ELECTION LAW JOURNAL

Election Law Journal: http://mc.manuscriptcentral.com/electionlaw

Can State Courts Cure Partisan Gerrymandering: Lessons from League of Women Voters v. Commonwealth of Pennsylvania (2018)

Journal:	*Election Law Journal*
Manuscript ID	Draft
Manuscript Type:	Articles
Keyword:	Redistricting, Field: Law, Field: Political Science, Political Parties, Voting Rights (general)
Manuscript Keywords (Search Terms):	Redistricting, Gerrymander, Elections
Abstract:	In League of Women Voters v. Commonwealth of Pennsylvania (2018) the Pennsylvania Supreme Court struck down a congressional plan that embodied one of the most egregious partisan gerrymanders of the 2010 redistricting round. It did so entirely on state law grounds after a three-judge federal court had rejected issuing a preliminary injunction against the plan. Here we examine the ruling and the expert evidence it relied on. We also contrast the criteria used to evaluate partisan gerrymandering by this court with those used by the federal three judge panels that found unconstitutional levels of partisan gerrymandering in plans in Wisconsin and North Carolina, and we compare it to the so far unsuccessful challenge in Maryland. In our concluding discussion we review the key lessons to be learned from this case, and the implications of the case for other jurisdictions, especially those in the twelve states whose state constitutions have provisions similar or identical to the one relied upon by the Pennsylvania Supreme Court.

Can State Courts Cure Partisan Gerrymandering: Lessons from *League of Women Voters v. Commonwealth of Pennsylvania* (2018)

April 22, 2018

WORD COUNT: 17,834 (including footnotes); 15,402 (without footnotes)

KEYWORDS: Gerrymandering; Redistricting; Elections

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

ABSTRACT

In *League of Women Voters v. Commonwealth of Pennsylvania* (2018) the Pennsylvania Supreme Court struck down a congressional plan that embodied one of the most egregious partisan gerrymanders of the 2010 redistricting round. It did so entirely on state law grounds after a three-judge federal court had rejected issuing a preliminary injunction against the plan. Here we examine the ruling and the expert evidence it relied on. We also contrast the criteria used to evaluate partisan gerrymandering by this court with those used by the federal three judge panels that found unconstitutional levels of partisan gerrymandering in plans in Wisconsin and North Carolina, and we compare it to the so far unsuccessful challenge in Maryland. In our concluding discussion we review the key lessons to be learned from this case, and the implications of the case for other jurisdictions, especially those in the twelve states whose state constitutions have provisions similar or identical to the one relied upon by the Pennsylvania Supreme Court.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

I. Introduction

Legendary song writer Woody Guthrie's ballad about "Pretty Boy Floyd" contains the immortal line: "Some will rob you with a six-gun, and some with a fountain pen." If Guthrie were writing today, he might have written 'Some will rob you with a six-gun, and some with computer-aided sophisticated partisan gerrymandering'. Until this year, the consensus among redistricting experts was that, unless we got a 5-4 decision from the Supreme Court that ratified lower court findings about unconstitutional partisan gerrymandering in Wisconsin and North Carolina, partisan gerrymandering would not be remedied in our lifetime. The other route, suing under state rather than federal law, was expected not to work, since the anti-gerrymandering provisions in the Florida Constitution used in 2015 to strike down districts in that state as unconstitutional partisan gerrymanders did not exist in other state constitutions.

In early 2018 everything changed. In *League of Women Voters v. Pennsylvania,* decided January 22, 2018 (henceforth abbreviated *LWV*), the Pennsylvania Supreme Court offered a new and distinctive test for unconstitutional partisan gerrymandering under which it found the state's congressional districts to violate the state's own constitution. In contrast, in 2017, a challenge to Pennsylvania's congressional districts brought in federal court under the Elections Clause of the U.S. Constitution, in a 2-1 vote, the three judge court rejected the viability of bringing a challenge on this legal basis.[1] Because the standard used to strike down the Pennsylvania congressional plan is quite distinct from (and arguably weaker than) any standard likely to be adopted by the U.S. Supreme Court, even if the Supreme Court upholds *Gill v. Whitford*, and enunciates a federal standard for unconstitutional partisan gerrymandering, LWV remains a very important case. Language virtually identical to the provision of the state constitution that the Pennsylvania Supreme Court relied on is found in the constitutions of another dozen states.

In the next section, we discuss the history of the litigation and identify the state constitutional provision that was the basis for the ruling. Then we review the expert witness evidence presented at trial and discuss how the trial court magistrate summarized

[1] The appeal of that case has arguably been mooted by the state court opinion.

and evaluated this evidence. Then we consider the Pennsylvania Supreme Court's instructions about the criteria to be used for plans to be submitted to it intended to remedy the constitutional violation and its decision about which plan to adopt. We also compare and contrast the criteria used to evaluate partisan gerrymandering by this court with those used by the federal three judge panels that found unconstitutional levels of partisan gerrymandering in plans in Wisconsin and North Carolina. And we briefly comment on the potential relevance of this decision to the challenge in Maryland that dealt with only a single congressional district.

In the succeeding section, we turn to comparisons of the adopted plan (henceforth, Court Remedial) with other plans: including three plans rejected by the Court, the 2011 plan, the plan submitted to the Democratic governor by the state's Republican legislative leaders (Joint Plan), the plan submitted by the governor (Gov. Wolf); as well as a plan created by the *DailyKos*, and some plans that we created ourselves (Authors' Plans 1, 2 and 3). Here we highlight an important distinction between *neutral plans* and *fair plans* and consider where on this continuum the Court plan lies -- an issue that has been debated in the blogosphere.[2]

In the concluding discussion, we consider the implications of this case for other jurisdictions in which partisan gerrymandering issues may arise, especially those in the twelve states whose state constitutions have provisions similar or identical to the one relied upon by the Pennsylvania Supreme Court. We argue that there is one clear lesson of *LWV v. Pennsylvania*: if you are going to gerrymander in the future, do it subtly, not blatantly! But this ruling also provides hints that, in the future, even *stealth gerrymandering* may not save line drawers from the vigilance of state courts and the detection skills of redistricting experts. If we are right about that, it will be easier in the future to prove partisan gerrymandering in a state court in Pennsylvania and in other states that adopt the Pennsylvania approach than is likely to be true under federal law, even if the Supreme Court adopts standards similar to those proposed by the three-judge court in *Whitford v. Gill*.

[2] e.g., New York Times Upshot, NYT Upshot

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

II. The Road Traveled

a brief outline of the litigation

As noted above, in *League of Women Voters v. Pennsylvania*, the Pennsylvania Supreme Court found that the congressional plan in Pennsylvania was an unconstitutional partisan gerrymander under Pennsylvania state law. The plan that was struck down was passed in 2011 when both branches of the legislature and the governor were Republican. Prior to the lawsuit, the 2011 plan was widely regarded by experts as among the most egregious partisan gerrymanders of the 2010 round of redistricting (see e.g., Royden and Li, 2017), with some even identifying it as the most egregious (Wang, 2016a, esp. Figure 3).[3]

Before the Pennsylvania Supreme Court exercised its powers of extraordinary relief, it ordered the Commonwealth Court to assign a judge to hold a trial and then issue findings of fact and conclusions of law. In his review, the trial judge found, in essence, that the legislative process through which the plan had been chosen demonstrated the indicia of a partisan gerrymander, that the weird shapes of a number of the districts demonstrated the indicia of a partisan gerrymander, that the dismemberment of municipalities, townships and counties in the state demonstrated the indicia of a partisan gerrymander, that the cracking and packing of Democrat-leaning voters demonstrated the indicia of a partisan gerrymander, and that the frozen 13R-5D results over the course of three elections in a state that is highly politically competitive statewide demonstrated the indicia of a partisan gerrymander. And the trial judge also found that none of these features of the map could be explained as necessitated by the electoral geography and demography of the state, though he suggested the possibility that consideration such as incumbency protection might have mattered. Having reached these findings of fact, the trial court then went on to say that, as a matter of law, there was nothing it could do about the problem under the Pennsylvania Constitution: "Petitioners have not articulated a judicially manageable standard by which this Court can discern whether the 2011 Plan crosses the line between permissible partisan considerations and unconstitutional partisan gerrymandering under the Pennsylvania Constitution" (slip op. No. 261 M.D. 2017, p.

[3] See also McGann et al. (2015, 2016).

126). The trial court magistrate did, however, express the hope that Supreme Court might, in pending cases, identify manageable standards to provide future guidance to other courts.

But, *mirabile dictu,* somewhat like a magician pulling a rabbit from a hat, in reviewing the magistrate's ruling, on January 22, 2018, the Pennsylvania Supreme Court majority found in the "Free and Equal Elections" clause in the Pennsylvania Constitution a bar against partisan gerrymandering, and one which could be used to craft a manageable standard for unconstitutional gerrymandering. In dramatic contrast to the lower court's conclusion of law, the Court held that "the fact that the 2011 Plan cannot, as a statistical matter, be a plan directed at complying with traditional redistricting requirements is sufficient to establish that it violates the Free and Equal Elections Clause [of the Pennsylvania State Constitution] (slip op. at p.128, emphasis added).[4] In support of its readings of this clause, the Court majority provided a long and very carefully crafted historical and legal narrative about the democratic spirit in Pennsylvania.

On February 5, 2018, Justice Alito of the U.S. Supreme Court denied a request by the legislative leadership for a stay of the state court order.[5] The Pennsylvania Supreme Court subsequently gave the Commonwealth of Pennsylvania until February 15, 2018 to submit to a plan that passed muster both with a Republican controlled legislature and the now Democratic governor Tom Wolf, elected in 2014. As expected, the Democrats and Republicans were unable to reach agreement on a plan.

Had such a plan been submitted, the Court would have reviewed its constitutionality. In the absence of a State plan, the Court invited submissions of plans from the parties to the lawsuit, as well as the intervenors. The Republican legislative leaders jointly, the Democratic Governor, Democrats within each branch of the legislature, and the Democratic Lieutenant Governor submitted plans. The Court then rejected all of them, adopting a plan of its own drawn by the Court's consultant,

[4] "Elections shall be free and equal; and no power, civil or military, shall at any time interfere to prevent the free exercise of the right of suffrage." Pa. Const. art. I, § 5.

[5] http://www.philly.com/philly/news/scotus-denies-gop-lawmakers-attempt-to-delay-drawing-new-congressional-map-20180205.html

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Professor Nathan Persily (Stanford Law School). Professor Persily, who has written extensively about election law in general and redistricting in particular, had previously served several times as a court-appointed Special Master in congressional redistricting litigation, drawing plans that were adopted by federal courts. On January 26, 2018, he had been hired to help the Court evaluate submitted plans and, if needed, to provide options to the Court for a congressional plan of its own.

While those, like the present authors, oppose partisan gerrymandering regardless of which party it benefits, will find much to celebrate in the *LWV* opinion, it is worth noting that the antagonism to this opinion by Republican state legislators took on an extreme character. This included not just litigation to block the implementation of the map, but also threats to impeach the judges who voted in the majority,[6] and labeling of the opinion as simply an illegitimate power grab by judges who are Democrats. House Speaker Mike Turzai, R-Allegheny, and Senate President Pro Tempore Joe Scarnati, R-Jefferson, in a joint statement issued on February 20, 2018 condemned the Pennsylvania Supreme Court's map. They maintained that the state justices do not have the authority to draw a map. "Implementation of this map would create a constitutional crisis where the Pennsylvania Supreme Court is usurping the authority of the Legislative and Executive branches," they said. They vowed to fight the map in court. "The League of Women Voters maintains that it filed this suit in order to take partisan politics out of the Congressional redistricting process. This map illustrates that the definition of fair is

[6] https://www.washingtonpost.com/news/wonk/wp/2018/02/22/pennsylvania-republicans-lost-the-redistricting-battle-now-theyre-declaring-war-on-the-courts/?utm_term=.96185402972e

http://www.philly.com/philly/news/politics/will-republicans-impeach-pennsylvania-supreme-court-justices-20180222.html

On March 20, 2018, Rep. Cris Dush, of Jefferson County introduced a measure that would impeach four Democratic members of the court that voted to replace the map. His argument is that the members of the court overstepped their rights and violated the Separation of Powers principle of the state.

simply code for a desire to elect more Democrats."[7] In a February 20, 2018 tweet, President Trump strongly encouraged Pennsylvania Republicans to carry the fight to the Supreme Court. "Hope Republicans in the Great State of Pennsylvania challenge the new "pushed" Congressional Map, all the way to the Supreme Court, if necessary. Your Original was correct! Don't let the Dems take elections away from you so that they can raise taxes & waste money!"

Later in this essay we will discuss the adopted plan and other proposed plans in some detail. But, for space reasons, and because these were seen as the alternatives having the most political support among the submitted plans, we limit ourselves to the plan proposed by the Governor and that proposed by the Republicans in the legislature. Then we consider non-partisan plans drawn according to good government criteria by the liberal blog, the *DailyKos*, and also provide comparisons to two exclusively good government plans that we ourselves draw that did not use partisan data of any kind, and to one plan that we drew almost entirely along good government lines, but with some concern for overall partisan fairness.

Now we turn to the expert witness in the case presented at trial.

examining the expert witness evidence in more detail

We consider five aspects of gerrymandering about which evidence was presented at trial: (1) violation of good government criteria to a degree that exceeded chance, (2) packing and cracking of minority party voting strength, (3) severe disparate partisan impact, to a degree that exceeded chance, after taking into account the electoral geography, (4) durability of the 2011 plan's partisan effects, and (5) intent to discriminate.

(1) violation of good government criteria to a degree that exceeded chance

[7] quoted in Katherine Reinhard, the *Morning Call*, http://www.mcall.com/news/nationworld/pennsylvania/capitol-ideas/mc-nws-trump-tweet-congressional-map-20180220-story.html

Using voting tabulation districts as building blocks, Professor Jowei Chen, testifying on behalf of plaintiffs, developed computer simulation results of plan features under two different sets of assumptions. The first prioritized the traditional criteria in the following order: (1) equal population; (2) contiguity of districts; (3) minimization of counties split between districts; (4) minimization of municipality splits; and (5) compactness, under two different compactness measures ("Reock" and "Popper-Polsby").[8] The second, Set II, added to these criteria the requirement that at least 17 of the incumbents were placed in separate districts. The Set I simulated plans split 11 to 16 counties and split 40-58 municipalities, whereas the 2011 Plan split 28 counties and 68 municipalities. Professor Chen took this as demonstration that these outcomes could not plausibly have emerged from a districting process that prioritizes traditional districting criteria. (Magistrate's Recommended Findings of Fact and Conclusions of Law, slip op., pp. 58-59).

What Professor Chen is doing is classical statistical null hypothesis testing. He posits that the legislature was following traditional districting criteria re minimizing county splits and then checks to see if the county and municipal splits in the 2011 plan are compatible with this null hypothesis. To do so, he generates a distribution of plans drawn with this good government criteria in mind and shows that the observed plan is so far out on the tail of this distribution that we can reject the claim that it was drawn with these criteria in mind. Professor Chen does essentially the same thing vis-à-vis compactness, showing that the 2011 Plan is significantly less compact than every single one of the Set I simulated districting plans, and we can thus safely reject the null hypothesis that it was drawn with compactness considerations in mind.

Professor Chen's demonstration provided a statistical demonstration that the 2011 plan cannot be directed at complying with traditional redistricting requirements, and the Court held that such a statistical finding is sufficient to establish that it violates the "Free

[8] One of these can be taken to be a measure of what Niemi et al. (1990) refer to as an a*real measure of compactness*, i.e., how close a district is to a circle, as judged by comparing the area of the district to the area of a circumscribing circle ; the other can be taken as a measure of what Niemi et al. refer to as a *perimeter measure of compactness*, i.e., how irregular is the perimeter of the district—to be determined by comparing the area of the district to the area of a circle whose circumference is equal to the district perimeter. See further discussion later in the text.

and Equal Elections Clause." Of course, in our view it is highly unlikely that the evidence offered by plaintiffs would have defeated a motion for dismissal had it not also included evidence that much more directly bore on the partisan effects and intent of the plan. In particular, as noted previously, the Court's Final Order refers to the 2011 plan as "severe and durable" partisan gerrymander, and Judge Brobson's opinion in the Commonwealth Court includes in his findings of fact a plethora of evidence that went well beyond its failure to satisfy traditional districting criteria to demonstrate that the plan had the indicia of a partisan gerrymander.

(2) packing and cracking of minority party voting strength and other forms of geographic manipulation

"Cracking" and "packing" are two of the fundamental tools of gerrymandering. *Packing* refers to concentration of other party's electoral strength so that it wins some districts by lopsided margins, thus leaving its voting strength inefficiently distributed. *Cracking* refers to dispersal of a party's electoral strength to maximize the number of districts in which it controls a majority. While packing and cracking are basically two sides of the same coin, as we discuss later, in deciding whether any given partition of a party's electoral strengths across districts involves either packing or cracking of its voting strength, it is necessary to look at the long run consequences of the distribution, both at the level of individual districts and overall. There is no single magic number which constitutes packing, and no single magic number which constitutes cracking. Analyses must be context specific.

Nonetheless, there are two ways in which attempts at packing and cracking might be demonstrated. The first and most straightforward involves examination of actual geographic boundaries of individual districts and a review of the districting choices that were made in creating those districts, i.e., a district by district analysis. The second involves use of some proposed statistical measure of overall packing and cracking. In this subsection we limit ourselves to the first of these; we discuss the *efficiency gap* as a statistical measure of packing and cracking in a later subsection.

Another form of geographic manipulation when redistricting takes place is disparate *pairing*, where incumbents of the minority party are more likely to be paired

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

with another incumbent of their own party than is true for incumbents of the majority party, or where there are pairings of incumbents of opposite parties are done in such a way that the incumbent from the minority party is more likely to lose. Yet another form of geographic manipulation is what one of us has previously labeled *kidnapping*, where incumbents of the minority party are more likely than incumbents of the majority party to see the boundaries of voters within their districts substantially changed in a way that reduces their likelihood of reelection, even though they might not be paired with any other incumbents.

The key expert witness for plaintiffs on the plan features identified above was Professor John J. Kennedy.[9] While usually Professor Kennedy discussed indicators of partisan support in the various "pieces" that made up a given district, so as to demonstrate either packing or cracking, much of his analysis, discussing a dozen districts in some detail, dealt with the strange shapes of many of the districts in the 2011 plan and/or with the failure of districts to reflect common communities of interest. Illustrative is his discussion of removal of Democratic voting strength from Pennsylvania's 6th congressional district to make the district more Republican. As summarized by the trial court (Recommended Findings of Fact and Conclusions of Law, slip op., p. 74):

> Dr. Kennedy explained that the 6th Congressional District, which is likened by some as resembling the State of Florida with a more jagged and elongated panhandle, includes communities in southern Chester County, western Montgomery County, Berks County, and Lebanon County. When asked whether there is anything that unites these communities other than all being located within the 6th Congressional District, Dr. Kennedy opined that they are all separate and distinct communities of interest that have been combined into the 6th Congressional District and not maintained as a whole. Dr. Kennedy also explained that the City of Reading, which is the county seat of Berks County, has been carved out of the 6th Congressional District. Dr. Kennedy opined that this changes the partisan makeup and performance of the 6th Congressional District considerably because the City of Reading is a very Democratic city.

[9] Professor Kennedy used the term *kidnapping* to refer to what we have called same party "pairing." Since the latter term is the more common one in the redistricting literature, it is the one we use here, and we use the term kidnapping for another purpose.

Relatedly, Professor Kennedy pointed out that the City of Reading and another Democratic area, the City of Coatesville, were placed in with heavily Republican areas in the 16th CD (based in Amish country), with a "net political effect of diluting Democratic precincts and Democratic performance in Reading and Coatesville" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 75).

Perhaps the most compelling of Professor Kennedy's dissections of district boundaries was his discussion of CD 7. The trial court (Recommended Findings of Fact and Conclusions of Law, slip op., p. 74) summarized this discussion as follows:

> [T]he 7th Congressional District, which is commonly referred to as the "Goofy Kicking Donald Duck" district, has become famous as one of the most gerrymandered districts in the country. Dr. Kennedy described the 7th Congressional District as essentially 2 districts (an eastern district and a western district) that are held together at 2 locations: (l) a tract of land that is roughly the length of 2 football fields and contains a medical facility; and (2) a Creed's Seafood & Steaks in King of Prussia. Dr. Kennedy also indicated that the 7th Congressional District contains 26 split municipalities.

Because Pennsylvania lost a congressional seat in the 2010 reapportionment, one pairing of incumbents was necessary unless there was a retirement. The Republican legislature and governor chose to pair two Democratic incumbents in the 2011 plan.

(3) statistical measures of severe disparate partisan impact

There were three different measures of disparate impact on which there was expert witness testimony: (a) comparison with simulation results, (b) mean-median gap, and (c) the efficiency gap

(a) comparison with simulation results of partisan outcomes under good government plans

In examining the statistical likelihood that partisan outcome disparities as extreme as those observed under the challenged 2011 plan could have occurred by chance alone, each of plaintiff's expert witnesses on this issue (Professor Jowei Chen, a political scientist, and Professor Wesley Pegden, a mathematician) offered distinctive computer simulation methods.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Professor Chen used essentially the same method for assessing the extremism of the partisan effects of the 2011 plan as he had used for assessing the degree to which the 2011 plan could have been said to satisfy good government criteria. He wrote a computer program to create a large number of good government plans and then evaluated the partisan consequences of each those plans, based on a simple dichotomous measure, namely what was the party with the majority of projected votes in each of the districts and how many seats for each party did such a projection yield for the delegation as a whole. In one of his algorithms he paid no attention to incumbency protection; in the second he did. Then he compared the 13R-5D outcome in the actual plan with the distribution of delegation wide partisan outcomes in the simulations.[10]

He found "that the 2011 Plan creates several more congressional voting districts with partisan performance calculations favoring Republicans, which resulted in several more Republican seats than what is generally achievable under a map drawing process respecting nonpartisan, traditional districting criteria" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 62).

Repeating this analysis with a simulation that took incumbency protection into account, showed that plans were similar to the 2011 districting plans in terms of incumbency protection could achieve "this protection at the cost of only a small increase in split counties and a modest decrease in district compactness" from plans drawn without taking incumbency protection into account. (Recommended Findings of Fact and Conclusions of Law, slip op., p. 66). For example, his second simulation "split between 12 to 19 counties, with the vast majority splitting 15, 16, or 17 counties, whereas the 2011 Plan split 28 counties" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 66).

Professor Chen also took into account racial vote dilution concerns by restricting his analyses to the 259 of his 500 simulated plans that contained a congressional district in it with at least a 56.8% African-American voting age population (VAP) -- the African-American VAP share in the single majority minority district in the 2011 plan. Among these plans, "not a single simulated districting plan remotely came close to creating 13

[10] He also used his simulation results to look at the *mean minus median gap*, but we will discuss that aspect of his findings in the section on partisan asymmetry measures below.

congressional voting districts with partisan performance calculations favoring Republicans" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 69).

Professor Pegden asked a different question from Professor Chen. Professor Chen asked, "How does the 2011 plan compare to good government plans in terms of its partisan disparity." Professor Pegden asked: "If you were to generate plans that were as good as the 2011 plan vis-à-vis good government criteria, how likely would such plans be to have the extreme partisan disparity (13 to 5) found in the challenged plan?" Thus, rather than generating plans from "scratch", i.e., from *de novo* line drawing, Professor Pegden started by taking the actual 2011 plan, and had his computer program make incremental changes in its boundaries to create a new plan in a fashion intended to minimize changes in overall plan compactness. By repeating this process many times, with random starting points for the changes, he was able to create a multiplicity of plans whose good government properties were similar to those of the challenged plan with respect to compactness. Moreover, although county and city cuts were not being into account, in fact, the plans that were created were similar in those respects to the challenged plan that served as its base.

Once we have established the simple fact of a discrepancy between partisan vote share and partisan seat share, Professor Pegden proposed two statistical tests for the existence of partisan gerrymandering. First, the overwhelming majority of the alternative plans should have less extreme partisan imbalance than the challenged plan. Second, because he was creating new plans in an explicitly incremental fashion, Professor Pegden could ask whether or not the changes that were made tended to reduce the level of disparate impact.[11] In his view, if small random perturbations in a plan made it less extreme in its partisan consequences that would suggest that the plan had been a carefully

[11] Here we should note that we have reversed the order of presentation of Professor Pegden's test, since we regard the first of these, the simpler one, as the more important. However, while we do believe that incremental shifts that make switches in partisanship in a direction that reduces the extremism of results are indicative of a plan that is a partisan gerrymander, we also believe that some strongly gerrymandered plans can be created that will pass both tests – ones where districts for both parties are safe, but the minority party districts are even safer than those of the majority party (cf. Owen and Grofman, 1988). But, since we have not actually done such tests using Professor Pegden's algorithm, that can only be a conjecture.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

crafted gerrymander. Professor Pegden found that both his tests demonstrated partisan gerrymandering. Not only was the first test met, in that "The 2011 Plan exhibits more partisan bias than roughly 99.999999% of the simulated possible alternative districting plans created by his computer algorithm," but he also found that "small random changes to the 2011 Plan rapidly decrease the partisan bias of the 2011 Plan" (Recommended Findings of Fact and Conclusions of Law, slip op., pp. 82-83).

The algorithmic approaches taken by both Professor Chen and Professor Pegden were challenged by defendant's expert Professor Wendy Tam Cho. Chen's algorithm was criticized as deterministic and not probabilistic, but that assertion was rebutted by Chen. Pegden's algorithm was criticized, *inter alia*, for its failure to take into account municipal slits and incumbency protection. The trial court concluded that "Dr. Cho's testimony not credible with regard to her criticisms of the algorithms used by Dr. Chen and Dr. Pegden, but credible with regard to her observation that Dr. Pegden's algorithm failed to avoid municipal splits and did not account for permissible incumbency protection" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 91). It then went on to say "Dr. Cho's testimony does not lessen the weight given to Dr. Chen's testimony that adherence to (what he considers to be) traditional redistricting criteria does not explain the partisan bias of the 2011 Plan" And "Dr. Cho's testimony does not lessen the weight given to Dr. Pegden's conclusion that the 2011 Plan is an outlier when compared to maps with nearly identical population equality, contiguity, compactness, and number of county splits" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 91). As we see it, the Chen and Pegden analyses, especially the former, offer the most nearly definitive statistical evidence about partisan gerrymandering because they compare the level of disparate impact in the actual plan to other plans in a way that takes into account the actual electoral geography.

From our viewpoint, the most important feature of Professor Cho's criticisms of the work of Professor Chen and the work of Professor Pegden was that these were being attacked as methodologically flawed. Pointing out methodological flaws is a very important part of the process of doing science, and it can be especially important at trials where there is not a process of peer review of methods and analyses presented. It is always useful, however, to remember Wuffle's 4th Law of Political Science Methodology

(Wuffle, 2015), "It is far easier to criticize than to do better." The failure of Professor Chen to do analyses of her own that added the features that she criticized as being missing from the analyses of plaintiff's experts made here testimony less compelling than it otherwise might have been.[12]

In order to simulate partisan outcomes under alternative maps, as was done by both Professor Chen and Professor Pegden, we need to be able to make projections into districts in which no congressional election has been held. To do this it is necessary to have past election data or some other measure of partisan propensities available to use at the same (low) level of geography that was used to create the plans.[13] There are three main social science issues. First, we need to determine whether we wish to use election data or registration data or some combination of the two as the basis for the projection. Second, we need to determine what election data is best to use for projection purposes, e.g., do we use state-wide bottom of the ticket contests, top of the ticket state-wide contests, presidential contests, or contests for elections of the same type as are under challenge in the case, or do we use some composite set of elections. Third, we may need to determine whether or not it adds significantly to predictive power to include other variables such as home ownership, or racial composition of the district, or incumbency factors, or trends. Our view is that there is no correct *a priori* answer to these questions. Rather, experts can offer predictive models and the test of those models is quite simply "Do they predict (past) elections?"[14]

Both Professor Chen and Professor Pegden made use of political outcome projections to evaluate the partisan consequences of their simulated plans. Professor Pegden used the 2010 Pennsylvania U.S. Senate race between Pat Toomey and Joe

[12] One of the present authors has a great deal of sympathy for the position Professor Cho found herself in, namely being brought in as a purely rebuttal witness, probably with less time than needed to do original analyses of her own, because he has been in that uncomfortable situation himself on several occasions.

[13] This is true whether the plans in question are created via computer simulation algorithm or created "by hand" by parties to the case or intervenors, or by interest groups, or by members of the public simply interested in offering an alternative map.

[14] But "Occam's Razor" also suggests that simpler models are, *ceteris paribus*, better, lest we "overfit" historical data in a way that will not generalize.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Sestak, "because it was a statewide race, there was no incumbent in the race, and it was among the most recent data available to mapmakers when drawing the 2011 Plan." (Recommended Findings of Fact and Conclusions of Law, slip op., p. 80). Professor Chen used "the aggregated results of the 2008-2010 statewide elections" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 63).

Testifying for Defendants, Professor Nolan McCarty's principal criticism of the simulation methodologies used was that the estimates of partisan outcomes they produced were inadequate. His analyses of the 2011 plan, based on a set of aggregated earlier elections used to create a partisan vote index (PVI),[15] suggested that "Democrats should have won 8 seats under the 2011 Plan, and that their failure to do so was based upon other outcomes, such as candidate quality, incumbency, spending, national tides, and trends within the electorate" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 93). However, when Professor Chen's rebuttal demonstrated that his predictive model yielded 54 of 54 correct predictions over the elections of 2012, 2014 and 2016, while Professor McCarty made predictions that were contradicted by actual election results and highlighted the importance of factors that were not actually incorporated into his model, the trial court took an Occam's Razor approach to Professor McCarty's testimony. "The Court finds Dr. McCarty's testimony not credible with regard to criticism of Dr. Chen's report, as the methodology employed by Dr. Chen to calculate partisan performance appears to have been a reliable predictor of election outcomes in Pennsylvania since the enactment of the 2011 Plan" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 93).

(b) mean-median gap

In addition to directly calculating expected outcomes, Professor Chen used his simulation studies of districts drawn utilizing good government criteria to provide statistical comparisons of mean minus median values. The mean minus median test is a

[15] The Cook PVI is based on past presidential voting patterns and is used (at the state or congressional level) measure of relative partisan performance compared to the nation as a whole.

well-known simple measure of asymmetry, closely related to *skewness*. Its use has recently been advocated as a test for partisan gerrymandering (McDonald and Best, 2015; *cf.* Wang, 2016b). When we report the two-party vote shares of a given party, if the mean value of all districts is much larger than its median value, this indicates that party vote shares are not symmetrically distributed (for a normal distribution the mean and the median are identical) and, more particularly, the party in this situation is disadvantaged. The larger the gap, the greater the partisan asymmetry, since the value at the median district tells you which party will control the majority of the seats. "Dr. Chen opined that the skew of the mean-median gap in the 2011 Plan created a significant advantage for Republicans by giving them stronger control over the median district" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 64). Chen found that the range of skew in his simulations that did not take incumbency into account was "always much smaller than the 5.9% mean-median gap observed in the 2011 Plan" and he "concluded, based on his analysis of the mean-median that the 2011 Plan … created an extreme partisan outcome that cannot be explained by Pennsylvania's voter geography or by any of the traditional districting criteria" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 65). Moreover, Chen's further analyses, using simulations that were attuned to the requirement of preserving incumbencies, reinforced this conclusion.

(c) efficiency gap

Professor Christopher Warshaw presented testimony at trial about the *efficiency gap*.[16] "The efficiency gap is defined as the difference between the parties' respective 'wasted votes', divided by the total number of votes cast in the election. In calculating the efficiency gap, all of the losing party's votes are wasted if it loses the election. As to the

[16] While his primary testimony was about the *efficiency gap*, Professor Warshaw also presented testimony more directly about partisan asymmetry. In particular, he pointed out that "in 2012, … Republican candidates won only 49% of the statewide vote, but they won 13 of 18 (72%) of Pennsylvania's congressional seats. Of course, a mere disparity between vote share and seat share, standing alone, is not enough to demonstrate partisan gerrymandering, much less partisan gerrymandering that rises to the level of unconstitutionality. But that fact was only a small part of the evidence for partisan gerrymandering produced at trial.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

winning party, the wasted votes are those above the 50% plus 1 vote required to win" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 85). Here the measure of *cracking* is taken to be the measure of wasted votes given above for the losing party, and the measure of *packing* is taken to be measure of wasted votes given above for the winning party. The sign on the *efficiency gap* will be positive when a given party has a higher proportion of seats than its proportion of votes. Professor Warshaw drew comparisons between its values in elections held under the 2011 congressional map versus elections held under previous congressional maps, and comparisons with efficiency gap values from other states. He concluded that *efficiency gap* measures under the 2011 plan were outliers both with respect to congressional elections held within the state under plans adopted in previous decades and in comparison to current efficiency gap values in other states.

In our view, the *efficiency gap* (EG) is the least useful of the three approaches to measuring disparate impact we discuss in this subsection since it is the furthest from the face meaning of that concept. While it is commonly presented as a measure of wasted votes in terms of vote efficiency lost to packing and cracking, as was done by Professor Warshaw in his testimony, this is somewhat misleading. As the inventor of the measure, Eric McGhee (2017: 6) correctly observes

> "The EG may be framed in terms of wasted votes, but it translates mathematically into a particular relationship between vote share and seat share. The EG will be zero as long as a plan's seat margin (the seat share minus 50 percent) is twice its vote margin (the vote share minus 50 percent). In other words, the ideal plan's seats-votes curve should pass through the (50, 50) point and have a responsiveness of two. This commitment to a specific responsiveness is a critical difference from both symmetry and the MMD [mean median difference]. Both of those measures allow any responsiveness so long as half the votes lead to half the seats."

This description of the efficiency gap by one of its leading proponents is sufficient to demonstrate its limited direct legal irrelevance in the context of partisan gerrymandering. While the idea that symmetry of treatment requires that both parties be treated alike is key to all the measures of disparate impact, the efficiency gap's requirement that the ideal responsive level is two has no obvious legal justification. If partisan vote shares were normally distributed, with a mean of 50% and the standard

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

deviation we would expect for a binomial distribution, for example, the expected responsiveness would be three, a la the famous cube law (see e.g., Kendall and Stuart 1950, see also Theil, 1970). Moreover, the responsiveness level in the plan is tied to the number of highly competitive seats, and thus, if we increase the number of highly competitive districts, this increase may (depending upon the nature of the partisan vote distribution) lead to a higher efficiency gap value. Thus, to the extent that having competitive districts is viewed as desirable then, for any given level of *partisan bias*, having a responsiveness level higher than two, and thus a non-zero *efficiency gap,* is also desirable. Yet reliance on the efficiency gap treats such higher levels of responsiveness as legally suspect.[17]

The potentially inappropriate treatment of competitive districts by the efficiency gap measure was noted by the trial judge, who asserted as a "lingering concern" (Recommended Findings of Fact and Conclusions of Law, slip op., p. 89) that

> "[T]he efficiency gap devalues competitive elections. Specifically, if a "fair" district is one in which the Republican and Democratic candidates have a roughly equal chance of prevailing in the election, a close contest will yield a substantial efficiency gap in favor of the prevailing party. In this regard, the efficiency gap treats a "fair" and competitive district as unfair and possibly unconstitutionally gerrymandered."[18]

[17] A philosophical argument might be constructed as to why two was an "ideal" responsiveness level for state legislative districts, since, if there are "too many" competitive districts, then relatively small shifts in voter preferences may lead to dramatic changes in party seat share, and such violent oscillation in controlling majorities may make it hard to create consistent public policies. But such a debate is well beyond the scope of the present essay.

[18] Its treatment of competitive districts was not the only complaint the trial judge had about the *efficiency gap*. He asserted that the *efficiency gap* "does not take into account some relevant considerations, such as quality of candidates, incumbency advantage, and voter turnout' (Recommended Findings of Fact and Conclusions of Law, slip op., p. 89). The factors pointed by Commonwealth Court Judge Brobson can indeed affect election outcomes, and thus can affect any measure that uses election results as input. But how important are these factors? Here, as with Judge Brobson's similar criticisms of other measures, there is a legal issue about burden shifting. At what point does the evidence for extreme partisan gerrymandering become sufficiently clear and compelling that it

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Having identified these caveats about the efficiency gap, we also want to identify its positive features. First and perhaps most importantly, like the mean-median gap, it is so simple to calculate for an election that has already taken place, where issues of projection of previous election results do not come into play, that law clerks can calculate it for themselves without recourse to expert witness testimony.[19] Second, because of this fact, we can imagine a rather different use of the *efficiency gap*. Here we would use the efficiency gap as a "miner's canary."[20] For example, courts concerned with wasting their time and the taxpayers' money on frivolous challenges to partisan gerrymandering could simply say that any value of efficiency gap below a certain level did not rise to the level of a constitutional violation.

While proportionality is, of course, not required by the Constitution, the Supreme Court in *Gaffney v. Cummings* 412 U.S. 735 (1973) suggested that it was not improper for jurisdictions to seek partisan fairness by drawing plans that represented parties in reasonable proportion to their support in the voting electorate. Just as we might use the efficiency gap as a miner's canary, values of responsiveness below two suggest that the disparate burden on the disadvantaged party/point of view is not that large, with a value of one showing perfect proportionality. Thus, courts may wish to simply refuse to hear challenges to partisan gerrymandering where the seat share above fifty percent of the seat majority party is less than twice the vote share above fifty percent of that party as long as the party controlling a majority of the seats also has more votes than the opposing party.

---

now becomes the responsibility of the defendant jurisdiction to present rebuttal to show that legitimate factors not taken into account in the analyses presented by plaintiffs can explain the patterns in the data that provided the indicia of gerrymandering? No single analysis can do everything, and it is often easy to come up with a laundry list of hypothetical complications or omitted variables, without actually presenting evidence that the potential confounds actually mattered on terms of conclusions about whether or not there had been egregious partisan gerrymandering.

[19] The only real issue is the treatment of uncontested districts and whether to code this as 0% and 100% or to use some method of imputation based on previous election results or results in other simultaneous elections.

[20] We are indebted to Lani Guinier (2002, 2009) for calling this useful metaphor to our attention.

However, while low values of the efficiency gap or of responsiveness might be used to demonstrate that the level of barrier to "fair and effective" representation raised by the line drawing would not be at the level of a potential constitutional violation but fall instead within the realm of "ordinary" politics, this does not at all mean that high values of responsiveness or high values of the efficiency gap prove unconstitutional partisan disparate impact. Partisan impact would remain to be litigated. What we are proposing is a one-sided test, since we are using responsiveness or the closely related efficiency gap measure to tell us about the magnitude of difference between seat share and vote share, not whether that difference creates a disparate impact that is attributable to partisan bias. The latter is presumably the question about which courts are most concerned.

In contrast to the *efficiency gap*, the partisan asymmetry approach that traces its roots to scholars such as Robert Dahl (1957) and Edward Tufte (1973), and which is perhaps now best represented in the work of scholars such as Gary King and Andrew Gelman (Gelman and King 1994), distinguishes *responsiveness* (sometimes called *swing ratio*) from *partisan bias*, and regards only the latter as the proper measure of partisan disparity.[21] This literature has been either neutral with respect to the normative desirability of responsiveness, or has treated higher levels of responsiveness as desirable. Unfortunately, there was not expert witness evidence presented at this trial using approaches to partisan gerrymandering of the kind that would have allowed for simultaneous measurement of both responsiveness and bias.[22] Nonetheless, the kinds of evidence provided by Professor Kennedy of very specific geographic instances of packing and cracking, the statistical evidence of the extremity of the 13R-5D outcomes relative to what would be found in computer simulations (either relative to truly good government plans a la Chen, or relative to plans that met good government standards to roughly the same extent as the challenged plan, a la Pegden), and the evidence that the 2011 plan's *mean-median gap* was also a statistical outlier, taken in toto, provide

[21] See review in Grofman and King (2007), see also Grofman (1983). Many other scholars have used variants of this approach. See esp. McGann et al. (2016) for a recent application.

[22] In our later discussion of alternative plans, we do not provide evidence about partisan bias, since it was not one of the indicia introduced at trial or considered by the trial court.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

compelling evidence of partisan gerrymandering. These indicia, in our view, could readily be taken, collectively, as one nail in a multi-part test for partisan gerrymandering that more specifically addressed partisan impact -- as opposed to the standard adopted by the Pennsylvania Supreme Court which looks solely at the degree to which good government criteria are satisfied. This is especially true given the evidence in this plan of its persistent partisan effects (13R-5D), as discussed below.

In this context we want to emphasize a very important point. Just as students in elementary statistics classes are taught that there are different features of a distribution, e.g., mean, variance, skewness, kurtosis, and are also taught that no single number allows us to fully capture the features of a distribution, although some, such as the mean and variance, capture features that are commonly regarded as of the greatest interest, there is no single "magic number" re an analysis of gerrymandering. That is true not just for the efficiency gap, but for the other measures as well. That is why having multiple measures the conclusions from which reinforce one another is so useful, but also why we need to separately examine magnitude of disparate effect and likely durability of that effect.

(4) durability of the 2011 plan's partisan effects

In judging plan durability, it is important to recognizing that, in principle, you can get different outcomes in competitive districts depending upon the nature of electoral tides, i.e., "good" years for Democrats (e.g., 2008) versus "good" years for Republicans (e.g., 2016). The simplest way to deal with this is to look at the actual (or projected) votes in years varying in the level of their partisan support and ask how resistant the outcomes in a given plan would be to perturbations in party support over these past elections. This is the approach we have taken in the discussion in Section III below.

Another more general way to examine likely durability would be to look at the magnitude of all recent past inter-election shifts in two party vote shares at the state level and look at the changes in district outcomes in a given plan that would result were such shifts occurring in each district. Here we would look at the directionality of previous shifts and either the mean or maximum value previously observed. While we do not believe that this type of evidence was presented at trial in this case. the fact that three elections had already been conducted under the plan vitiated the need for direct

examination of inter-election shifts. Quite simply, the fact that the outcomes had been 13R-5D in three consecutive elections, regardless of the partisan tides in those years, demonstrated the durability of the 2011 plan's egregiously disparate partisan effects in the existing political climate in Pennsylvania.

Ceteris paribus, districting effects can be expected to vary substantially over the course of a decade only if there are a substantial number of competitive seats. In the modern era of data rich computer-based redistricting, and highly polarized voting patterns, skillful gerrymandering can use the combination of *packing* and *cracking* gerrymandering techniques to eliminate the possibility of meaningful competition for most of the districts. By largely eliminating competitive districts, it is possible to freeze in place outcomes that will be more or less unchanged over the course of a decade, in a fashion that is largely independent of changes in voter preferences. Thus, absent a wave election, one party's partisan control over most districts, with concomitant control of a legislative chamber or a congressional delegation as a whole, can be maintained. As Keith Gaddie has pointed out, this lack of responsiveness punishes not just supporters of the disadvantaged party, but also punishes swing voters, since they find themselves unable to affect outcomes even when their own partisan preferences have shifted.[23]

Line drawers are now more knowledgeable about the need to avoid drawing *dummymander* (plans drawn by one party that end up benefiting the other party when electoral tides shift) of the sort warned against by Grofman and Brunell (2005), than they were in past decades, as that term has become an increasingly familiar part of the redistricting lexicon. Line drawers also have gained more technical sophistication in mapping historical election data into proposed districts and then checking to make sure that they do not make a dummymandering kind of mistake.

In general, to assess durability, information about the number of potential competitive seats in the plan is very relevant (cf. Brady and Grofman, 1991), and should be viewed in conjunction with (historical) information about the magnitude and distribution of *inter-election shifts*. The common-sense idea is very simple. If we know the historical patterns in changes across elections, then we can see how many districts

[23] Remarks at the Roundtable on Partisan Gerrymandering, Annual Meeting of the American Political Science Association, Chicago, April 6, 2018.

would be affected by changes in the vote share in districts in the challenged plan that are of the same order of magnitude as the inter-election shifts in vote shares we have seen in the jurisdiction in the past. If districts are drawn sufficiently non-competitive, but with one party's voting strength even more concentrated than that of the other party, uncertainty about future outcomes can be kept to a minimum, or eliminated entirely, and the preservation of partisan bias imposed by differential cracking and packing perpetuated. We can, in principle, have large scale disparities in treatment that are likely to be fleeting, and large scales disparities in treatment that can be expected to last an entire decade, but the same is also true for small scale disparities. Magnitude and expected duration can be unrelated to one another, and thus each much be examined separately.

As we will discuss in more detail later in the paper, it is also well known that incumbents of a given party can run ahead of the party's support in statewide contests. For example, if you simply project President Trump's 2016 vote into the 2011 districts you would get six Democratic victories in 2016 as compared to the five Democratic congressional wins. We can explain that difference by the fact that Clinton won pluralities in two seats that Republicans incumbents won, in districts 6 and 7, while Trump won the plurality in district 17, which was won by the Democratic incumbent**.** Because of incumbency advantage effects,[24] we might see a reduction in the seats minus votes gap for the major party in elections after the first under a new plan if that first election results in the election of more incumbents of the dominant party. If we are trying to develop realistic predictions of future elections it might be important to take incumbency effects into account, but there are also good reasons to disregard those in computer simulations, since in simulations we want to look at potential plans as a *tabula rasa*.

[24] Looking at districts over time, incumbents generally run better than non-incumbents of their same party because incumbents, on average, have more developed campaign organizations, sources of funding, and better name recognition, among other advantages.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

(5) intent to discriminate.

There are multiple ways to get at legislative intent. First, there might be smoking gun evidence in the form of public statements by legislators. Remarkably, such statements were made in the 2010 round because legislators regarded federal courts as paper tigers with respect to policing egregious partisan gerrymander due to the Supreme Court's repeated failure to enunciate a manageable standard for unconstitutionality. Second, there could be process related evidence, e.g., secrecy of line-drawing, exclusion of input from the minority party, irrelevance of public input, or in those states where interrogation of legislators and staff involved in redistricting is permitted, evidence of instructions to line drawers that suggested a blatantly partisan process. Third, there could be indirect evidence of intent based on taking the form (e.g., conformity to good government criteria) and partisan consequences (or expected consequences) of a plan and positing that what is achieved or foreseeable by any reasonable person is purposive. Here, of course, we would need to distinguish purposively partisan effects of line drawing from consequences of "natural gerrymandering."

In our views, the evidence presented at trial is clear. You could get the 13R-5D outcomes observed in Pennsylvania in every election beginning in 2012 only with a really carefully designed pro-Republican gerrymander that, in Professor Chen's words, "subordinat[ed] traditional districting criteria in the drawing of congressional voting districts." In the words of the Pennsylvania Supreme Court's words, such a plan was an "extreme and durable" partisan gerrymander.

instructions about the criteria to be used for plans to be submitted to the Court

The Court Order in February 2018 inviting submissions of feasible maps mentioned only good government criteria: minimizing the number of county, township and municipality splits, and maximizing district compactness under several different measures of compactness. Because the Court requested no information about partisan consequences, this left open the possibility that one or more of the parties would submit the kind of "stealth gerrymander" that a *Monkey Cage* op-ed by Grofman and Cervas (February 9, 2018) warned the Court to watch out for. As they point out,

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

"gerrymandering is like cancer: sometimes the signs are obvious" (e.g., suspicious lumps); "sometimes the signs are more or less invisible" (prostate cancer). As they wrote: "Districts that are bizarrely drawn and unnecessarily fragment existing political boundaries such as townships and counties almost always indicate … gerrymandering. But sometimes districts satisfy good government criteria on their face, and nonetheless have egregious partisan intent and effects." The latter is what we mean by a *stealth gerrymander*.

Imposing good government constraints is something like imposing a handicap in a horse race, it forces skillful gerrymanders to work harder to win. As A Wuffle (personal communication, April 1, 2017) has opined: "If there are no constraints on gerrymandering imposed by good government standards rooted in geography, then it doesn't take a genius to gerrymander."[25] But as we also show, imposing good government constraints can affect the playing field by making it harder to overcome the bias imposed by "natural gerrymandering" in the interest of fairness.

neutral vs. fair redistricting

Defining gerrymandering is far from simple, but there is not that much disagreement that, carried to an extreme, it is a bad thing. Extreme gerrymandering has been referred to as a "pathology of democracy" (Shapiro 1985). The common-sense definition is that a partisan gerrymander places unequal burdens on voters' opportunity to elect their representatives based on the party with which they associate. A term often used in conjunction with gerrymandering is *vote dilution*. The LWV final order asserts (February 19, 2018, at p.1):

> This adjudication was based upon the uncontradicted evidentiary record developed in the Commonwealth Court, wherein the Petitioners established that the 2011 Plan was a partisan gerrymander and that this gerrymander was extreme and durable. It was designed to dilute the votes of those who in prior elections voted for the party not in power in order to give the party in power a lasting electoral advantage. … On this record, it is clear that the 2011 Plan violates Article I, Section 5, since *a diluted vote is not an equal vote* (emphasis added).

[25] Wuffle refers to this as the "5G principle."

But should we take into account the electoral geography that might operate to constrain the ability of a party whose strength is highly geographically concentrated from translating votes into seats, and thus, arguably seeing its votes diluted? In Pennsylvania, for example, Philadelphia is overwhelmingly Democratic in voting. In particular, if you draw two congressional districts entirely within Philadelphia County, one of them is very likely to give Democratic candidates around 90% of the vote, and for sure, as long as both are wholly within the County, the average vote in the two will be around 80% Democratic no matter how you draw the two districts. There are no other large, concentrated pockets of equivalently overwhelmingly Republican voting strength. Thus, more Democratic votes will "naturally" be "wasted" in Philadelphia than Republican votes will be "wasted" elsewhere in the state. Additional Democratic wasted votes come in heavily Democratic Allegheny County, the county in which the city of Pittsburgh is located.

One approach to defining gerrymandering that addresses this issue is to define partisan gerrymandering as occurring when a districting plan penalizes members of the minority party in their ability to translate votes into seats *compared to what might be expected from a plan drawn on the basis of neutral principles*. By *neutral principles* we mean satisfying the one person, one vote standard, avoiding diluting the voting strength of protected racial or ethnic minority groups, minimizing unnecessary political subunit splits, and providing reasonably compact districts. The definition given above proposes to define partisan gerrymandering with respect to a baseline of "neutral" treatment based on the actual electoral geography and based on satisfying good government criteria for redistricting. Such a standard is different from a standard of "fairness" based on having some reasonable relationship between the proportion (or expected proportion) of congressional or legislative seats captured by a given party and that party's share (or expected share) of the statewide two-party vote, such as the expectation that the gap between a party's seat share and 50 percent party should not be much more than twice the difference between its vote share and 50 percent) as posited, for example, by the *efficiency gap* literature) .

The reason that a standard of neutral treatment and a standard of fair treatment cannot be expected to be the same is that, as in Pennsylvania, one party may have its

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

electoral strength less efficiently allocated geographically than is true for the other party, and the imposition of good government criteria may make it more difficult to compensate for this bias in the direction of partisan fairness. Thus, a "neutral" plan may reflect what has sometimes been called a *natural gerrymander*, i.e., a plan that results in a discrepancy between a party's vote share and its seat share that can be attributed solely to the geographic distribution of a party's electoral support and not to any intentional partisan gerrymandering. In Pennsylvania, for the U.S. House of Representatives, *ceteris paribus*, as we discuss below, Republicans pick up about a one seat advantage from "neutral" redistricting using good government criteria, so if the state is 50-50 in partisan vote share, a "neutral" plan might not be 9R-9D, but more like 10R-8D.[26]

Because electoral geography can matter, distinguishing purposive effects of line drawing from consequences of "natural gerrymandering" can be critical for courts that are seeking to assess partisan intent indirectly by looking at partisan consequences and positing that what is achieved/expected, and foreseeable by any reasonable person, can be seen as intended. But even if partisan intent is not at issue, as is true in Pennsylvania, a court that is paying attention to partisan consequences when it is drawing a remedy plan might have to decide to what extent it is seeking to draw a "neutral" plan as opposed to drawing a "fair" plan as its remedy -- using the definition of those terms given above.[27]

What does it take to comply with the Pennsylvania State Constitution?

Under the standard for unconstitutionality laid down by the Pennsylvania Supreme Court (a) though particular districts might be particularly egregious, evidence of unconstitutional gerrymandering can be found in a plan as a whole; (b) evidence of purposeful discrimination was not necessary in order to find unconstitutionality;[28] (c)

[26] Chen and Rodden (2013), for example, estimate this bias as about 1.45 seats (8%).

[27] We provide further discussion of the partisan consequences of electoral geography when we discuss the adopted map in comparison to other proposed maps later in the article.

[28] However, if we look at the Court's final Order of February 18, 2018, at p. 1, the Court asserts that the 2011 plan "was *designed* to dilute the votes of those who in prior elections

evidence of discriminatory partisan effects was not necessary to show a violation; [29] and (d) evidence that a plan complied with good government criteria is almost certainly not sufficient to rule out a finding of unconstitutionality, even though a finding that a plan failed to comply with such criteria was sufficient to invalidate the plan.

As noted above, the trial court found many indicia of partisan gerrymandering, but declined as a matter of law to hold that these allowed it to formulate a manageable standard for when partisan gerrymandering rose beyond politics as usual to the level of unconstitutionality. Moreover, when experts (such as Professor Pegden or Professor Kennedy) offered an opinion on the "ultimate question of law –*i.e.*, whether the 2011 Plan is an unconstitutional political gerrymander, that opinion was "disregarded" by the trial court (see e.g., Recommended Findings of Fact and Conclusions of Law, slip op., p. 83). As for going beyond the statistical facts to assert that these were indicia not merely of partisan gerrymandering but of unconstitutional partisan gerrymandering, we agree fully with the trial court that determination of the latter is a matter of legal judgment as to what is the appropriate standard of unconstitutionality. However, while courts must be the architects of manageable standards by deciding which features of a plan need to be examined as part of a test for unconstitutional partisan gerrymandering, we believe the lower court was wrong to dismiss the findings of experts that various features of the plan (e.g., undue fragmentation of county boundaries) were so extreme that one could statistically rule out the likelihood that this feature could have been due to chance along. The kind of two (or three) standard deviation test used by various of the expert witnesses in this case is compatible with federal court jurisprudence in areas of civil rights such as employment discrimination (e.g., *Bazemore v. Friday*, 478 U.S. 385, 398-402, 1986) and racial discrimination in jury selection *(e.g., Castaneda v. Partida*, 430 U.S. 482, 492-95, 1977).

What expert witnesses can do is (a) to provide the analytic framework and metrics that courts can draw upon in crafting legal standards, (b) provide case-specific

---

voted for the party not in power in order to give the party in power a lasting electoral advantage" (emphasis added),

[29] However, if we look at the Court's final Order of February 18, 2018, at p. 1, the Court asserts that the 2011 plan was "a partisan gerrymander… that was extreme and durable."

measurements of the factors that courts come to regard as the relevant ones, and (c) provide analysis of the degree to which given features of a plan are extreme in statistical terms.[30] Unlike the lower court, the Pennsylvania Supreme Court made use of expert witness testimony in exactly these ways, using not just common sense measurements such as the number of city or county cuts, but relying also on the results of expert witness use of analytic tools for measuring compactness, and statistical tools for determining when observed patterns in the data counted as statistical outliers for potential use in a court-drawn test to distinguish the legitimate use of legislative discretion from choices that were unconstitutional.

Once the Pennsylvania Supreme Court found there to be unconstitutional partisan gerrymandering, in choosing a remedy plan we had thought that it would need to confront the legal question of whether a remedy for that unconstitutional plan is a "neutral" plan or a "fair" plan, as those terms are defined above. But the Pennsylvania Supreme Court successfully ducked this question by the simple expedient of not saying anything at all about the partisan implications of the plan it adopted, or even about the partisan implications of the plans they rejected that were submitted by the parties and intervenors.[31]

In a case about partisan gerrymandering, quite remarkably, the only evidence about partisanship referenced in the *LWV* opinion and orders of the Pennsylvania Supreme Court is about the partisan implications of the 2011 plan. Instead, since violation of good government criteria is held to be legally sufficient to strike down a plan,

[30] This is exactly the incremental path followed by courts in areas in such as minority vote dilution. There, expert witness testimony about factors such as racially polarized patterns of voting and level of minority electoral success eventually leads (in *Thornburg v. Gingles* 478 US 30, 1986) to an enumeration of necessary conditions for a violation of Section 2 in the form of a three-pronged test.

[31] If the Pennsylvania Supreme Court had chosen to inquire about the partisan consequences of proposed plans, in examining those projected consequences, the Court would also have needed to determine how much weight, if any, it will give to having certainty about projected outcomes. If you want to draw districts that are maximally predictable, then you draw few if any competitive seats, since seats with margins close to 50% can flip in either direction depending upon the partisan directionality of electoral tides. But, if you draw few competitive districts, then you are essentially freezing outcomes.

the focus of the Court's inquiry has been on the good government aspects of the remedy plans submitted to it. Indeed, its endorsement of the plan that it ultimately adopted is couched entirely in good government terms (Opinion and Order, February 19, 2018, at pp. 6-7, with internal footnotes omitted), with no discussion of its partisan implications.[32] Describing the adopted plan, the Court says:

> It is composed of congressional districts which follow the traditional redistricting criteria of compactness, contiguity, equality of population, and respect for the integrity of political subdivisions. The Remedial Plan splits only 13 counties. Of those, four counties are split into three districts and nine are split into two districts. The parties, intervenors, and *amici* differ in how they calculate municipal and precinct splits, and, as noted earlier, the Legislative Respondents suggest that updated data on precinct and municipal boundaries does not exist. The Remedial Plan is superior or comparable to all plans submitted by the parties, the intervenors, and *amici*, by whichever Census-provided definition one employs (Minor Civil Divisions, Cities, Boroughs, Townships, and Census Places)
>
> The compactness of the plan is superior or comparable to the other submissions, according to the Reock, Schwartzberg, Polsby-Popper, Population Polygon, and Minimum Convex Polygon measures described in the Court's January 26 Order. Here, too, the parties, intervenors, and *amici* disagree on the precise ways to calculate these measures, and some failed to deliver compactness scores with their submissions. By whichever calculation methodology employed, the Remedial Plan is superior or comparable. Finally, no district has more than a one-person difference in population from any other district, and, therefore, the Remedial Plan achieves the constitutional guarantee of one person, one vote.
>
> Accordingly, this 19th day of February 2018, the Court orders as follows: First, the Pennsylvania primary and general elections for seats in the United States House of Representatives commencing in the year 2018 shall be conducted in accordance with the Remedial Plan…

[32] We should also note that the Court in its adopted plan pairs two Democratic incumbents in Philadelphia (though one, Rep. Brady, was already intending to retire,), three others who are retiring (Reps. Meehan, Dent, Shuster), and the vacant incumbent seat in old district 18 with an incumbent now running for the Senate (Rep. Barletta). Its final Order does not even bother to mention the consequences for incumbents of the adopted plan. Presumably this is because the 2011 was so tainted overall by what we might call "partisan greed" that no deference is required to the contorted lines in it that yielded those incumbencies. But, since the Court says nothing about its reasons for not mentioning incumbent effects, we can only guess.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

As we show later, in a good year for Republicans, the Joint plan submitted to the Court by the state legislature is as bad or almost as bad for Democrats as the 2011 plan, and in a good year for Democrats it is just as bad or even worse (see Table 1 below). Moreover, the Governor's plan is uniformly either better or no worse for Democrats than the Joint Legislative plan (see Table 1). Thus, it is not surprising that Republican Joint legislative plan was rejected by the Governor, presumably because of its partisan implications. But this plan was also rejected by the Court, with no mention of its partisan implications. The Court's stated reasons for rejecting the Joint Legislative plan was because it was not as good on good government grounds as the Court-ordered plan.

Thanks to the change in partisan control of the governorship, all the plans submitted to the Court lacked the imprimatur of being an adopted plan and thus could be more easily dismissed. In particular, had the joint plan offered by the legislature been the Commonwealth of Pennsylvania's actual plan in 2011, it is an open question as to whether traditional norms of legislative deference might have spared it from a finding of unconstitutionality. Except with respect to one person, one vote, legislatures are not usually seen as required to draw the "best possible" congressional plan. But, in the absence of a State plan with a claim to legitimacy, the Court picked the plan among those before it that most closely satisfied good government criteria – which happily, thanks to Professor Persily's expertise, turned out to be its own plan, and thus a plan which the Court could know with certainty was not intended as a partisan gerrymander for either party.

comparison of the LWV approach to that of federal judges in Wisconsin and North Carolina

The closest federal voting right law parallel to the standard laid down by the PA Supreme Court is "one person, one vote" case law. Under federal law, you do not need to show that malapportionment matters for any particular end, you just need to show it exists.[33] Now, in Pennsylvania, you do not need to show that violation of good

[33] Litigants also do not need to show that the malapportionment found was intended but, since calculating malapportionment is a matter of 4th grade arithmetic, it is hard to see how any malapportionment found could not have been intended.

government redistricting criteria had substantive political or racial consequences, or that it was intentional --all you need to show is that it gave you lines on a map that lack adequate legitimate justification in terms of traditional good government districting criteria. In contrast, the two federal courts that have struck down plans as unconstitutional partisan gerrymanders, in Wisconsin and North Carolina, have taken quite different approaches to what is needed to show unconstitutionality -- ones requiring plaintiffs to establish extreme and durable partisan effects that were intentionally created. Of course, we must also acknowledge that the Pennsylvania Supreme Court, in rejecting the 2011 plan, also characterized it as having "extreme and durable" partisan effects. But as the quote provided earlier demonstrates, the finding of its unconstitutionality did not rest on those partisan aspects of the 2011 congressional plan but instead on a blatant violation of good government criteria that was inexplicable by any legitimate considerations, and whose parameters could be characterized as statistical outliers.

However, we would also emphasize one important parallel between the state court and the federal court approaches in the states where a gerrymandering plan was overturned. Federal courts, like the Pennsylvania Supreme Court, rooted their opinions in a profound repugnance to egregious partisan gerrymandering and a belief that it violates the spirit of democracy. This idea is captured in the North Carolina majority opinion's assertion (*Common Cause v. Rucho,* No. 1:16-CV-1026, January 9, 2018, slip op. at p. 3 quoting *Ariz. State Leg.*, 135 S. Ct. at 2677) that "On its most fundamental level, partisan gerrymandering violates 'the core principle of republican government . . . that the voters should choose their representatives, not the other way around'".

One last point re comparisons of the Pennsylvania decision with that of other courts recently considering partisan gerrymandering challenges: We believe that the approach taken in Pennsylvania has potential relevance to the Supreme Court reviews of the decision in Maryland that dealt with only a single congressional district. In Maryland, a preliminary injunction to stop the implementation of the challenged map was rejected by the trial court *(Benisek v. Lamone*, Civ No, JKB-13-3233, D. Maryland, August 24, 2017), in an opinion that looked toward a resolution of the constitutional issues when the Supreme Court decides *Gill*.

The unsuccessfully challenged district, while having issues both with respect to existing boundaries and with respect to compactness, is not as "ugly" vis-à-vis compactness and other good government grounds as some of the districts that were not challenged in the case. Rather, the heart of the challenge to the district was that particular areas of the state were included in or excluded from the challenged district entirely on partisan grounds, with one and only one motive, to change the partisan outcome in of the challenged district. Essentially, the argument is that changes in a plan that can be understood only as an exercise of unbridled partisan lust are *ipso facto* unconstitutional. This argument has much the same flavor as the Pennsylvania Supreme Court's view that plans that blatantly violated good government criteria with no legitimate justification were *ipso facto* unconstitutional. And, evidence at trial in Pennsylvania by Professor Kennedy addressed the partisan implications of boundary shifts for some of the districts in the Pennsylvania congressional plan with evidence that was similar to that produced in the Maryland case. There are, of course, two key differences. First the Pennsylvania Supreme Court rejected an entire plan and relied on evidence from the plan as a whole. Second, the Pennsylvania decision relied on criteria that did not require assessment of either partisan consequences or of partisan intent.

While we believe that it is best to consider partisan gerrymandering in a plan as a whole, we recognize that racial gerrymandering in the form of packing or race as a preponderant motive is often addressed in a district specific way,[34] and there are legal approaches to partisan gerrymandering that do consider the potential for unconstitutionality based solely on the way in which lines have been drawn in a single district. In particular, the Florida Constitution has a provision inserted by initiative, Art. III, § 20(a), that mandates that "[n]o apportionment plan or individual district shall be drawn with the intent to favor or disfavor a political party or an incumbent." Such a standard, as interpreted in *League of Women Voters v. Detzner*, 172 So. 3d 363, 427 (Fla. 2015), could lead to the invalidation of a district such as Maryland's 6th Congressional district. But, if the Supreme Court were to affirm one or both of the lower court opinions in the Wisconsin and North Carolina cases, or issue a standard of its own to judge

[34] See e.g., *Thornburg v. Gingles*, 478 U.S. 30 (1986), *Shaw v. Hunt* 517 U.S. 899 (1996), *Page v. Va. State Bd. of Elections* (Page II), (E.D. VA. June 5, 2015).

statewide results, and if that standard treated a consistent one seat effect as potentially being unconstitutional, it would make sense for Maryland plaintiffs to file an amended complaint challenging the statewide congressional map under that new standard. Wang (2016a), among others, has analyzed the Maryland map, and found evidence of gerrymandering. His results indicate that "[congressional] redistricting gained Democrats a one-seat advantage in 2012, a strong Democratic year, and that this advantage was retained in the national wave election of 2014 that swept dozens of Republicans into office in states outside Maryland" (Wang, 2016a: 1314).

III. Evaluation of the Adopted Plan in Comparison with Other Plans

To place the actual Court ordered plan in contrast with the 2011 plan in context with other possible remedy plans, we will look at the Joint plan proposed by the Republicans leaders of the two legislative chambers; the Governor's proposed plan as a Democratic alternative; a plan drawn by the liberal political entity, *DailyKos*, that relies entirely on good government criteria; and three plans drawn by the authors, two of which (V1 and V2) are also drawn entirely on the basis of good government criteria,[35] and one of which, V3, though relying heavily on good government criteria, looks also to the issue of partisan fairness. For each of these plans, we provide in Table 1 a comparison of a number of features: the number of county splits, since the Court specifically sought to limit the number of counties spanning multiple districts; two measures of compactness, since compactness is also a criteria specifically referenced by the Court; and anticipated two-party vote shares for each of the districts. based on the one hand, on a summary of the projections of the presidential results in 2008 (a good Democratic year) and 2016 (a good Republican year), respectively, in each of the districts and, on the other hand, on a composite of state-wide races in 2008 and 2016, respectively. An overview of the plans according to each of these criteria is provided in Table 1.

[35] We include two good government plans drawn by the authors to demonstrate that good government plans can have a range of potential political consequences. Of course, this point is made even more compellingly by the expert witness testimony at trial, using computer-based simulations of hundreds of potential "good government" maps.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

<<Table 1 about here>>

We turn next to the implications of the data provided in Table 1 for the degree to which the plan satisfies the good government criteria of minimizing county splits and drawing compact districts. For the latter consideration, we will briefly review the properties of the two measures of compactness we use. Then we look at the more accurate synthetic projection method of likely election results, based on five statewide elections in 2016, and examine the results of that projection method at the district level. Then we discuss how the various election projections we have provided bear on the question of the degree to which each plan shows evidence of being a partisan gerrymander.

COUNTY SPLITS

That the 2011 map is an eyesore vis-a-vis unnecessary splits of county boundaries is revealed clearly in Table 1. All the other maps, including the Joint Legislative map proposed by Republicans in response to the Court's request for submission of possible remedy plans, and the Governor's map, a Democratic response, do a dramatically better job in preserving county boundaries. However, among potential remedy plans, the latter two maps are the worst with respect to this criterion, though the differences across the new plans we compare are relatively small.

COMPACTNESS

To measure the compactness of the districts, we utilize two well-known measures, the Polsby-Popper measure, which shows how smooth or contorted the boundaries are, and the Reock ratio, which is a degree to which a district is spread out from its centroid (Reock 1961, Polsby and Popper 1991).[36] We took the average value across the 18 districts for each plan. The Polsby Popper measure ranges from zero to one, with zero

[36] A(D) = area of district, P(D) = perimeter of district, Circle(D)= minimum circumscribing circle; POLSBYPOPPER = $4\pi A(D)/P(D)^2$**,** REOCK = A(D)/A(Circle(D))

indicating non-compactness and one indicating a perfect circle. It is simply the ratio of the district area to the area of a circle with the same perimeter. Reock is the ratio of a district's area divided by the area of the minimum circumscribing circle. Like Polsby-Popper, higher numbers indicate more compactness.

As Table 1 shows, the old map adopted in 2011 is the least compact, by far, according to both measures. The Republican Legislature's Joint plan is the second least compact among the plans analyzed, again according to both measures. While there are plans that are marginally better than the Court adopted plan on one of the two measures, in the set we consider, there is no plan that is superior to the Court plan on both measures.

Figure 1 provides an intuitive visual way to understand how the two measures of compactness need not give the same evaluation. To do so we have rescaled the size of the illustrative district relative to the two comparison circles. The main difference between the two is that the Reock is the area of the district relative to a circle that fits the maximum width of district, where the Polsby-Popper is the area relative to the circle with the same perimeter. To better illustrate the features of the various plans we compare, we present maps for each, drawn to the same scale, in Figure 2.

<<Figures 1 and 2 about here>>

PARTISAN EFFECTS OF THE PLANS

evaluation of the 2018 plan in comparison with the 2011 plan

At trial, as previously noted, Professor Chen used a baseline of five elections from before 2011 for his projections; for the analyses offered here, we primarily use a composite of five statewide elections in 2016 to do projections, though we also report projections based solely on presidential data. District estimates are projected from precinct level results. The five 2016 elections used in the composite are President, U.S. Senate, Attorney General, Treasurer, and Auditor. We utilize the average of the five contests because Republicans won the statewide vote in two of them (President, U.S. Senate), while the Democrats won the statewide vote in three of them (Auditor,

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Treasurer, and Attorney General).[37] We also include, for comparison purposes, the projected data for four composite elections 2008. The 2016 composite measure gives a realistic expectation of baseline partisanship in the districts.[38] The composite of the five 2016 elections correctly predicted the five Democratic victories in 2016 under the map adopted in 2011. [39]

Using the 2016 composite, among the 18 districts in the Court map, five are strongly favored by the Democrats, seven are strongly favored by the Republicans, and six are competitive (+ or - 5%).[40] In the 2011 map, Democrats were strongly favored in only four districts, while Republicans were favored in eight. In the map adopted in 2011,

[37] We compiled the composite measure using both the sum of the raw votes by party and an average of the two-party vote in the four or five races at the district level. Both ways result in the same district partisanship breakdowns in all plans except for the Joint legislative plan in 2008. In that election, the breakdown goes from 11R-7D for raw votes, as reported in Table 1, to 10R-8D for the composite measure averaged across districts.

[38] As noted above, there is dispute in the academic literature about what elections to use in making projections, and in what other factors, if any, to include in the projections. Nonetheless, there can be no dispute that those doing the redistricting make use of past election data to make judgements about expected partisan outcomes and commonly do so with considerable confidence. Our view is that there is no *a priori* way to determine optimal projection methods prior to actually running analyses of their predictive accuracy. We also recognize that projection methods that appear deterministic are nonetheless probabilistic in their predictions, even if that point is not made explicitly. Moreover, for most plans, we would expect to get different outcomes in very good years for Republicans than in very good years for Democrats. As discussed above, what makes for a durable gerrymander is the absence of competitive districts. This absence operates to constrain the responsiveness of the plan to changes in voter preferences and puts limits on the possible success of the minority party in other than tsunami–like wave elections. For alternative perspectives see e.g., McGhee (2018), Best et al (2018b: esp. at p. 83).

[39] The precinct level data come from Harvard Dataverse for 2008, and from Nathaniel Kelso's github for 2016.

[40] For a discussion about thresholds of competitiveness, see Cervas and Grofman (2017).

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

the Republicans were able to consistently win 13 of the districts in every election, despite a statewide partisan split that straddled the 50% line.[41]

We show projections under this composite five election measure in Table 2. Here, we provide more detail than in Table 1 by ranking the districts from least Democratic to most Democratic in projected two-party vote share. We see from Table 2 that the Court plan, *ceteris paribus*, is on average expected to yield a 9D-9R split. However, because of its high number of highly competitive seats this single number is somewhat misleading. As noted above, using the five-election model as our predictor, there are fewer safe Democratic seats (5) than safe Republican seats (7) in the Court plan, using a five-percentage point definition of a competitive seat. This is one fewer safe seat for Republicans and one additional safe Democratic seat than is found in the 2011 plan, but the Court plan also has five districts balanced on a knife edge, with a projected winner's margin under two percentage points, making it highly responsive to changes in voting patterns. Thus, in the Court drawn plan, changes in vote choice or turnout [42] might result in different outcomes, especially in the most marginal of districts. [43]

---

[41] In 2012, the Republican share of the popular vote in the congressional elections in Pennsylvania was 49.2%, in 2014 it was 55.5%, and in 2016 it was 53.9%.

[42] We also do not take incumbency into account in our projections. If there is an incumbent in place, and if the district boundaries are not that much changed by redistricting, districts with highly competitive baseline partisan support will be biased toward the reelection of that incumbent. However, incumbency advantage has declined in recent years to lows which have not been seen since the 1950s (Jacobson 2015). The 2016 general election average we used to calculate the projected votes include three state-wide races where the Democratic candidates did particularly well compared both to Hillary Clinton and to the Democratic candidate for U.S. Senate. Similarly, in Pennsylvania in 2016, Republican congressional incumbents out-polled President Trump in 9 of the 13 districts won by Republicans.

[43] Pennsylvania Supreme Court Judges are elected or retained to ten-year terms in elections with party labels on the ballot. Judges with Democratic affiliations now constitute the majority of the Court. As noted earlier, the Court plan was attacked by Republicans as a partisan gerrymander on behalf of Democrats. For the reasons given above we are quite skeptical of this claim of partisan bias. Furthermore, given the geographic concentration of Democrats in Pittsburgh and Philadelphia, a true pro-Democratic gerrymander would have unnecessary splits of county and municipality boundaries in the areas of Allegheny and Philadelphia counties that are absent in the Court plan. Given the probabilistic nature of projections, our own simulations reinforce

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

<< Table 2 about here>>

evaluation of the 2018 plan in comparison with Republican and Democratic plans rejected by the Court

The Joint Submission, which was a map sent to the governor but not passed by the legislature, was really a collaboration of the Republican leaders of the two chambers, Rep. Turzai and Sen. Scarnati. While the plan adhered far more closely to the principles laid out by the court than the 2011 map, the governor called it a partisan gerrymander and rejected it. He issued a statement saying "… my preference would have been for the General Assembly to send me a fair map" and submitted his own map to the Court.[44] The Court rejected both in favor of the plan drawn by the Court's consultant.

Using the 2016 five election composite projection, the Joint Legislative plan would have delivered an average of seven Democratic seats, the same as the Governor's plan. However, there are very important differences between the two plans once one looks at the details. One main difference between the Joint plan and the Governor's is that among the seven districts likely to be won by Democrats, the Democratic percentages are cut more narrowly in the Republican plan than in the Governor's plan in three districts so that, in a good Republican year, Republicans will do better under their plan than under the Governor's plan. The second main difference between the plans, which operates in the same direction, is that the next four most competitive seats, where a Republican is projected to win under our composite measure, have the Democratic vote margin lower under the Joint legislative plan than under the Governor's plan. Thus, in a good Democratic year, the Democrats will likely not fare as well under the Republican plan as under the Governor's plan. In sum, whether it's a good year for Republican or a bad year for Republicans, the Republicans will fare better under their own Joint plan than under the Governor's plan. [45]

---

the view that the partisan consequences of the Court plan are within the parameters expected of a non-partisan plan.

[44] https://www.governor.pa.gov/governor-wolf-submits-a-fairer-congressional-map-to-supreme-court/

[45] The degree to which the Joint plan is what we have been calling a *stealth gerrymander* can be best understood by looking at the good Democratic year of 2008 in Table 1 where,

evaluation of the 2018 plan in comparison with DailyKos plans

The liberal blog, the *DailyKos*, also submitted a plan seeking to implement the criteria that the Court laid out.[46] Using our composite measure, as shown in Table 2, their plan created five safe Democratic districts, eight safe Republican districts, and five competitive districts. Based on the five-election model, in 2016 the Democrats would have been predicted to win eight seats and the Republicans 10 under this plan.[47]

evaluation of the 2018 plan in comparison with plans drawn by the authors

We have drawn two alternative plans for the House using the good government criteria required by the court in which we paid absolutely no attention to partisan information, and a third plan that started with plan 1 as a base and attempted to increase the number of Democratic districts by making small changes based on precinct level maps of 2016 election results at the presidential level. All three plans score well on good government criteria. These plans have between five and six competitive seats. In a good Democratic year such as 2008, as shown in Table 1, the projected result in these plans would have been 7R-11D in Plan 1 and 8R-10D in Plans 2 and 3 using presidential results. In a good Republican year such as 2016, we estimate the result would have been 11R-7D in Plan 1, 10R-8D in Plan 2, and 11R-7D in Plan 3, using only presidential results. Using the 2016 five-election synthetic projection method, we find results of 10R-8D in Plan 1, 11R -7D in Plan 2, and 9R -9D in Plan 3. Comparing 2008 and 2016 projections, these plans demonstrate responsiveness to shifting preferences of the electorate.

---

in the non-partisan plans (Court Remedial, the two Authors' Plans, DailyKos), the Republicans averaged under 7 districts using 2008 presidential and composite election projections, but in the two Republican-drawn plans, they win 11 seats in both plans. In contrast, in the Governor's plan, the Democrats are projected to win three fewer seats in a good Republican year (2016) than in a good Democratic year (2008). See Table 1 and discussion of alternative plans below.

[46] *DailyKos* has prepared a number of other plans, but for space reasons we will only consider their non-partisan map. The DailyKos has long advocated for redistricting reform that creates good government plans that resemble proportional representation.

[47] There are also illustrative maps proposed by academics, by various interest groups, and by journalists (such as *FiveThirtyEight)*, that we also have not space to review.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

IV. Lessons for the future

foreseeing litigation in the 12 states with similar state constitutional provisions to the one relied upon in Pennsylvania

While the Pennsylvania court opinion is limited to Pennsylvania, and thus it might seem of only limited importance, in footnote 71 of the Opinion (slip op. pp.116-117), the court took what we regard as a rather unusual step. It issued what can only be called an invitation to other state courts to use the same logic it used to invalidate partisan gerrymanders in their own state. The court noted that there are twelve states whose constitutions contain election clauses identical to the Pennsylvania charter, requiring elections to be "free and equal": Arizona, Arkansas, Delaware, Illinois, Indiana, Kentucky, Oklahoma, South Dakota, Oregon, Tennessee, Washington, and Wyoming.[48] While only a handful of these states are ripe for partisan gerrymandering challenges -- in that (ca. 2017) three lack unified party control of the state, two are single district states, and some already have a commission drawing plans and, in others, indicia of gerrymandering are missing -- it is hard to imagine that the February 7 Pennsylvania decision won't trigger new partisan gerrymandering challenges in at least one state court in this set of states, perhaps even quickly enough to potentially impact some elections in 2020.

We show in Table 3 various information that can help us form beliefs about whether or not the congressional plans in these dozen states are partisan gerrymanders However, as pointed out in the table notes, some measures are not well suited for calculation when the seats at issue are few in number.

<< Table 3 about here >>

Moreover, in looking to the future, we should note that the two states not yet sued that are generally regarded as among the most pernicious of the not yet so far challenged

[48] In its Opinion, the Court provided specific citations to each of these provisions which we have not bothered to reproduce.

partisan gerrymanders, Michigan and Ohio (see e.g. Wang 2016a; Royden and Li, 2017) are not included among the twelve, nor is Maryland. Furthermore, we should also should note that a "free and equal" elections clause is not the only avenue state courts might use to attack partisan gerrymanders in the future. As University of Kentucky College of Law Professor Joshua Douglas has pointed out, virtually every state constitution protects voting rights more explicitly than the U.S. Constitution does. In addition to the thirteen states that require elections to be "free and equal," an additional thirteen have state constitutional provisions that require elections to be "free and open," and this clause could, in principle, be used in exactly the same way as the "free and equal" clause.[49]

stealth gerrymandering, yes; blatant gerrymandering, no

In Pennsylvania in the next decade of redistricting, and in any other state that takes the same approach to partisan gerrymandering as has the Pennsylvania Supreme Court, it will simply not be possible to repeat the kind of obvious gerrymander vis a vis tortuous line and fragmentation of county and municipal boundaries that we saw in the 2011 Pennsylvania congressional plan that was struck down – and that we saw in many other states. Moreover, even in states that lack the specific state constitutional provision relied on by the Pennsylvania Supreme Court, after the LWV opinion demonstrated the vulnerability of plans to state court challenge, we expect that risk averse legislators now may not want to take chances that their plan -- carefully crafted gerrymander or not --will be struck down. Thus, we expect good government criteria to be adhered to much more strictly in the 2020 round of redistricting than was the case in the 2010 round.

But, as noted previously, following good government criteria has two consequences. First, it requires more skill to do a partisan gerrymander, though it certainly does not eliminate partisan gerrymandering if plans are being evaluated solely on good government criteria and not in terms of partisan impact. In this case, *stealth gerrymandering* (i.e., gerrymandering operating within the constraints of satisfying good government criteria), of the sort we saw in the Joint Plan in Pennsylvania, will replace overt gerrymandering. Second, if the minority party is disadvantaged by having its

[49] We are indebted to Jonathan Lai of the *Philadelphia Inquirer* (personal communication, April 2018) for calling this information to our attention.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

electoral strength more geographically concentrated than that of the majority party, it is very likely to suffer the consequences of what we above referred to as *natural gerrymandering*. Indeed, even if it is the party with free rein to do the districting because of its control of both the legislature and the governorship, the party disadvantaged by having it electoral strength more geographically concentrated than that of the other party may not be able to fully compensate for this disadvantage by skillful line drawing.

However, LWV also offers hope that, at least in Pennsylvania, even stealth gerrymanders will not be invulnerable. The reason for this optimism is that this case makes violation of good government criteria to a statistical level of extremeness a sufficient condition for a plan to be an unconstitutional partisan gerrymander but does not make it a necessary condition. Those who wish to put limits on partisan gerrymandering can take heart from the Pennsylvania Supreme Court's finding that the 2011 congressional plan was an "extreme and durable" partisan gerrymander, even though that finding was apparently not directly relevant to the Court's decision to strike down the plan as unconstitutional.

In the presence of a stealth gerrymander rather than the blatant gerrymander before it, the Pennsylvania Court might well have relied on partisan impact evidence to reach a conclusion of unconstitutionality. The language of the Final Order implementing the Court's own plan strongly suggests this possibility. There, the Court said about the 2011 plan that it "was designed to dilute the votes of those who in prior elections voted for the party not in power in order to give the party in power a lasting electoral advantage. In stark contrast, Article I, Section 5 of the Pennsylvania Constitution provides: 'Elections shall be free and equal; and no power, civil or military, shall at any time interfere to prevent the free exercise of the right of suffrage.' Pa. Const. art. I, § 5. On this record, it is clear that the 2011 Plan violates Article I, Section 5, since a diluted vote is not an equal vote" (slip op. p. 2). Moreover, given the Pennsylvania Supreme Court's willingness to determine that a plan was a severe and extreme gerrymander on the basis of the expert witness presented at trial in LWV, we believe that may be easier in the future to prove partisan gerrymandering in a state court in Pennsylvania and in other states that adopt the Pennsylvania approach than is likely to be true under federal law,

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

even if the Supreme Court adopts standards similar to those proposed by the three judge court in *Whitford v. Gill*.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

REFERENCES

Best, Robin E., Shawn J. Donahue, Jonathan Krasno, Daniel B. Magleby, and Michael D. McDonald. 2018a. "Considering the Prospects for Establishing a Packing Gerrymandering Standard." Election Law Journal, 17(1): 1-20.

Best, Robin E., Shawn J. Donahue, Jonathan Krasno, Daniel B. Magleby, and Michael D. McDonald. 2018b. "Authors' Response–Values and Validations: Proper Criteria for Comparing Standards for Packing Gerrymanders." Election Law Journal, 17(1): 82-84.

Brady, David W. and Bernard Grofman. 1991. Modeling the Determinants of Swing Ratio and Bias in U.S. House elections, 1850-1980. Political Geography Quarterly, 10(3): 254-262.

Chen, Jowei and Jonathan Rodden. 2013. ''Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures.'' Quarterly Journal of Political Science, 8(3): 239–69.

Cervas, Jonathan R., and Bernard Grofman. 2017. "Why Noncompetitive States are so Important for Understanding the Outcomes of Competitive Elections: the Electoral College 1868–2016". Public Choice, 173: 251-265.

Cervas, Jonathan R., and Bernard Grofman. "Pennsylvania has to draw new congressional districts, but getting rid of gerrymandering will be harder than you think." *Monkey Cage*, *Washington Post,* 9 Feb. 2018.

Gelman, Andrew, and Gary King. 1990. "Estimating Incumbency Advantage without Bias". American Journal of Political Science, 34(4): 1142-1164.

Gelman, Andrew, and Gary King. 1994. "Enhancing Democracy through Legislative Redistricting". American Political Science Review, 88(3): 541-559.

Grofman Bernard. 1983. Measures of Bias and Proportionality in Seats-Votes Relationships. Political Methodology, 9: 295–327.

Grofman, Bernard and Thomas Brunell. 2005. "The Art of the Dummymander: The Impact of Recent Redistrictings on the Partisan Makeup of Southern House Seats." In Galderisi, Peter (Ed.) Redistricting in the New Millennium, New York: Lexington Books, pp. 183-199.

Grofman, Bernard and Gary King. 2007. "Partisan Symmetry and the Test for Gerrymandering Claims after *LULAC v. Perry"*. Election Law Journal, 6 (1): 2-35.

Guinier, L. and G Torres. 2009. The Miner's Canary: Enlisting Race, Resisting power, Transforming Democracy. Harvard University Press (originally published 2002).

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Jacobson, Gary C. 2015. "Decline of the Incumbency Advantage". Journal of Politics, 77(3): 861-873.

Kendall, M. G. and A. Stuart. 1950. The Law of the Cubic Proportion in Election Results. The British Journal of Sociology 1(3): 183-196.

King, Gary. 1989. "Representation Through Legislative Redistricting: A Stochastic Model". American Journal of Political Science, 33(4): 787-824.

McDonald, Michael D., and Robin E. Best. 2015. Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases, Election Law Journal, 14: 312-330.

McGann, Anthony J., Charles Anthony Smith, Michael Latner, and J. Alex Keena. 2015. "A Discernable and Manageable Standard for Partisan Gerrymandering". Election Law Journal, 14(4): 295-311.

McGann, Anthony J., Charles Anthony Smith, Michael Latner, and J. Alex Keena. 2016. Gerrymandering in America: The House of Representatives, the Supreme Court, and the Future of Popular Sovereignty. Cambridge University Press.

McGhee, Eric. 2017. "Measuring Efficiency in Redistricting.'' Election Law Journal, 16(4): 417–442.

McGhee Eric. 2018. Rejoinder to "Considering the Prospects for Establishing a Packing Gerrymandering Standard." 17 Election Law Journal. 73-82.

Niemi, Richard G., Bernard Grofman, Carl Carlucci and Thomas Hofeller. 1990. Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering. Journal of Politics, 52(4): 1155-1181.

Owen, Guillermo and Bernard N. Grofman. 1988. Optimal Partisan Gerrymandering. Political Geography Quarterly, 7(1): 5-22.

Reock, E. (1961). A Note: Measuring Compactness as a Requirement of Legislative Apportionment. Midwest Journal of Political Science, *5*(1): 70-74.

Royden, Laura and Michael Li. 2017. "Extreme Maps: Partisan Bias and Gerrymandering in the Congressional Maps of the 2010 Cycle," Brennan Center Report, New York, NY. https://www.brennancenter.org/publication/extreme-maps

Shapiro, Martin. 1985. "Gerrymandering, Unfairness, and the Supreme Court." 33 UCLA Law Review. 227-256

Theil Henri. 1970. The Cube Law Revisited. Journal of the American Statistical Association, 65: 1213–1219.

Tan, Netina and Bernard Grofman. "Electoral Engineering and Electoral Manipulation." Presented at the 7th Annual Conference on Mathematics and the Social Sciences, University of Auckland, New Zealand, February 18, 2016.

Tufte Edward R. 1973. The Relationship between Seats and Votes in Two-Party Systems. American Political Science Review, 67: 540–47.

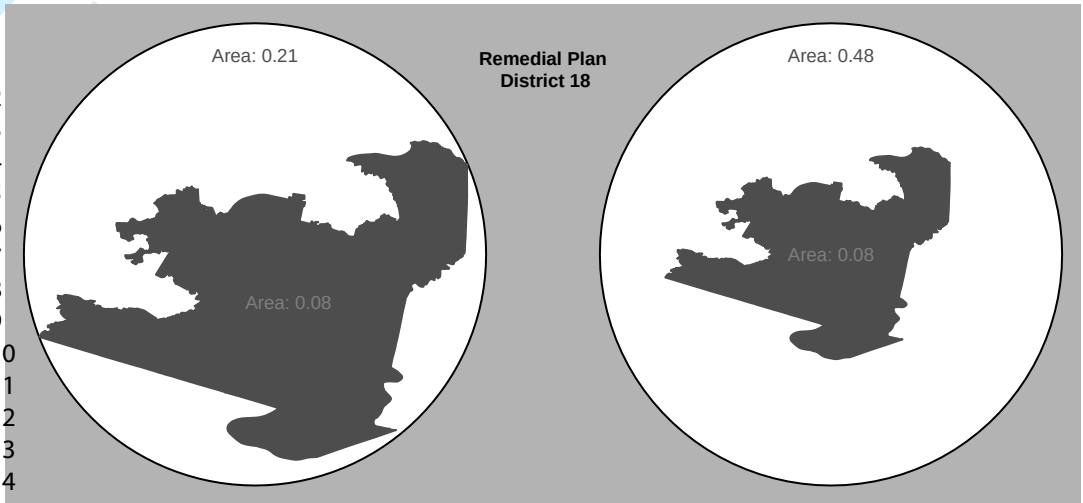
Wang, Samuel 2016a. Stanford Law Review. 68: 1263-1321.

Wang, Samuel S. H., 2016b. Three Practical Tests for Gerrymandering: Application to Maryland and Wisconsin. Election Law Journal, 15: 367-384.

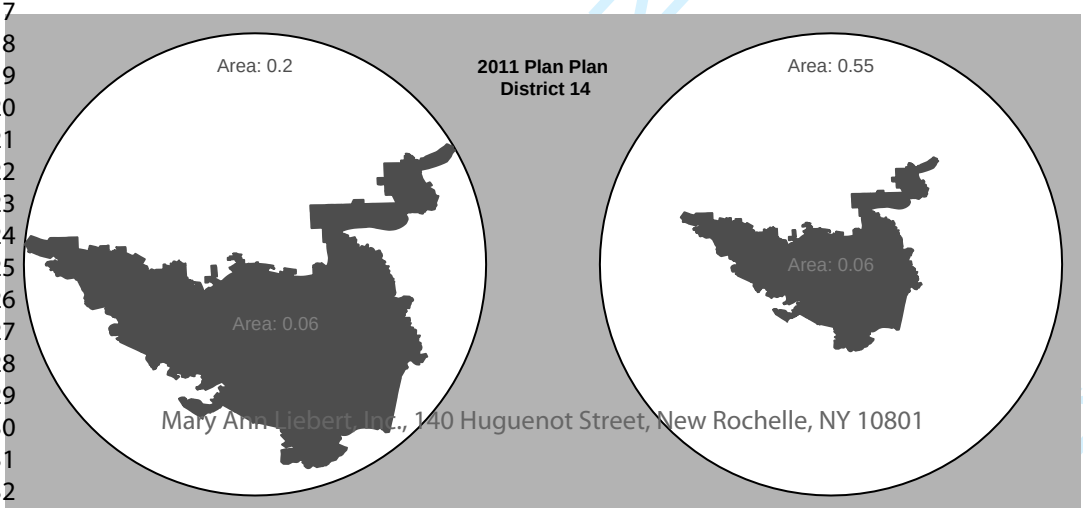
Warrington, Gregory S. 2018. "Quantifying Gerrymandering Using the Vote Distribution." Election Law Journal, 17(1): 39-57.

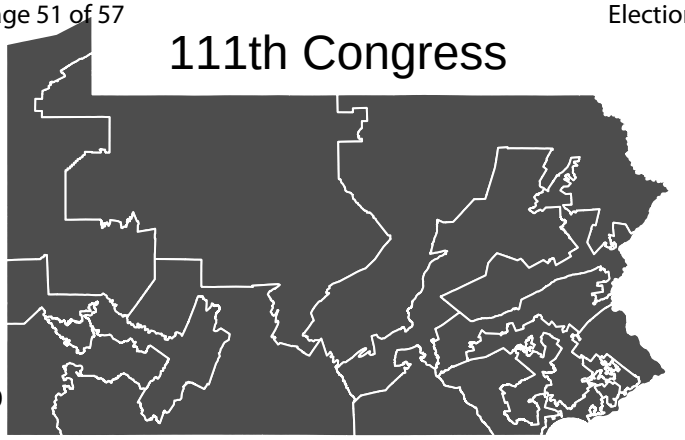
Wuffle, A. 2015. Uncle Wuffle's Reflections on Political Methodology. PS. January: pp. 1-7.

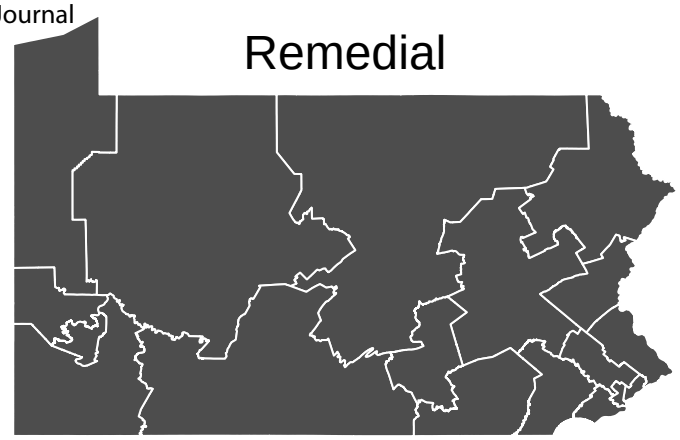
Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

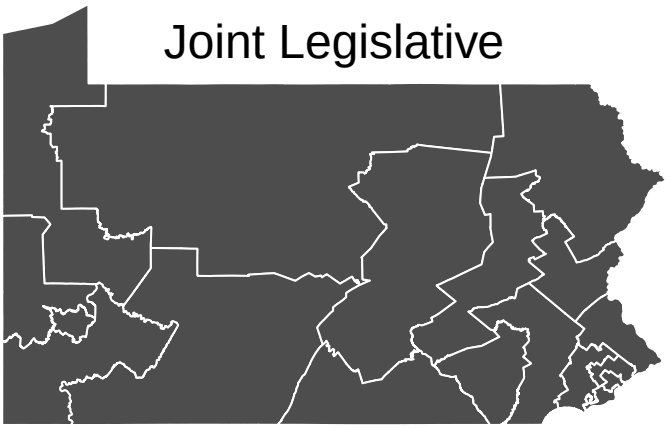


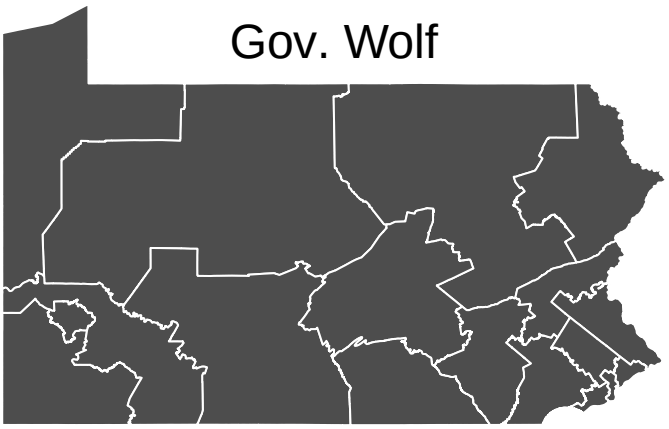
Reock **Polsby−Popper**

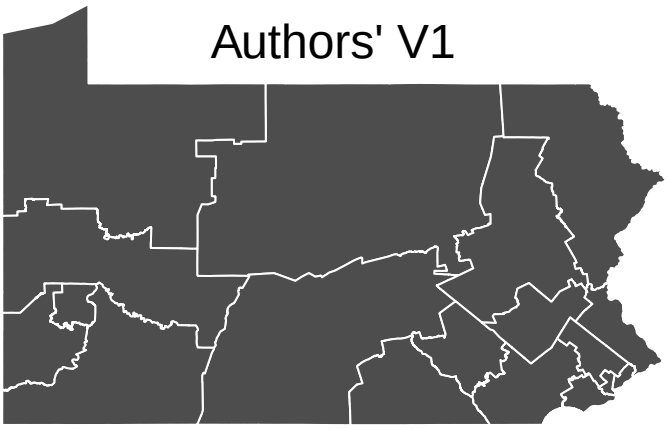












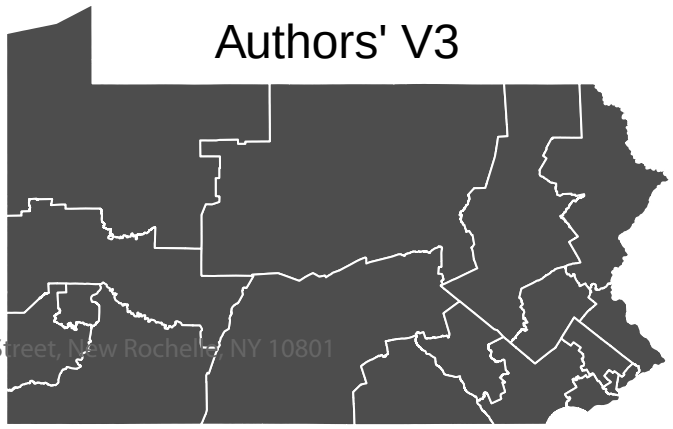


Table 1: County Splits, Compactness Scores of the Plans, and Projected Outcomes using Elections in 2008 and 2016

	County Splits	Polsby-Popper	Reock	Projected using 2008 Presidential Results	Projected using four state-wide elections in 2008	Projected using 2016 Presidential Results	Projected using five state-wide elections in 2016
2011 Map	41	0.164	0.271	8R-10D	11R-7D	12R-6D	13R-5D
Court Remedial	17	0.321	0.425	8R-10D	6R-12D	10R-8D	9R-9D
Joint Submission	19	0.282	0.368	9R-9D	11R-7D	12R-6D	11R-7D
Gov. Wolf	19	0.307	0.394	8R-10D	8R-10D	11R-7D	11R-7D
DailyKos	15	0.356	0.414	8R-10D	7R-11D	11R-7D	10R-8D
Authors' 1	18	0.342	0.383	7R-11D	6R-12D	11R-7D	10R-8D
Authors' 2	16	0.362	0.415	8R-10D	7R-11D	10R-8D	11R-8D
Authors' 3	15	0.359	0.407	8R-10D	6R-12D	11R-7D	9R-9D

Note: County splits include all the pieces in which a county is split, not just the total number of counties split (The latter number is the one most often reported in both court documents and in the media.) The 2008 four-election average includes the Presidential, Attorneys General, Auditor, and Treasurer contests. The 2016 five election average includes President, U.S. Senate, Attorney General, Treasurer, and Auditor. State-wide Democratic percentage of the two-party vote are as follows: 2008 Presidential (55.2%), 2008 Composite (54.7%), 2016 Presidential (49.6%), 2016 Composite (51.3%).

Table 2: Projected Partisanship of the 18 Districts among the Different Plans According to a Five-Election Composite of Statewide 2016 elections (districts arranged from least to most Democratic)

2011 Map	Court Remedial	Joint Legislative	Gov. Wolf	DailyKos	Authors' good govt V1	Authors' good govt V2	Authors' V3
34.30%	30.80%	34.70%	33.60%	32.10%	29.60%	34.90%	29.60%
34.80%	33.90%	34.80%	34.50%	34.00%	36.90%	34.90%	37.20%
38.50%	34.70%	35.30%	36.10%	35.00%	37.50%	36.30%	37.50%
40.20%	36.80%	35.50%	37.10%	35.90%	39.20%	37.00%	39.20%
40.30%	37.10%	41.40%	38.60%	38.80%	40.60%	37.90%	40.60%
42.10%	42.10%	42.30%	41.60%	41.50%	41.00%	41.20%	40.90%
44.40%	42.70%	43.00%	41.80%	42.90%	41.90%	42.80%	41.00%
44.70%	46.10%	43.80%	44.10%	43.30%	45.60%	44.90%	45.70%
45.50%	49.80%	46.20%	48.00%	47.90%	47.00%	46.00%	46.90%
46.70%	51.20%	47.30%	48.10%	49.90%	49.60%	49.30%	50.10%
47.80%	51.50%	49.20%	49.10%	50.10%	50.30%	49.80%	50.30%
47.90%	51.50%	50.10%	52.20%	51.30%	50.50%	50.40%	50.40%
48.70%	51.50%	50.90%	53.10%	52.20%	50.90%	50.60%	50.90%
50.70%	57.30%	51.00%	56.90%	55.70%	59.00%	59.30%	59.00%
66.80%	61.90%	65.90%	67.40%	60.20%	60.20%	60.50%	60.20%
71.40%	67.00%	71.30%	69.40%	69.20%	66.30%	70.80%	66.30%
81.30%	75.90%	82.30%	74.90%	85.80%	75.90%	74.80%	75.90%
90.50%	91.40%	91.10%	88.40%	89.60%	91.40%	92.70%	91.40%
13R-5D	**9R-9D**	**11R-7D**	**11R-7D**	**10R-8D**	**10R-8D**	**11R-7D**	**9R-9D**

Least Democratic ----->

<-----Most Democratic

Table 3: Information about Twelve States with Constitutional Provisions similar to Pennsylvania

State	CDs	Unified Control (2011)	Unified Control (2016)	Seats (2016)	Votes (2016)	Mean DEM District Vote Share	Median DEM District Vote Share	Efficiency Gap	Who does districting
Arizona	9	X (R)	X (R)	44.4%	48.1%	50.4%	49.4%	0.08	Independent commission
Arkansas	4	X (D)	X (R)	0%	35.7%	35.5%			state legislature
Delaware	1	X (D)	X (D)	100%	55.9%	56%			NA
Illinois	18	X (D)		61.1%	59.0%	59.6%	59.8%	0.08	state legislature
Indiana	9	X (R)	X (R)	22.2%	39.9%	40.1%	35.9%	0.09	state legislature
Kentucky	6		X (R)	16.7%	34.3%	33.8%			state legislature
Oklahoma	5	X (R)	X (R)	0%	30.7%	30.7%			state legislature
Oregon	5		X (D)	80%	56.2%	56%			state legislature
South Dakota	1	X (R)	X (R)	0%	34.0%	34%			NA
Tennessee	9	X (R)	X (R)	22.2%	36.4%	37.4%	31.3%	0.03	state legislature
Washington	10	X (D)	X (D)	70%	58.8%	56%	52.3%	-0.05	5-member independent commission
Wyoming	1	X (R)	X (R)	0%	24.3%	24.3%			NA

NOTE: Seats and votes are based on the 2016 five-election projection (to deal with the existence of non-contested congressional districts). Percentages are of the Democratic two-party vote. Columns labelled "Votes" and "Seats" are based on state-wide raw vote results, while "Mean" and "Median" are central tendencies of district level results. This difference is why the percentages reported in columns 6 and 8 are not identical. Data from *DailyKos*, *All about Redistricting*, and *Ballotpedia* web sites. States with fewer than 9 districts do not have efficiency gap or median values reported because of the potential unreliability of those calculations given the small number of districts involved.

Methodological Appendix

This appendix describes in more detail how we measure different indicia found in the literature. These measures include the *efficiency gap* (McGhee, 2017), the *mean/median gap* (Best et al., 2018a), and partisan symmetry calculations (Grofman and King, 2007; McGann et al. 2015, 2016). The first two of these measures are straightforward, though estimates will differ depending upon the elections from which they are calculated.

The Efficiency Gap is calculated as defined in McGhee (2014), where all the party's votes are wasted if they lose the district, and all the winner's votes over 50% are wasted. The difference between each party's wasted votes is then divided by the total votes cast to produce the Efficiency Gap, with a value of zero denoting what is regarded as ideal. As noted in the text, this is equivalent to taking an aggregate swing ratio of 2 as ideal.

The Mean/Median gap is the difference between the average vote percentage for a party and its share in the median district when districts are sorted according to two-party vote share. This measure is a variant of skewness, such that when the mean is substantially higher or lower than the median, this is indicative of bias.

However, there are a number of different ways to estimate partisan bias based on the shape of the votes to seats distribution. Here we calculate partisan bias by centering the contest at a 50% vote share for both parties, then incrementally adding (or subtracting) one percentage point to find aggregate seat outcomes under these differing mean vote shares. The resulting (s,v) points on the votes to seats curve are then converted to a log odds form by using log(s/1-s) as the dependent variable, and log(v/1-v) as the independent variable, and the points that fall between 40% and 60% vote share are entered into a regression. Partisan Bias is then calculated from the intercept of this regression using an exponential transformation (for details see Grofman, 1983). Partisan bias is calculated from the Democratic perspective, so a negative bias indicates bias against Democrats. We also examine the standard errors to determine the probability that the observed bias is not due to random chance. Table 3 also reports the *swing ratio*, a measure of responsiveness that is defined as the slope of the same regression used to generate partisan bias.

Table 3 in the text reports the measures of gerrymandering described above. The 2011 adopted plan is the worst on all three measures, by far. The Joint Legislative plan proposed by

the Republican leaders of the state legislature is second worst on two of the three measures, and third worse on the other. By way of contrast, the Remedial Court adopted plan of 2018 scores the best on two of three measures, and second best on the other.

We also offer an alternative way of presenting district specific vote share or projected vote share data, of the sort provided in Table 2 of the main text. Figure A1's shadings indicate the extent to which the districts are tilted toward one or the other party.

<< Figure A1 about here >>