ELECTION LAW JOURNAL
Volume 18, Number 4, 2019
Mary Ann Liebert, Inc.
DOI: 10.1089/elj.2018.0537

Locating the Representational Baseline: Republicans in Massachusetts

Moon Duchin, Taissa Gladkova, Eugene Henninger-Voss, Ben Klingensmith, Heather Newman, and Hannah Wheelen

ABSTRACT

Republican candidates often receive between 30% and 40% of the two-way vote share in statewide elections in Massachusetts. For the last three Census cycles, Massachusetts has held 9–10 seats in the House of Representatives, which means that a district can be won with as little as six percent of the statewide vote. Putting these two facts together, it is striking that a Massachusetts Republican has not won a seat in the U.S. House of Representatives since 1994. We argue that the underperformance of Republicans in Massachusetts is not attributable to gerrymandering, nor to the failure of Republicans to field House candidates, but is a structural mathematical feature of the actual distribution of votes observable in some recent elections. Several of these elections have a remarkable property in their vote patterns: Republican votes clear 30%, but are distributed so uniformly that they are locked out of the possibility of representation. Though there are more ways of building a valid districting plan than there are particles in the galaxy, every single one of them would produce a 9–0 Democratic delegation.

Keywords: redistricting, area: United States, political parties

INTRODUCTION

Gerrymandering IS THE PRACTICE of using the formation of electoral districts to create a representational advantage for some subsets of the population, or to favor certain kinds of candidates. In recent years, gerrymandering has received increasing levels of attention and public indignation. There are essentially two indicators that are taken by the public and by many commentators as red flags for gerrymandering: *bizarre shapes* and *disproportional outcomes*. For instance, the enacted 113th congressional districting plan in Pennsylvania contained a notorious district nicknamed "Goofy kicking Donald Duck," whose contorted shape was taken by many as prima facie evidence of redistricting abuse. Under this map, Pennsylvania elections exhibited nearly 50–50 splits in party preference, while Republicans held 13 out of 18 seats, or over 72% of the House representation. While there is indeed compelling evidence that Pennsylvania was gerrymandered in a partisan manner (Pegden 2017; Duchin 2018), this fact is not established by either shapes or disproportions alone. In this article, we show that there can also exist benign and structural obstructions to securing representation that have to do with not just the number of votes but how they

Moon Duchin is an associate professor of mathematics at Tufts University in Medford, Massachusetts, and co-director of the Voting Rights Data Institute (VRDI) research program. Taissa Gladkova, Eugene Henninger-Voss, Ben Klingensmith, Heather Newman, and Hannah Wheelen are VRDI Fellows.

We gratefully acknowledge the Amar G. Bose Research Grant and PI Justin Solomon for major support of the Voting Rights Data Institute. We thank Gabriela Obando and William Palmer at the Massachusetts Secretary of State's office for their help collecting and interpreting data. Thanks also to Jowei Chen for sharing a dataset that approximates precinct-level vote counts in 2000, to Gary King for extremely useful feedback, and to Max Hully and Ruth Buck for excellent data and technical support.

are distributed around the state. We mean this in the technical sense of "obstruction"—in a departure from much of the political science literature, we are not discussing a tendency or likelihood, but a mathematical certainty of securing zero representation.

This article is framed to study a "riddle" about Republican voting patterns in Massachusetts: *why is 1/3 of the vote proving insufficient to secure any representation?* By contrast, the classic "cube law" predicts 1/9 of the seats, the logit model from Chen and Rodden (2016) predicts roughly 2/9 of the seats, and proportional representation would be 1/3 of the seats for Republicans in this situation. In fact, the partisan lopsidedness of voting in Massachusetts is in some ways comparable to that in several other states of similar size (e.g., Arizona, Maryland, and Tennessee have comparable U.S. Senate statistics[1]) but none of those other states has ever sent a one-party delegation to the House at any point in the last 20 years, while Massachusetts does so in every election. The core of our analysis is a rigorous proof that certain actual observed voting patterns guarantee this lockout effect, regardless of the districting plan. This illustrates the intuitive principle that *uniformity itself can block desired representational outcomes for a group in the numerical minority* (like Republicans in Massachusetts), considering both the numbers and the geometry. Though this is mathematically obvious when taken to an extreme, exhibiting actual voting patterns with this level of uniformity is a novel finding.

Massachusetts is one striking case in point, but the broader message is that once the rules have been set, it becomes a scientific question to study the breadth of partisan outcomes left available to the districters. This case study describes a surprising limitation on the power to control the representational outcome. In other cases there will be other surprises, such as an extremely wide latitude of seats that a party can secure with a given pattern of votes by carefully constructing the district lines, or simply a baseline of seat outcomes in a non-intuitive range. This article contributes to the emerging viewpoint that it is only legitimate to compare an observed partisan outcome against the backdrop of actual possibility.[2]

It is very important that we state clearly that this analysis rests on the study of votes and not voters. We make no claims about the true baseline partisanship of the *people* of Massachusetts, but rather we fix particular election outcome data, one race at a time, and vary district lines. For instance, Massachusetts voters are clearly willing to elect Republican governors, and have done so in three of five elections since 2000. We focus our study on presidential and U.S. Senate votes, with no attempt to control for incumbency or other factors, because they give an ample supply of instances with R share in the 30%–40% range in recent years, and the article is focused on how the distributional effects of cast votes interact with the ability to gerrymander in that range.

Numerical uniformity. We use the phrase "numerical uniformity" to describe a situation in which the vote shares across the building-block units are extremely consistent. In the second section, we examine the numerical distribution of votes in 13 statewide elections in Massachusetts, showing that for five of them, the numbers alone make it literally impossible to build a R-favoring collection of towns or precincts with enough population to be a congressional district. Because this type of analysis is run on the numbers only, this result is very strong: no district-sized grouping can be formed, even without requiring contiguity, compactness, or any other spatial constraint on districting. The reason is that elections in which Republicans are locked out exhibit extremely *low variance* in the town- and precinct-level voting results.[3] At the very extreme, you could imagine that Republicans have 35% of the vote statewide, and in each town, and in each precinct—and the reality is closer to that extreme than one might guess. In particular, even in some elections in which a Republican received 30%–40% of the overall vote, the R vote share rarely exceeded 50% in any precinct, leaving not enough R-favoring precincts to assemble into a grouping of the size of a congressional district.

[1]No two states have exactly matched voting data, but all four of these have a Senate tilt in the 61%–63% range, a consistent presidential tilt in the same direction, and 8–10 House seats over the last two Census cycles. See GitHub (2018), <https://github.com/gerrymandr/party-tilt>, for Senate data. Uncontested congressional races make that vote data unsuitable to compare directly.

[2]J. Chen, W. Cho, M. Duchin, J. Mattingly, and W. Pegden have all recently supplied expert reports to that effect for legal action in states from Florida to Pennsylvania to Wisconsin to Ohio to North Carolina.

[3]Note that this is the variance of the dataset itself, not of a fitted distribution.

Geometric uniformity. On the other hand, "geometric uniformity" describes a situation in which partisan preference does not correlate strongly to location within the state, reflected in the absence of partisan enclaves or clusters. In the third section we will add a spatial component to our analysis. Even when it is numerically feasible to collect enough precincts to form an R-favoring district, the precincts may not be spatially located in such a way that this can be accomplished in a connected (i.e., contiguous) fashion. We first show visualizations that illustrate the lack of a Republican enclave in the low-variance elections, suggesting low correlation between location and partisanship in these Massachusetts elections.[4] To corroborate this, we compute clustering scores (which measure segregation of Republican votes from Democratic votes). We find that the actual vote distributions in 2000–2010 have clustering levels that are similar to those that would be observed if placing the Republican votes by drawing randomly from a uniform distribution around the state, but that clustering has increased over time, which suggests directions for future work that relates vote clustering to district shape criteria. Geometric uniformity may be contributing to partisan underperformance above and beyond numerical uniformity, and focused study of that effect on its own will be quite valuable, but it does not drive the effect that we observe here.

In short, the conclusion is that extreme representational outcomes are not always attributable to gerrymandering, nor to spatial clustering in the arrangement of voters from either party. Generally, counterintuitive limitations on representation can emerge from a complicated interplay of the numerical and spatial distribution of voter preferences; in the case of Massachusetts, the numerical distribution is so uniform that it makes the spatiality insignificant. The effects on representation of the distribution (and not just the share) of votes is a difficult mathematical question and is richly worthy of further study.

While public observers may expect proportional representation as a matter of fairness, even seasoned political scientists have often measured fairness in terms of universal representational indices. For instance, the efficiency gap, or *EG*, can be described as measuring parity of wasted votes, but is fundamentally measuring whether the seat share S is close to $2V - 1/2$, where V is the vote share. The efficiency gap, $EG = 2V - S - 1/2$, has been argued to flag a legally actionable gerrymander when its magnitude is more than 8%. But the Massachusetts data contain five actual vote distributions (Pres 00, Pres 04, Sen 06, Pres 08, Sen 08) for which even an omniscient redistricter with the honorable goal of $EG = 0$ could not succeed: not a single one of the many quintillions of possible nine-district plans has an efficiency gap below 11% in any of those five races. This shows that finding a reasonable baseline to decide when gerrymandering has occurred is a subtler problem than has so far been appreciated in the public discourse or in some of the mainstream political science literature. A broader and more detailed review of the literature can be found below in the subsection titled "Relationship to previous literature."

Data

Massachusetts is home to about 2%–3% of the nation's population, with 6,349,097 people in the 2000 Census and 6,547,629 people in 2010. After the 2010 Census, the number of congressional delegates apportioned to Massachusetts dropped from 10 to 9 because the state's population growth did not keep pace with the country's.

Massachusetts is made up of 351 jurisdictions that we will call *towns* (also written in some places as townships or municipalities), which have not changed over the timespan covered here. Towns do not overlap, and they completely cover the state; in this language, cities are large towns. We obtained a town shapefile from MassGIS that has population attributes from the Census (MassGIS 2019).

Each town is subdivided into some number of *precincts*—the level at which election results are reported—ranging in number from 2,166 in 2002 to 2,174 in 2016 according to the Secretary of State's database (Massachusetts Secretary of State 2019). In 2016, 125 towns were not subdivided (the town equals one precinct), and at the other extreme, Boston was made up of 254 precincts, followed by Springfield with 64. Note that the precincts used to administer actual elections are similar but not identical to Voting Tabulation Districts or VTDs, which are snapshots of precincts reported by the Census Bureau every ten years

[4]In physics terms, partisanship has low entropy in Massachusetts.

(U.S. Census Bureau 2016). The Secretary of State's office provided us with a state-modified VTD shapefile from 2010 and directed us to Census resources that cover the earlier period. Unfortunately, small changes to precincts are common between elections. We cleaned the data to obtain precinct shapefiles that can be held constant over Census cycles, which involves making a small number of merges. This produced two shapefiles: one with 2,156 precincts and 2002–10 election results, and one with 2,151 precincts and 2012–16 election results.

In the tables below, the cast vote data comes from the Secretary of State's website (Massachusetts Secretary of State 2019). They offer town-level election results back to the year 2000 and much earlier, but precinct-level results only back to 2002. For population numbers, Census 2000 population figures were used for elections taking place 2000–08, and Census 2010 for 2010–16. Town-level population was included in the MassGIS data. The Secretary of State's shapefile included VTD population numbers with a nearly exact name match to the tabular data; a script was only needed to switch from a multi-column format to a single-column format, and the very few non-exact matches were handled by hand with no ambiguity by human standards. We double-checked the population data against the Census Application Programming Interface (API) to be confident in its quality. For the six elections (2002–10) not covered by that data, we used a python preprocessing tool to compare the shapefiles of census blocks and precincts (Metric Geometry and Gerrymandering Group 2019c). This computes the assignments of blocks to precincts and aggregates block population up to precincts.

All of our data, together with scripts needed to run the various algorithms described here, can be found in the public GitHub repositories of the Voting Rights Data Institute (Metric Geometry and Gerrymandering Group 2019a, 2019b, 2019c).

Setup choices: Election data, number of districts, smallest units, constraints

In order to illustrate this effects of uniformity observed in real voting data, we run a districting-feasibility analysis (described in full detail in the Appendix) on election results from 13 presidential and U.S. Senate elections in Massachusetts. Endogenous (congressional) election results are not considered here because many of the recent races are uncontested. For example, in the 2016 U.S. House election, five out of nine districts had no Republican who filed to run (Ballotpedia 2016). Therefore, two-way vote share analysis would not be meaningful for these races, though we note that our focus on Republican share of 30%–40% is generous with respect to available congressional data. The analysis could certainly be extended to other statewide races, including governor, attorney general, and secretary of state if desired; we chose a collection of races that demonstrates interesting distributional effects in the 30%–40% range of Republican share.

Many political scientists have debated whether statewide races are good predictors of congressional voting patterns, and if so, which ones are most predictive. That debate is beside the point for this analysis, which is focused on the range of representational outcomes that are possible for given naturalistically observed partisan voting patterns. We will also choose to analyze the seat share possible out of nine congressional districts for the sake of consistency, even though our timespan of electoral data includes a period over which the apportionment varied between nine and ten. Neither decision blunts the impact of the findings, which study the extent to which empirical patterns in actual voting data can restrict the range of representation that is possible for a group in the numerical minority.

In the numerical feasibility section we will only require that districts hew close to the standard of equal population and that they are made of whole units, such as towns and precincts. Because of the central role of real voting data in this analysis, we are bound to use precincts as the smallest building blocks, since that is the smallest level at which vote returns are available. In practice, the 2011 congressional plan held 2119 precincts intact while splitting 32, which means that fewer than 1.5% were split. Using towns or precincts as unsplittable building blocks does have some precedent in law and practice. As a historical matter, the state constitution of Massachusetts (2019) did require in Article XVI that state councillors be elected from contiguous districts that keep towns and city wards intact, but this system of councillors is now obsolete.

There is a still-active contiguity requirement for state legislative districts, and a rule to preserve towns as much as is "reasonable," but no formal contiguity or unit-preservation requirement for congressional districts. In fact, only 23 states have a

contiguity requirement for congressional districts, while 49 require contiguity for legislative districts. Nonetheless congressional district contiguity is essentially universal in practice.[5] Shape constraints will only be discussed in the geometric section of the paper (the third section).

One possible interpretation of our fencing-out findings is that they primarily identify the partisan consequences in Massachusetts of putting a heavy weight on the traditional districting principle that guides districters to avoid splitting municipalities in their plans. A rule requiring at least some weight on respecting political boundaries is almost as common as the contiguity rule, featuring in some 19 states for congressional districts and 49 states for legislative districts.[6] We do not think that this is the extent of the conclusion that can be appropriately drawn, however. First, as we will develop below in the second section, there is practically no change in the feasibility analysis when moving from towns to precincts as building blocks, even though there are more than six times as many precincts. Second, the strength of the findings here, which show that in fenced-out elections the most Republican-favoring collection of precincts falls far short of ideal district size, all but guarantees that under actual current districting practices (contiguity, reasonable compactness, and under 1.5% of precincts split) the fence-out would remain in force.[7] Thus we find robust support for the broader conclusion that the representational baseline for single-member districts is strongly dictated by the specific political geography of each time and place.

Relationship to previous literature

This article concerns a surprising relationship between the vote share V for a political party across the districts of a polity and the seat share S that it is *possible* to earn, as the district lines vary. It is worth situating this work with respect to a sizeable political science literature that broadly seeks to capture (V, S) data points observed in actual elections by treating S as a function of V whose graph is a curve.

Published in 1950, Kendall and Stuart (1950) is the classic reference on the origins and status of the so-called *cube law* that holds that $\frac{S}{1-S} = \left(\frac{V}{1-V}\right)^3$. The cube law was still the dominant framework for understanding the votes-to-seats relationship in the 1970s, with papers such as Taagepera (1973) seeking generalizations. Gudgin and Taylor's (2012) important text in 1979 rebooted the cube law with an enhanced analytical derivation and some discussion of the role of geography in deviations from its predictions. However, the authors did not develop tools to measure the contributions of spatial statistics. The logic of the book circulated around the cube law itself and the idea that deviations are reasonably *defined* as partisan bias, whether attributable to spatial distribution or gerrymandering.

Rae's 1967 text avoids a fully cubic fixation but devotes itself to a philosophically similar search for rules and principles in the votes-to-seats conversion effected by districts and other systems in the form of twenty "Propositions." Tufte's 1973 article is particularly theoretically rich and stakes out a very influential new direction. Instead of a search for a grand unified law or rule, Tufte makes the case for curve fitting. He develops the mathematics for a power relationship between vote odds and seat odds but ultimately argues for fitting in the linear class, with particular attention to the slope or "swing ratio" of the best-fit line to the (V, S) data. (In a sense, efficiency gap falls in the Tufte tradition but approaches bias with the opposite logic, *declaring* an optimal slope of 2 and measuring deviation from that line.) King and Browning (1987) is one article in an abundant literature taking up the curve-fitting charge. The authors seek joint estimates of the exponent and coefficient from (V, S) data points, then develop statistical devices for introducing second-order "disturbance terms."

Grofman et al. (1997) set out to do something new, disaggregating the effects of geographic distribution, turnout, and malapportionment. But their ingredients are once again only district vote totals and seat shares, which makes them unable to learn what is possible from a given distribution of votes as district lines change. Recognizing this, they explain that their distribution term will not distinguish

[5]District contiguity can be made somewhat complicated by water and by smaller geographic units that are themselves disconnected, but these issues are relatively easy to resolve in Massachusetts. Districting rules may be found in the Massachusetts Constitution (2019) and at Levitt (2019a), <http://redistricting.lls.edu/states-MA.php>.

[6]See Levitt (2019b), <http://redistricting.lls.edu/where-tablefed.php>.

[7]For a detailed analysis of the effects of raising and lowering the priority on various districting criteria with another set of empirical vote data, see also Duchin (2018).

"the chance effects of geography [from] intentional gerrymandering" (461). More recently, a cluster of papers has asked new kinds of questions about the relationship between votes and seats that do probe the effects of geographic distribution over subunits, starting with the observation in Calvo and Rodden (2015) that the American political science literature "has given surprisingly little attention to the geography of party support" (791). That paper is still premised on a functional approach to the (V, S) relationship, but notably introduces a Gini coefficient (a statistic that is not geometric but numerical in the sense of this paper) as part of a discussion of predicting the degree of majoritarian bias.

Finally, two papers of Chen and Rodden (2013; 2016) seek to scope out typical districting outcomes by generating samples of districting plans with computers. However, they simultaneously use a logit model that introduces significant noise to the observed pattern of votes. Since their algorithms do not sample exhaustively nor representatively, these methods could not produce a finding that a party is locked out of representation, even if the voting pattern is fixed. But by also adding significant noise to the votes, the authors decouple their analysis from the actual vote distribution even more starkly. For instance, (Chen and Rodden 2016, fig. 8) shows the results of their simulated elections in the 2008 presidential race in Massachusetts. They find Republicans securing over 20% of the seats on *average*, whereas we demonstrate below that no district lines whatsoever could have produced more than one-ninth (11.1%) Republican seats in that particular race. A logit model, among other probabilistic models of votes, is often employed to represent predicted change or uncertainty from one race to the next, but any such model risks obscuring lockout effects, such as the ones found below in Massachusetts Senate and presidential vote patterns over a full ten-year Census cycle.

Table 1. Statistics of Republican Two-Way Vote Share in 13 Statewide Elections in Massachusetts

		R share by town		*R share by precinct*	
Election	*R share*	*mean*	*variance*	*mean*	*variance*
Pres 2000	35.2%	39.70%	.0073	–	–
Sen 2000	25.3%*	29.15%	.0043	–	–
Sen 2002	18.7%	20.29%	.0019	17.43%	.0028
Pres 2004	37.2%	39.99%	.0093	34.53%	.0140
Sen 2006	30.5%	33.23%	.0076	27.62%	.0118
Pres 2008	36.8%	39.00%	.0117	33.85%	.0179
Sen 2008	31.9%	34.40%	.0094	28.91%	.0141
Sen 2010	52.4%	53.78%	.0201	47.76%	.0307
Pres 2012	38.2%	41.05%	.0145	34.85%	.0227
Sen 2012	46.2%	49.19%	.0168	42.64%	.0274
Sen 2013	44.8%	48.89%	.0217	41.84%	.0311
Sen 2014	38.0%	41.14%	.0141	34.22%	.0205
Pres 2016	35.3%	40.17%	.0165	33.04%	.0235

Numbers are truncated rather than rounded. Lower-variance elections are marked in gray.
*Libertarian vote share included with R in 2000 Senate race.

The chief novel contribution of the present article is an extremely elementary technique that rigorously establishes much stronger bounds than had been previously available on the achievable partisan outcomes for a given distribution of votes; in particular, we show for the first time that multiple actual historical voting patterns featured a minority party with over 30% of the votes but no possibility of securing any seats at all, no matter how the lines are drawn.[8]

ARITHMETIC OF REPUBLICAN UNDERPERFORMANCE

In this section, we describe a method to determine theoretical bounds on the number of districts with a Republican majority, given only the geographical units, their population, and their vote totals for D and R candidates in a particular election. For this part of the analysis we impose no spatial constraints at all; we do not even require contiguity, but would allow a district constructed out of an arbitrary collection of towns or precincts from around the state. We show, for example, that *even though George W. Bush received over 35% of the two-way vote share against Al Gore, it is mathematically impossible to construct a collection of towns, however scattered, with at least 10% of the population and where Bush received more collective votes than Gore*. (See Figure 4.)

Remark (The Boston Effect). Note that in Table 1 the town-level mean R share reliably overshoots the statewide R share, while the precinct-level mean errs in the other direction. To see why, recall that there are 351 towns in the 2016 election, subdivided

[8]Taagepera (1989) addresses the question of how much vote share has been needed to produce the nonzero seat share, but the study considers only observed outcomes of historical elections and does not deal with alternative district lines. That is, it only looks at what has happened, and not at what is possible with alternative districts.

into 2,151 precincts. Boston is composed of 254 precincts; Springfield has 64; and most other towns have fewer than 25 precincts, with 126 towns (more than a third) having only one. This means Boston is an outlier in size, and it is also an outlier in the lopsidedness of its Democratic voting majority. (In the 2016 presidential election, Boston had only a 14.7% R two-way vote share.)

The town-level averaging underweights Boston because it is weighted equally to tiny towns like Gosnold (population 75). The precinct-level results overweight Boston because its average precinct population is under 2,500, lower than the statewide average of over 3,000. (Exact figures vary between the two census cycles.) This accounts for the direction of error in the mean of each statistic relative to the statewide share, which is naturally population-weighted.

As Table 1 illustrates, the elections from 2000 to 2008 had consistently lower variance in their town- and precinct-level vote shares than can be observed since 2010. Below, we will connect that to the representability of Republicans across these elections.

Numerical feasibility of R districts

Let's first review the limitations on the power of gerrymanderers that are produced by the numbers alone. We begin with very simplified algebraic argument that we will refer to as the *naive bounds* on gerrymandering. In an abstract districting system with equal vote turnout in its districts, if Party X receives share $0 \le V \le 1$ of the vote, its possible seat share S is constrained to a range, with the actual outcome depending on how the votes are distributed across the districts. At its most ruthlessly efficient, Party X could in principle have barely more than half of the vote in certain districts and no vote in the others, thus earning seat share up to $2V$, or twice its vote share. At minimum, a party with less than half of the vote can be shut out entirely by having less than half in each district; if Party X has more than half of the vote, then its opponent has a vote share of $1-V$ and a maximum seat share of $2(1-V)=2-2V$, so the minimum seat share for Party X is $1-(2-2V)=2V-1$. For example, a party with 40% of the vote can get anywhere from 0%–80% of the seats, while a party with 55% of the vote can get anywhere from 10%–100% of the seats. The naive bounds would project that districters could in principle arrange for Beatty voters in the 2008 Senate race to convert their 32% of the votes to 0%–64% of the seats. In sum, we have

Naive Bounds on Gerrymandering:

$$\begin{cases} 0 \le S < 2V, & V \le 1/2; \\ 2V-1 \le S \le 1, & V \ge 1/2. \end{cases}$$

But the naive bounds do not take into account constraints introduced by the fixed number of districts, by the variation in turnout, or by the discreteness of the building blocks. The feasibility analysis in this paper does account for all of those factors. Table 2 shows that in Ed Markey's 2013 special election to the Senate, his opponent's pattern in obtaining 38% of the vote could not have earned him any more than three district wins out of nine, no matter how the districts were drawn, despite the naive bounds that suggest up to six district wins could have been possible. And even more strikingly, though Jeff Beatty earned nearly a third of the vote against Kerry in the Senate race of 2008, Beatty voters in that distribution are actually locked out of representability entirely. The actual observed turnout patterns, and the effect of the mandate to build districts out of intact precincts, have lowered Beatty's ceiling from five districts out of nine all the way to zero. Smaller building blocks should mean more flexibility, but shrinking the building blocks from towns to precincts didn't in this case help Beatty at all.

Here is our method for measuring feasibility in our setup. Suppose that the ideal district size (state population divided by number of districts) is denoted by I. Then we will declare that it is *numerically feasible* for a party to get k seats in a certain election if there exists a collection of units (towns or precincts) with population at least $M=kI$ and in which that party has a majority of the two-way vote share. A feasibility bound for the party is the largest such k that has been demonstrated.

By contrast, we will say that it is *numerically infeasible* for a party to get m seats in a given election if there is proven to be no collection with population at least $M=mI$ and a majority for the party. An infeasibility bound is the smallest such m that has been demonstrated.

We use a simple sorting algorithm to get feasibility and infeasibility bounds for the elections considered here, presenting the results in Table 2. Often, but not always, the algorithm produces tight bounds,

TABLE 2. IF DISTRICTS WERE TO BE MADE OUT OF TOWNS OR OUT OF PRECINCTS, WITH NO REGARD TO SHAPE OR EVEN CONNECTEDNESS, HOW MANY R OR D DISTRICTS COULD BE FORMED? FEASIBILITY AND INFEASIBILITY BOUNDS ARE SHOWN IN THIS TABLE

				R feas/infeas		*D feas/infeas*	
Election	*D candidate–R candidate*	*R share*	*Seat quota (9 seats)*	*Town*	*Prec*	*Town*	*Prec*
Pres 2000	**Gore**–Bush	35.2%	3.2	0/1	—	9/-	—
Sen 2000	**Kennedy**–Robinson/Howell	25.3%*	2.3	0/1	—	9/-	—
Sen 2002	**Kerry**–Cloud	18.7%	1.7	0/1	0/1	9/-	9/-
Pres 2004	**Kerry**–Bush	37.2%	3.4	1/2	1/2	9/-	9/-
Sen 2006	**Kennedy**–Chase	30.5%	2.8	0/1	0/1	9/-	9/-
Pres 2008	**Obama**–McCain	36.8%	3.3	1/2	1/2	9/-	9/-
Sen 2008	**Kerry**–Beatty	31.9%	2.9	0/1	0/1	9/-	9/-
Sen 2010	Coakley–**Brown**	52.4%	4.7	9/-	9/-	8/9	8/9
Pres 2012	**Obama**–Romney	38.2%	3.4	3/4	3/4	9/-	9/-
Sen 2012	**Warren**–Brown	46.2%	4.2	7/9	7/8	9/-	9/-
Sen 2013	**Markey**–Gomez	44.8%	4.0	7/9	7/8	9/-	9/-
Sen 2014	**Markey**–Herr	38.0%	3.4	3/4	3/4	9/-	9/-
Pres 2016	**Clinton**–Trump	35.3%	3.2	2/3	3/4	9/-	9/-

Lower-variance elections (see previous table) are marked in gray. Election winners shown in boldface; R share is with respect to two-way vote; seat quotas are proportional share of nine seats.
*Libertarian vote share included with R in 2000 Senate race.

in the sense that the infeasibility bound is one more than the feasibility bound.[9]

Our procedure is simply to greedily create the largest R-majority collection possible from the chosen geographic units (in our case, towns or precincts) by including them in order of Republican margin per capita:

$$\delta/p = (\#\text{Rvotes} - \#\text{Dvotes})/(\text{census population of unit}).$$

The proof supporting this test of feasibility is shown in the Appendix.

As stated above, we will fix the number of districts at nine throughout the analysis, matching the congressional apportionment at the time of writing. This means that ideal district size is $I = 705{,}455$ for races before 2010 and $I = 727{,}514$ for 2010–16.

We can make several observations from the table. Moving to finer granularity of building blocks did not have any impact on the feasibility bounds for most elections. In two cases (Sen 2012 and Sen 2013), the precinct-level bounds are sharper: in both cases, our method applied to towns produces an inconclusive result about a grouping of size $8I$. With precincts, we find that the uncertainty is eliminated and a grouping of size $8I$ is shown to be impossible.[10] The 2016 presidential election is the only one for which the finer granularity has *shifted* the feasibility bounds. It is not possible to find (even scattered) towns totaling three districts' worth of population which collectively favor Trump over Clinton, but it becomes possible if precincts are the building blocks. So in that case, it is narrowly possible to achieve proportional representation for Trump voters; note, however, that this still falls far short of the *seven* Trump districts that the naive bounds would have predicted to be accessible by extreme gerrymandering, and this is even before the contiguity requirement is applied.

Numerical uniformity: The role of variance

In statistics, the *mean* of a set of numerical data records its average value, and the *variance* (or second central moment) tells you how spread out the values are around this mean. We claim that variance in the vote share of a minority group (here, Republicans) can be a primary explanatory factor for poor representational outcomes in districting. At one extreme, this is obvious: if the variance is zero, then the preferences in the state are completely uniform, and every single unit has the same 35% (say) of Republican votes. In this case, we can easily see that districting has no

[9]It is possible that the feasibility bound actually overstates the number of districts that can be built with a majority for the designated party—because the collection of size kI may not be splittable into k appropriate collections of size I—but any infeasibility bound reflects a mathematically proven impossibility, which drives all the conclusions in this article. See Appendix for more details.

[10]Of course, the impossibility for precincts implies the impossibility for towns—because towns are made up of precincts—even though the sorting method did not discover this. We handle the inconclusive cases for towns algorithmically in the Appendix.

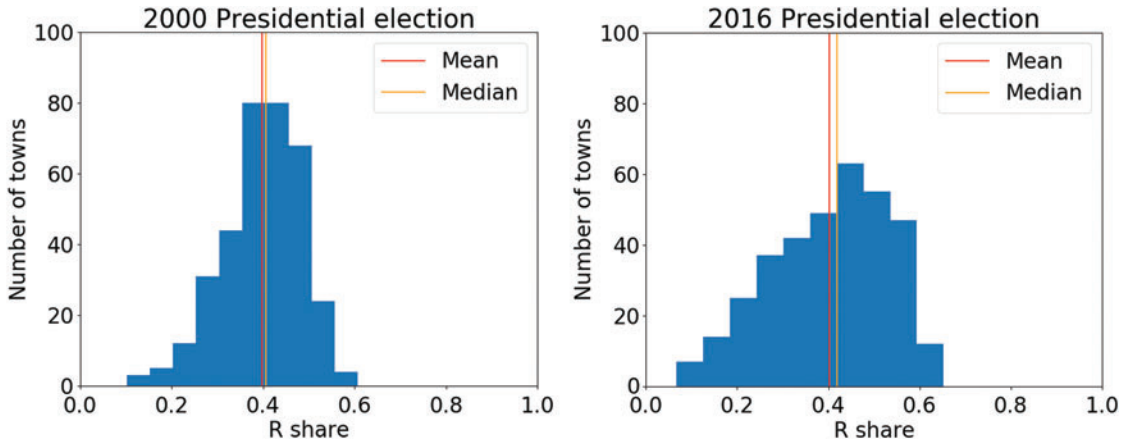

FIG. 1. These histograms show the distribution of Republican vote share by town in the 2000 and 2016 Massachusetts presidential contests, illustrating that these two elections had very nearly the same mean but different levels of variance. (The town-level variance is .0074 and .0165, respectively.) Color images are available online.

impact at all: every possible district will also have 35% R, and so will be won by Democrats.

Notably, the Gore/Bush election in 2000 had a two-way R vote share of 35.2% and results in zero possible R-majority districts. Meanwhile the Clinton/Trump election had a nearly identical 35.3% R vote share but produces the possibility for as many as *two* districts (built from towns) with a Trump majority.

The fundamental impact of variance is starkly illustrated in the histograms showing the actual vote patterns in Figure 1. A low-variance election with a minority of R votes may have very few units with R share over .5, which are precisely the building blocks needed to build an R-majority district.

Looking back to Table 2 corroborates this finding: 7 out of 13 elections exhibit a mathematical impossibility of representation or fall at least two seats short of proportionality—completely independent of the choices made by districters. These are precisely the seven elections in which the vote totals show lower variance, both at the town level and the precinct level. In five of the elections, this effect is so pronounced that the minority party is completely locked out of any possibility of representation.

Varying variance

To further probe these outcomes, we generated datasets with similar mean vote share to the 2000 and 2016 presidential elections (Figure 2), adjusting the variance of R-share per unit while maintaining voter turnout and population at actual levels.

We assigned R two-way vote shares chosen from a truncated skewed normal distribution with a set mean of 35.25% (the average of the Gore/Bush and Clinton/Trump R vote share) and variances ranging from 0.0020 to 0.0320, covering the range actually observed in Table 1.[11] From those datasets, we re-ran our procedure to produce bounds on the number of possible R seats.

The results, plotted in Figure 3, strongly corroborate the hypothesis that feasible representation is controlled by variance in vote share. In fact, a high enough variance can be seen to make it numerically feasible to overperform proportionality.

GEOMETRY OF REPUBLICAN UNDERPERFORMANCE

We now consider the spatial aspects of the vote distribution with respect to the possibilities for district formation.

[11]We used the scipy python library skewnorm.rvs function to generate random numbers from a skewed normal distribution with the chosen location, scale, and shape variable. *Truncation* means that any value outside of the [0,1] range was replaced by another value drawn from the same distribution. This truncation process changes the mean and variance of the distribution being produced, so we ran it iteratively, adjusting the mean and variance until the desired parameters were produced. Throughout, a shape variable of -8 was selected to best capture the observed distributions in historical elections. The resulting distributions can be seen in Figure 2.

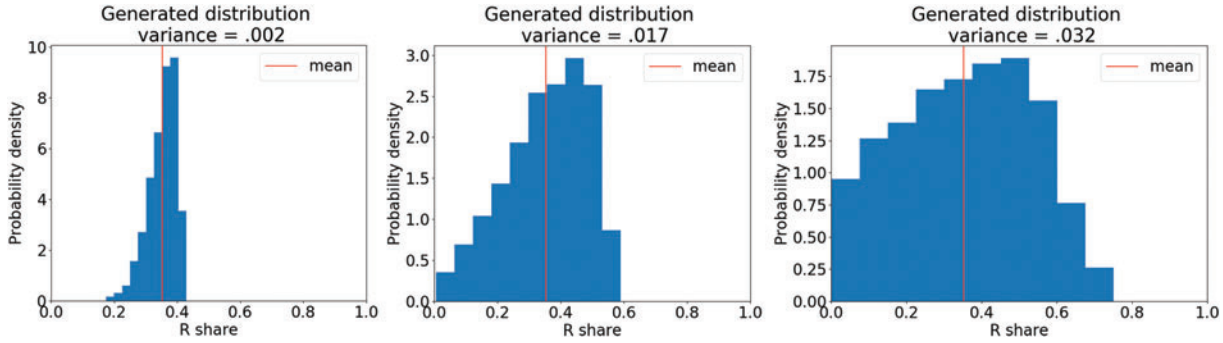

FIG. 2. Skewed truncated normal distributions are shown here with the same mean as the observed results. These were used to generate election data to test the hypothesis that vote datasets with higher variance would achieve higher levels of numerically feasible representation. Color images are available online.

Lack of Republican enclaves

Compounding the numerical effects described above is the spatial scatter of the areas preferring Republicans in Massachusetts. To illustrate this, consider forming a grouping of towns by collecting them in order of their R margin per capita δ / p, as above, until the collection is large enough to be a valid district. The result is a dramatically discontiguous assemblage spanning nearly the full state (Figure 4). A similar pattern can be observed in 2006 Senate returns.

In fact, very few of the building blocks seen in the figures are R-favoring at all. Strikingly, only an astonishing nine of 2,166 precincts in 2006 record a Chase majority.[12] Only 31 out of 351 towns had a G.W. Bush majority in 2000. The largest Bush-favoring collection of towns—which boasts an aggregate one-vote Bush margin—only has a population of 426,304, well short of the ideal district size of over 700,000.

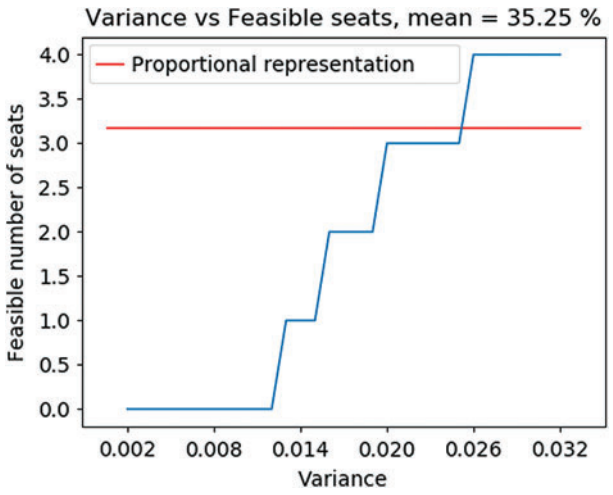

FIG. 3. Higher-variance datasets reliably produce greater numbers of feasible seats, even with the vote share held constant. This figure shows the results of three trials with the protocol described above; the results are indistinguishably close. Color images are available online.

Clustering

The voting data used here makes it possible to test whether, in addition to increased variance, the election results after 2010 exhibit more spatial clustering than before. To assess this we use an index called a capy (or clustering propensity) score, which closely resembles well-established assortativity scores in network science, generalized for use with demographic data.[13]

The geographical units that make up a jurisdiction have populations of different sizes and compositions. In geographical unit v_i, we use x_i and y_i to denote the populations from group X and group Y in that unit. We record the X population data as an integer-valued vector $\mathbf{x} = (x_i, \ldots, x_n)$ with entries for each unit's population, and likewise write $\mathbf{y}$ for the Y population figures. If unit v_i is adjacent to unit v_j, we write $i{:}\, j$. Then

[12]In fact, the real number is almost certainly eight precincts. The official results record a large Chase majority in Medford Ward 5 Precinct 2, which is not consistent with the voting behavior of that precinct in any other election on record, including the primary that year. Medford town officials were unable to provide corrected data.

[13]For details, see Alvarez et al. (2018) for a comparative mathematical survey of scores of clustering and segregation, and Newman (2003) for a survey of network science that includes assortativity.

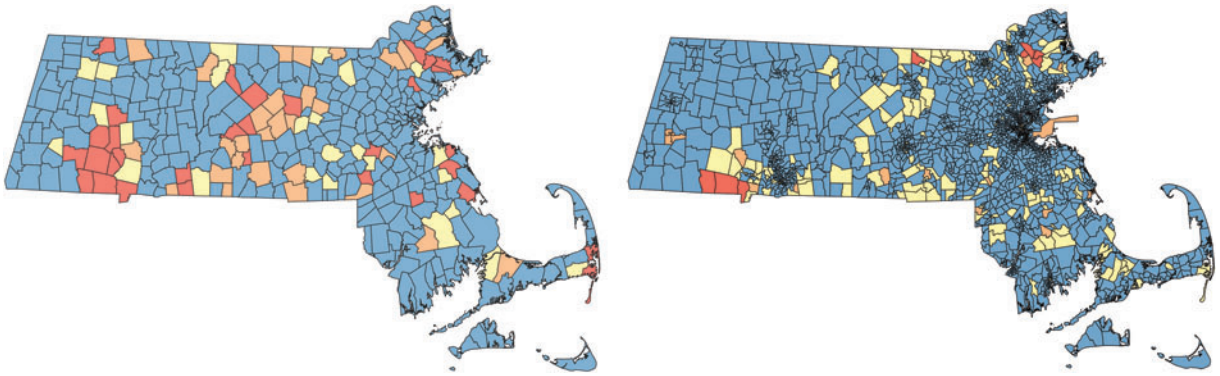

FIG. 4. These figures show the voting pattern for Republicans George W. Bush in the 2000 presidential race (*left*, by town) and Kenneth Chase in the 2006 senate race (*right*, by precinct). The *darkest red* units favored the Republican outright, and the *lighter red* shade shows the most Republican-favorable units available in assembling enough population for a congressional district. These quasi-districts still preferred Gore and Kennedy, respectively, by comfortable margins. Color images are available online.

let $\langle \mathbf{x}, \mathbf{y} \rangle := \sum_i x_i y_i + \sum_{i \sim j} x_i y_j + x_j y_i$. The idea is that $\langle \mathbf{x}, \mathbf{y} \rangle$ is a close approximation to the number of individuals of X type living next to an individual of Y type, either in the same geographical unit or in neighboring units.[14] With this, we define

$$H(\mathbf{x}, \mathbf{y}) := \frac{1}{2}\left(\frac{\langle \mathbf{x}, \mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle} + \frac{\langle \mathbf{y}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle}\right).$$

By construction, this score varies from 0 to 1 and measures the tendency of each of the two kinds of population to live next to another member of their own group, rather than the other. A perfectly uniform distribution where the x_i and the y_i were constant would earn the score $H = 1/2$, and a perfectly clustered distribution where the $x_i = 0$ in one region and the $y_i = 0$ in the complementary region would tend towards $H = 1$ in a sufficiently large network.

Table 3 shows the observed $H(R,D)$ clustering results for Republican compared to Democratic voters. For each election, we create two comparison points by experiment: the *uniform H* score is the highest score recorded in 30 trials in which Republican voters were scattered randomly under a uniform distribution until reaching the statewide R share observed in that election. The *clustered H* score is produced by applying a dynamical step that moves votes into a configuration with higher tendency for neighbors to have the same vote.[15] As a general matter, we see that the H scores from actual election data closely resemble the uniform trials, and that there is only a mild upward trend in the H scores over time. In some cases, there are interesting comparisons, such as in comparing the presidential outcomes in 2000 and 2016—there, we can see that Trump voters are appreciably more likely to live next to each other than Bush voters were, but still far from highly clustered.

There is a one-way relationship between numerical and geometric uniformity: if there is low

TABLE 3. CLUSTERING SCORES FOR REPUBLICAN VERSUS DEMOCRATIC VOTERS AT THE TOWN LEVEL IN EACH OF THE ELECTIONS DISCUSSED IN THIS ARTICLE

Election	*R share*	*Uniform H*	*Observed H*	*Clustered H*
Pres 2000	35.2%	.5001	.5135	.9456
Sen 2000	25.4%*	.5000	.5063	.9374
Sen 2002	18.7%	.5001	.5035	.8982
Pres 2004	37.3%	.5000	.5182	.9351
Sen 2006	30.6%	.5001	.5171	.9537
Pres 2008	36.8%	.5000	.5210	.9591
Sen 2008	32.0%	.5000	.5181	.9513
Sen 2010	52.4%	.5001	.5329	.9587
Pres 2012	38.2%	.5000	.5243	.9268
Sen 2012	46.2%	.5000	.5272	.9597
Sen 2013	44.9%	.5002	.5366	.9492
Sen 2014	38.0%	.5001	.5276	.9557
Pres 2016	35.3%	.5000	.5344	.9480

We show the score for a uniform trial, the actual observed votes, and a highly clustered trial, each with the statewide share that corresponds accurately to the given election. The numbers are truncated (not rounded) after four decimal places.
*Libertarian vote share included with R in 2000 Senate race.

[14]This approximation approaches equality as the populations get large. For details, see Alvarez et al. (2018).
[15]This belongs to an extremely standard toolkit from physics (cf. Glauber dynamics in the Ising model); replication code can be found in our GitHub repo, Metric Geometry and Gerrymandering Group (2019).

variance in observed partisan shares by unit, then all units tend to have the same shares, so there is necessarily no major spatial pattern to partisan preference. However, high variance in partisan share can occur in a way that is strongly spatially clustered (such as if there are pronounced enclaves) or in a way that is not (such as if there is a checkerboard pattern of strong support for each party). The findings here strongly support a conclusion that numerically uniform vote patterns create obstructions to representation for a group in the numerical minority. Further work is needed to study the spatial determinants of representability in the high-variance case.

CONCLUSION

The numerical and geometric/spatial distribution of voter preferences, and the local rules of redistricting, restrict and skew the possibilities for representation in an extremely complex way that one-size-fits-all normative ideals fail to capture. There has been significant recent progress attacking the mathematical challenges of identifying the representational baseline. New tools, such as the ones presented here, make it increasingly possible to separate the effects of choosing district boundaries from the consequences of political geography. A strong message is emerging: the range of possible representational outcomes under valid redistricting is not always in keeping with the range that classic modeling approaches have predicted from the vote share alone. Any meaningful claim of gerrymandering must be demonstrated against the backdrop of valid alternative districting plans, under the constraints of law, physical geography, and political geography that are actually present in a jurisdiction.

REFERENCES

Alvarez, E., M. Duchin, E. Meike, and M. Mueller. 2018. "Clustering Propensity: A Mathematical Framework for Measuring Segregation," preprint. <https://mggg.org/Capy.pdf>.

Ballotpedia. 2016. "U.S. House of Representatives Elections in Massachusetts, 2016." *Ballotpedia.* <https://ballotpedia.org/United_States_House_of_Representatives_elections_in_Massachusetts,_2016>.

Calvo, E., and J. Rodden. 2015. "The Achilles Heel of Plurality Systems: Geography and Representation in Multiparty Democracies." *American Journal of Political Science*, 59(4): 789–805.

Chen, J., and J. Rodden. 2013. "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures." *Quarterly Journal of Political Science*, 8: 239–269.

Chen, J., and J. Rodden. 2016. "The Loser's Bonus: Political Geography and Minority Party Representation," preprint. <http://www.jonathanrodden.com/s/losers_bonus_nov2016.pdf>.

DeFord, D., and M. Duchin. 2019. "Redistricting Reform in Virginia: Districting Criteria in Context." *Virginia Policy Review*, 12(2): 120–146.

Duchin, M. 2018. *Outlier Analysis for Pennsylvania*, LWV vs. Commonwealth of Pennsylvania, Docket No. 159 MM 2017 (February 15).

Grofman, Bernard, William Koetzle, and Thomas Brunell, 1997. "An Integrated Perspective on the Three Potential Sources of Partisan Bias: Malapportionment, Turnout Differences, and the Geographic Distribution of Party Vote Shares." *Electoral Studies*, 16(4): 457–470.

Gudgin, G., and P.J. Taylor. 2012. *Seats, Votes, and the Spatial Organisation of Elections*. With a new introduction by R.J. Johnston. ECPR Press (original publication: Pion, 1979).

Karp, Richard M. 1972. "Reducibility Among Combinatorial Problems." *Complexity of Computer Computations* (Proc. Sympos., IBM Thomas J. Watson Res. Center), 85–103.

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

King, Gary, and Robert X. Browning. 1987. "Democratic Representation and Partisan Bias in Congressional Elections." *American Political Science Review*, 81(4): 1251–1273.

"Knapsack Problem—Definition," 2019. *Wikipedia*, <https://en.wikipedia.org/wiki/Knapsack_problem#Definition>.

Levitt, Justin. 2019a. "Massachusetts." *All About Redistricting*. <http://redistricting.lls.edu/states-MA.php>.

Levitt, Justin. 2019b. "Where the Lines Are Drawn—Congressional Districts." *All About Redistricting*. <http://redistricting.lls.edu/where-tablefed.php>.

Massachusetts Constitution. 2019. <https://malegislature.gov/laws/constitution>.

Massachusetts Secretary of State. 2019. *Massachusetts Election Statistics*. <http://electionstats.state.ma.us>.

MassGIS. 2019. "Community Boundaries (Towns) from Survey Points." *MassDocs Pilot*. <https://docs.digital.mass.gov/dataset/massgis-data-community-boundaries-towns-survey-points>.

Metric Geometry and Gerrymandering Group. 2018. "PartyTilt." *GitHub*. <https://github.com/gerrymandr/party-tilt>.

Metric Geometry and Gerrymandering Group. 2019a. *GerryChain MCMC Python Package*. <https://github.com/mggg/GerryChain>.

Metric Geometry and Gerrymandering Group. 2019b. *Massachusetts Election Data Repository*. <https://github.com/gerrymandr/Massachusetts_underperformance>.

Metric Geometry and Gerrymandering Group. 2019c. *Preprocessing Tools*. <https://github.com/gerrymandr/Preprocessing>.

Newman, M.E.J. 2003. "The Structure and Function of Complex Networks." *SIAM Rev.*, 45(2): 167–256.

Pegden, W. 2017. *Pennsylvania's Congressional Districting Is an Outlier*. Expert report in LWV vs. Commonwealth of Pennsylvania. (November).

Rae, Douglas W. 1967. *The Political Consequences of Electoral Laws*. New Haven, CT: Yale University Press.

Taagepera, Rein. 1973. "Seats and Votes: A Generalization of the Cube Law of Elections." *Social Science Research*, 2: 257–275.

Taagepera, Rein. 1989. "Empirical Threshold of Representation." *Electoral Studies*, 8(2): 105–116.

Tufte, Edward R. 1973. "The Relationship Between Seats and Votes in Two-Party Systems." *American Political Science Review*, 67(2): 540–554.

U.S. Census Bureau. 2016. *TIGER/Line Shapefiles*. <https://catalog.data.gov/dataset/tiger-line-shapefile-2012-2010-state-massachusetts-2010-census-voting-district-state-based-vtd>.

Address correspondence to:
Moon Duchin
Tufts University
503 Boston Avenue
Medford, MA 02155

E-mail: Moon.Duchin@tufts.edu

Received for publication November 17, 2018; received in revised form September 14, 2019; accepted September 30, 2019; published online December 13, 2019.

Appendix: Rigorous Feasibility Bounds

Suppose you have a list of units with corresponding populations p_i and R margins $\delta_i = r_i - d_i$, the number of R votes minus the number of D votes. Re-index so that they are ordered from greatest to least by margin per capita:

$$\delta_1/p_1 \geq \delta_2/p_2 \geq \ldots \geq \delta_n/p_n$$

We will call a collection of units S a *grouping*, and let $p(S)$ and $\delta(S)$ be its population and R margin, found by summing the p_i and δ_i for its units. Let D_k be the grouping indexed by $\{1, \ldots, k\}$. Let K be the smallest integer k for which $\delta(D_k) \leq 0$. This means that D_{K-1} has a collective R majority, but if you add the Kth unit you get a grouping D_K that fails to have an R majority.

Theorem 1. *With the notation above, let M be any positive integer.*

Case 1. $M \leq p(D_{K-1})$. There *exists* an R-majority grouping of size at least M.

Case 2. $p(D_{K-1}) < M \leq p(D_K)$. Inconclusive: such a grouping may or may not exist.

Case 3. $p(D_K) < M$. There *does not exist* an R-majority grouping of size at least M.

Proof. In Case 1, it is clear that a Republican grouping can be created, because D_{K-1} is a Republican-majority grouping of sufficient size.

We present examples to illustrate that Case 2 is inconclusive.

i	r_i	d_i	p_i	δ/p_i
1	8	0	8	1
2	1	9	10	$-4/5$
3	0	5	5	-1

i	r_i	d_i	p_i	δ/p_i
1	8	0	8	1
2	1	9	10	$-4/5$
3	0	8	8	-1

For both examples, fix $M = 13$. We have $K = 2$ in both examples because $\delta(D_1) = 8 > 0$ and $\delta(D_2) = 0$. Both fall under Case 2 because $p(D_1) = 8$ and $p(D_2) = 18$, while $M = 13$. In the left-hand example there exists an R-majority grouping, made by putting together units 1 and 3 to form a grouping with $\delta = 3$ and population 13. But in the right-hand example there is none, which is easily confirmed by considering all of the combinations.

Finally, in Case 3, we have $p(D_K) < M$.

Claim. *Let $S = D_K$ and suppose that $p(S) < M$. Then for any $S' \subseteq \{1, \ldots, n\}$,*

$$p(S') > p(S) \Rightarrow \delta(S') < \delta(S).$$

The claim asserts that D_K has the optimal R margin among all groupings with at least as much population. Since we seek a grouping larger than $p(D_K)$ and since $\delta(D_K) \leq 0$, this implies that a R-majority grouping cannot be formed. So it just remains to prove the claim.

Let $A = S' \setminus S$ and $R = S \setminus S'$ denote the sets of indices added to and removed from S, respectively, to make S'. Since A and R are disjoint, and we have assumed that $p(S') > p(S)$, it follows that $p(A) > p(R)$. Let $\mu = \max\{\frac{\delta_i}{p_i} | i \in A\}$ and let $\mu\prime = \min\{\frac{\delta_i}{p_i} | i \in R\}$. Note that, since $R \subseteq S = \{1, \ldots, K\}$ and $A \subseteq S^c = \{K+1, \ldots, n\}$ and the $\frac{\delta_i}{p_i}$ are non-increasing, we have $\mu \leq \mu'$.

Note that every unit $i \notin S$ has a Democratic majority ($\delta_i < 0$). This is because Republican-majority units are added to S in decreasing order of $\frac{\delta_i}{p_i}$ until the overall margin satisfies $\delta \leq 0$, so by construction every unit with a Republican majority is in S. It follows, since $A \subseteq S^c$, that $\mu < 0$.

We have $\mu \cdot p(R) > \mu \cdot p(A)$ because $p(R) < p(A)$ and $\mu < 0$. Also, $\mu' \cdot p(R) \geq \mu \cdot p(R)$. So, transitively, $\mu' \cdot p(R) > \mu \cdot p(A)$.

Note that

$$\mu' \cdot p(R) = \sum_{i \in R} \mu' \cdot p_i \leq \sum_{i \in R} \frac{\delta_i}{p_i} \cdot p_i = \delta(R).$$

Similarly $\mu \cdot p(A) \geq \delta(A)$. Combining our inequalities, we have shown that $\delta(R) > \delta(A)$. It follows that $\delta(S) > \delta(S')$, as claimed. This completes the proof of the claim and the theorem.

Note that Case 2, the inconclusive situation, is more likely when there are units that are large relative to the population threshold, because the gap between $p(D_{K-1})$ and $p(D_K)$ is the population of the Kth unit. So if we consider the formation of districts, we are more likely to get an inconclusive result with large units like counties or towns and less likely with smaller units like blocks or VTDs/precincts.

This theorem suggests an algorithm for computing feasibility bounds that is no more complex than sorting, which makes it fast and efficient. The answers are not completely satisfying, however, because of the possibility of an inconclusive finding (Case 2) and because the existence of a grouping with an R majority and population that is m times the size of an ideal district does not imply that it can be split into m sub-groupings of equal size, each with R majorities. However, a refined algorithm that could close those loopholes is known to have forbidding computational complexity, because eliminating the inconclusive case is equivalent to the 0 – 1 *knapsack problem*, which is NP-complete.[A1] Sorting into m collections while tracking both weight and value, which would close the second loophole, is strictly harder.

Some problems may be solvable in reasonable time even when we lack an algorithm with a polynomial bound. We implemented a pseudo-polynomial dynamic programming knapsack algorithm, which ascertained in under a minute that the correct feasibility/infeasibility bounds in Senate 2012 and Senate 2013 were 7/8, removing the ambiguity left by the simple sorting algorithm in Table 2. However, we were unable to find or quickly devise a variant for sub-sorting in the six elections where multiple R districts are numerically feasible.

[A1] "Knapsack Problem—Definition," *Wikipedia* (2019), <https://en.wikipedia.org/wiki/Knapsack_problem#Definition>. For a formal reference, see, for instance, the classic Karp (1972).