbents are drawn into a newly majority Republican districts. It is also not the case in the current plan that the incumbents of only one party were protected. If Chen were to incorporate the traditional understanding of incumbent protection into his simulations, surely the number of Republican seats that result from the simulations would rise.

### **What do “Easily Accomplished” and “Reasonably Necessary” Mean?**

Since Chen did not create the proper set of comparison maps, he is not able to make the claim that “the enacted Act 131 plan divided far more counties than was reasonably necessary” or that “a valid plan with only 16 or fewer counties split can *be easily accomplished* without difficulty and without sacrificing other non-partisan districting criteria, such as equal population” (emphasis added). If any legal criterion was not considered in his map creation process but was part of the criteria used for the enacted map, he does not know whether satisfying those legal criteria would necessitate dividing more counties. In his simulation set, he may not have sacrificed population equality, but he did sacrifice minority representation. Minority voting rights are protected by the Voting Rights Act. Redistricting plans cannot legally dilute the minority vote. If his algorithm tried to comply with the VRA, would the maps have split more counties? Given the characteristics of VRA districts in the U.S., additional split counties are likely. Would the districts have been less compact? Given the characteristics of VRA districts in the U.S., the answer is again, likely. Chen implies that because he was able to find these maps “easily,” the enacted map is an unconstitutional gerrymander. At the same time, since Chen does not provide a large number of random legal maps, he has not “easily accomplished” what he claims.

It is also unclear what it means to be “reasonably necessary” or “easily accomplished.” Chen seems to imply some legal standard but it is unclear what that legal standard might be. In any case, it is germane that both of these terms relate back to the details of the algorithm. Some algorithms can “easily accomplish” tasks that other algorithms have great difficulty accomplishing. Does the phrase mean that the algorithm found these maps quickly or that the algorithm found a very large number of these maps or that the algorithm found any at all? If it found a large number of these maps, is there some assurance that the map finding was representative of what is possible and not biased toward oversampling some maps and not others. Or, does it only matter that maps can be found? If my supercomputer or algorithm can identify even better maps, does that make Chen’s

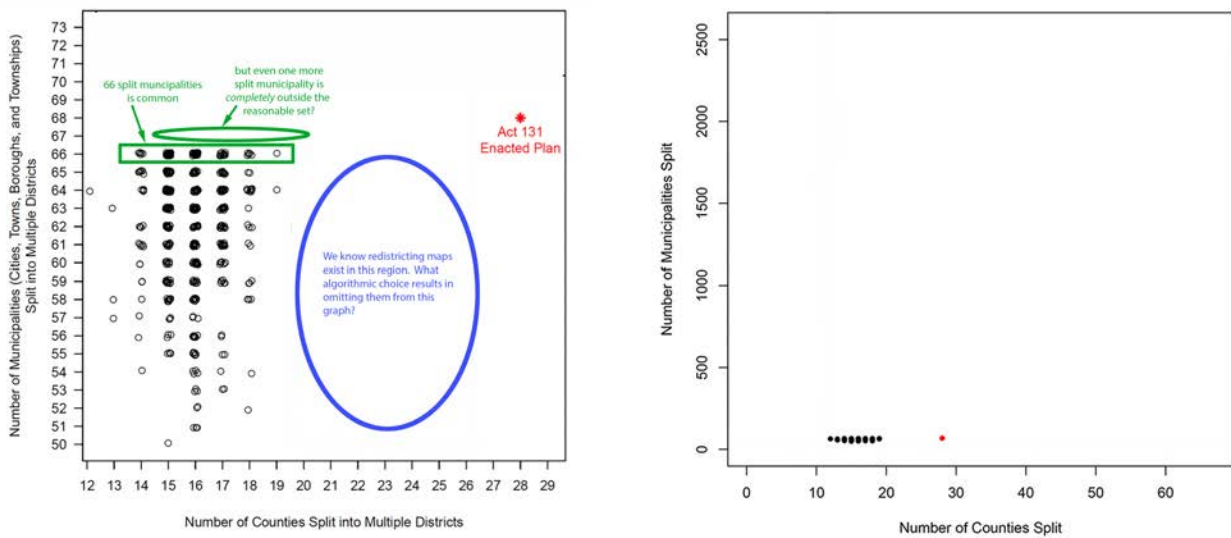


Figure 3: Plot Scales

maps unconstitutional? Is there some cutoff that precludes the output from some computers and some algorithms from this definition of “easily accomplished?” The logic behind his argument is unclear.

### Stylized Interpretations and the Threshold for Unconstitutional

The Supreme Court has not established a clear threshold for when a plan becomes unconstitutional due to “excessive” partisan gerrymandering, or how that threshold might be measured. However, it is clear that Chen wants to say that because Act 131 is unlike the plans in his simulation set, it is an outlier and unconstitutional. The plot on the left in Figure 3 comes from his Figure 6. Between his set of maps and Act 131, there is a gulf where no maps have been identified, implying that there are no maps here. But, of course there are possible redistricting maps here. Surely, if it is easy to create maps with 19 split counties, it is even easier to create maps with 20 split counties. However, the way Chen presents the results, there is an implication that there is his cloud of constitutional maps, then there is a chasm, then there is the unconstitutional Act 131. Given that there are maps everywhere between the set he highlights and Act 131, how precisely does one decide which to present and which to omit? Is there an algorithmic decision here that should be identified?

When one examines his figure along the  $y$ -axis where he plots the number of split municipalities, the subset of possible maps that are in his simulation is even more striking. Could it possibly

be true that it would be *common* and “reasonably necessary” to split 66 municipalities in many different maps, but that there is not a single identifiable instance when splitting even one more municipality would be reasonable? It is hard to believe that, given how many times Chen found it reasonably necessary to split 66 municipalities, there is some hard line between 66 and 67 that makes splitting 67 municipalities unreasonable. Auspiciously for his argument, then, the 68 split municipalities in the current map is completely outside *all* of his reasonable and constitutional maps. If this is a randomly drawn set of maps, he would then say that splitting 68 municipalities is a gross outlier because it is larger than 99.99% of *his* identified maps. This is an untenable claim.

The plot on the right shows the possible values that split counties and split municipalities can take. Chen’s plot was a magnification of this plot. Note that how he chose to portray his results plays with scale. For the range of possible number of municipalities that can be preserved, the simulated plans and the current plan are mathematically proximate—both preserve a high proportion of municipalities. The simulations preserve a few more, but given that there are more than 2,000 municipalities, both preserve a high percentage and split fewer than 3% of the total municipalities.

The number of county splits in the simulated plans and the enacted plan are further apart. However, notice how Chen’s figure shows them at opposite ends of the figure, which gives a misleading picture of the actual range of possibilities. Though there are possible redistricting maps in the vast proportion of the plot on the right, he shows only a small portion of the range of possibilities. By leaving large areas of the plot empty and magnifying one area, Chen implies that the line that divides constitutional from unconstitutional must lie in a particular region. These are games being played with plots and algorithmic decisions, not legal standards.

### **Use of Technical Terms**

Chen repeatedly but unjustifiably invokes technical terms in his report. He calls the enacted plan an “extreme statistical outlier.” This is not a “layman’s term.” It is a technical term with a precise meaning. It implies that some statistical theory and statistical analysis underlie the claim. However, Chen has never rigorously examined his method. His claim necessitates a comparison to a large, random, independent set of maps. His set is not large enough to ensure reasonable precision, does not produce a random and unbiased set of maps, and there is no indication of whether the set is independent or not. The algorithm has not been vetted by the academic com-

munity, published in a scholarly statistics journal, justified in the report, or even described in any detail in the report. This claim simply cannot be made without the evidence to back it up. At minimum, technical terms are not appropriate. Chen could say that he created some maps; the maps have certain properties; and the enacted plan is different from the maps he created. He cannot say, based on what he presents, that the enacted plan is an “extreme statistical outlier.”

Chen also implies that he has implemented an optimization algorithm. He does not explain the nature of the optimization algorithm. This is a term of art in operations research. It is unclear, however, that there is any type of operations research attempted in Chen’s work. Is he actually optimizing via some optimization algorithm or just picking and choosing among maps he finds in some unknown way? Does he implement a simulated annealing algorithm? Ant colony? Particle swarm? Evolutionary algorithm? What is the nature of the optimization algorithm and how is it implemented? Has he benchmarked the algorithm? Is it fast, efficient, or effective? Given how difficult it is to optimize on the space of possible redistricting maps, simply saying “I optimize” falls far short of any academic or scholarly standard. Chen implies rigor through his language. The language must be changed or the rigor must be shown.

## **Summary**

Although Chen’s algorithm has not been the subject of peer review by the academic community, he presents its output as one upon which we should base legal decisions. While there was an offer made after delivery of his report on November 27 to provide the file under a confidentiality agreement, a few business days is not sufficient time for a thorough vetting. The algorithm needs to undergo rigorous scholarly review before it can be accepted as a method that has the potential to have tremendous policy impact.

Even without knowing the precise details of the algorithm, the idea behind his algorithm is problematic. Chen purports to have an algorithm that randomly generates maps. He has never evaluated this claim in any rigorous way. In my assessment of this “random” framework algorithm on a very small toy redistricting data set, I found that the strategy generated a biased set of maps that oversamples some maps while undersampling other maps. The state of Pennsylvania presents a very large application of redistricting that poses a far more complex and intricate computational and statistical problem than the small toy redistricting problem. Since his algorithm is

unable to produce a reliable estimate for a very small redistricting example, it is unreliable for a larger problem.

Beyond the general framework of the algorithm, there are many smaller decisions that have a consequential impact on the output for this case. For instance, when the algorithm has a decision about which VTD goes in which district, or how to shift VTDs, or whatever mechanism Chen uses, and there is a conflict between preserving cities and improving compactness, does it prioritize cities over compactness? Does it prioritize compactness over cities? Does it probabilistically choose which to prioritize? What happens when several of the constraints are improved while several others are worse? Chen uses the language “reasonably necessary,” which is critical to his interpretation and creation of the output but gives no indication of how “necessity” is operationalized.

Even if the idea behind his algorithm was statistically rigorously and the implementation was sound, the criteria he chose to create his comparison set of maps is incorrect. The 946 maps that are not created with all of the legal criteria are not comparable to the current map. The remaining 54 maps were created with a problematic definition of incumbent protection. Given how these maps were created, Chen has not established, and I am not aware of any basis to believe, that these maps are useful in the legal realm for assessing the current Pennsylvania map.

Sincerely,



Wendy K. Tam Cho  
Professor  
Department of Political Science  
Department of Statistics  
Department of Asian American Studies  
College of Law  
Senior Research Scientist  
National Center for Supercomputing Applications  
University of Illinois at Urbana-Champaign



## References

- Bullock, III, Charles S. 2010. *Redistricting: The Most Political Activity in America*. Rowman & Littlefield Publishers, Inc.
- Cain, Bruce E., Wendy K. Tam Cho, Yan Y. Liu and Emily Zhang. 2017. "A Reasonable Bias Method for Redistricting: A New Tool for an Old Problem." *William & Mary Law Review* 59(5).
- Chikina, Maria, Alan Frieze and Wesley Pegden. 2017. "Assessing Significance in a Markov Chain without Mixing." *Proceedings of the National Academy of Sciences* 114(11):2860–2864.
- Cho, Wendy K. Tam. 2017. "Measuring Partisan Fairness: How well does the Efficiency Gap Guard Against Sophisticated as well as Simple-Minded Modes of Partisan Discrimination?" *University of Pennsylvania Law Review Online* 166.
- Cho, Wendy K. Tam and Albert H. Yoon. 2001. "Strange Bedfellows: Politics, Courts, and Statistics: Statistical Expert Testimony in Voting Rights Cases." *Cornell Journal of Law and Public Policy* 10(2):237–264.
- Cho, Wendy K. Tam and Albert H. Yoon. 2005. "Panethnicity Revisited: Asian Indians, Asian American Politics, and the Voting Rights Act." *UCLA Asian Pacific American Law Journal* 10:8–30.
- Cho, Wendy K. Tam and Yan Y. Liu. 2015. A High-Performance Approach for Solution Space Traversal in Combinatorial Optimization. *SC15: The International Conference for High Performance Computing, Networking, Storage and Analysis*.
- Cho, Wendy K. Tam and Yan Y. Liu. 2016a. A Scalable Evolutionary Algorithm with Intensification and Diversification Protocols Designed for Statistical Models. *SC16: The International Conference for High Performance Computing, Networking, Storage and Analysis*.
- Cho, Wendy K. Tam and Yan Y. Liu. 2016b. "Toward a Talismanic Redistricting Tool: A Fully Balanced Computational Method for Identifying Extreme Redistricting Plans." *Election Law Journal* 15(4):351–366.
- Cho, Wendy K. Tam and Yan Y. Liu. 2017. Massively Parallel Evolutionary Computation for Empowering Electoral Reform: Quantifying Gerrymandering via Multi-objective Optimization and Statistical Analysis. *SC17: The International Conference for High Performance Computing, Networking, Storage and Analysis*.
- Fifield, Benjamin, Michael Higgins, Kosuke Imai and Alexander Tarr. 2017. "A New Automated Redistricting Simulator Using Markov Chain Monte Carlo." Working Paper.
- Garey, M.R. and D.S. Johnson. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: W.H. Freeman and Co.
- Keane, M. 1975. "The Size of the Region-Building Problem." *Environment & Planning A* 7:575–577.
- King, Douglas M., Sheldon H. Jacobson, Edward C. Sewell and Wendy K. Tam Cho. 2012. "Geographs: An Efficient Model for Enforcing Contiguity and Hole Constraints in Planar Graph Partitioning." *Operations Research* 60(5):1213–1228.

- Liu, Yan Y., Wendy K. Tam Cho and Shaowen Wang. 2015. A Scalable Computational Approach to Political Redistricting Optimization. In *Proceedings of the 2015 Annual Conference on Extreme Science and Engineering Discovery Environment*. XSEDE15: Scientific Advancements Enabled by Enhanced Cyberinfrastructure St. Louis, MO: pp. 6:1–6:2.
- Liu, Yan Y., Wendy K. Tam Cho and Shaowen Wang. 2016. “PEAR: A Massively Parallel Evolutionary Computational Approach for Political Redistricting Optimization and Analysis.” *Swarm and Evolutionary Computation* 30:78–92.
- Mann, Thomas E. and Bruce E. Cain. 2005. *Party Lines: Competition, Partisanship, and Congressional Redistricting*. Brookings Institutional Press.