Voting and Elections: New Social Science Perspectives

Richard Holden

University of New South Wales Business School, Sydney, New South Wales 2033, Australia; email: richard.holden@unsw.edu.au

Abstract

This article discusses recent developments in the study of voting and elections. How people end up voting in an election depends on (a) how effective voting power is distributed among voters and (b) the strategic interactions between voters and other interested parties. These are, in turn, affected by institutional arrangements, such as the composition of voting districts, campaign finance laws, and constitutional restrictions on vote dilution. In recent years, new social science–based approaches, both theoretical and empirical, from economists, political scientists, and legal scholars have shed new light on the democratic process.

Keywords

campaigns, gerrymandering, incumbent, political geography, redistricting, registration, turnout
INTRODUCTION

At the very core of the democratic process is the way in which voters choose their elected officials. It is no surprise, then, that the properties of these processes have long been of great interest. Dating at least to eighteenth-century French mathematician Condorcet, who studied voting rules, social scientists have understood that the ways electoral institutions are designed have important effects on the incentives, and hence behavior, of individuals. This in turn affects the efficacy of electoral institutions.

Legal rules are particularly important for both the process and outcome of elections. How people end up voting in an election depends on (a) how effective voting power is distributed among voters and (b) the strategic interactions between voters and other interested parties. These are, in turn, affected by institutional arrangements, such as the composition of voting districts, campaign finance laws, and constitutional restrictions on vote dilution. Legal rules are fundamental to these two, broad factors.

The United States is a compelling laboratory for discussing these issues—although many of the general principles apply in other jurisdictions. In the United States, constitutional protections, especially those contained in the Bill of Rights, have at least the potential to shape both points above.

The free speech guarantees of the First Amendment have long had an important impact on campaign finance laws. Since the early 1970s there has been a dialogue between Congress and the US Supreme Court about limits on political contributions. Beginning with the Federal Election Campaign Act of 1971 and the 1974 reforms to it, there have been challenges to the constitutionality of such regulations. In Buckley v. Valeo, the Court struck down limits of expenditures by candidates themselves (applying strict scrutiny), but upheld limits on contributions to candidates (applying rational basis review). McConnell v. Federal Election Commission essentially upheld the Bipartisan Campaign Reform Act of 2002 (often referred to as McCain-Feingold). However, more recently, Citizens United v. Federal Election Commission found that the government is prohibited, on First Amendment grounds, from restricting independent campaign expenditures by not-for-profits, corporations, and labor unions. The Equal Protection Clause of the Fourteenth Amendment provides,

No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws. (Clause 1)

This has opened the door to a string of cases addressing whether certain redistricting, voter registration, and other electoral practices are constitutional.

The landmark decision in Baker v. Carr—a case that involved vote dilution due to congressional districts not being redrawn and population growth leading to disparities as large as 3 to 1—declared that redistricting claims were justiciable. The Court went further in Wesberry v. Sanders, holding that only congressional voting districts with populations “as nearly equal as possible” were consistent with the Equal Protection Clause. The court applied a similar standard to districts for statewide legislative bodies in Reynolds v. Sims and local governments in Avery v. Midland County.

Statutes have also had an important impact on elections, with the Voting Rights Act (VRA) (1965) being the quintessential example. From the perspective of elections and redistricting, there are two key sections of the VRA: sections 2 and 5. Section 2 prohibits states from using any voting practice “in a manner which results in a denial or abridgement of” minority voting rights [42 U.S.C. §2 (1973)]. Roughly speaking, this has been interpreted over time to mean that districting plans that dilute the interests of minority voters are prohibited (see, e.g., Thornburg v. Gingles).
What “dilute” means is a complex and contested concept, but it is fair to say that in some instances this has led to a requirement to create so-called majority-minority districts—i.e., districts in which a racial minority constitutes a majority of the voters.¹

Section 5—known as the reclearance provisions—requires that certain jurisdictions submit changes to districting plans to the Department of Justice (or the D.C. Circuit) for approval. The jurisdictions to which this provision applies are defined in the coverage formula of section 4(b), which was updated by Congress in 1970 and 1975. It was targeted at jurisdictions that had engaged in egregious vote-dilution practices. Section 5 was effectively invalidated in a relatively recent decision of the US Supreme Court in *Shelby County v. Holder*. That decision ruled the coverage formula unconstitutional for not being responsive to current political conditions.

Social science–based understandings of voting and elections play an important role in many of these cases, sometimes in the background, and often in the foreground. A threshold question is whether, for example, claims that political districting plans violate the Fourteenth Amendment are justiciable. This question dates to at least *Colegrove v. Green*, in which Justice Frankfurter, writing for the 4–3 plurality, famously observed, “To sustain this action would cut very deep into the very being of Congress. Courts ought not to enter this political thicket” (emphasis added). In *Davis v. Bandemer*, the Court found that partisan gerrymandering claims were justiciable, but did not offer a clear standard for assessing such claims. The plurality (4–1–4) decision in *Vieth v. Jubelirer* did not overturn *Bandemer*, but the plurality opinion did hold that the absence of a judicially manageable standard meant that partisan gerrymandering claims were not justiciable. The closeness of this case, and the change (and future change) of the composition of the Court, however, mean that this could be an important question going forward. As I discuss in the conclusion of this article, the very notion of a judicially manageable standard for partisan gerrymander claims puts social science at the heart of legal matters, for it is social science techniques and analysis that are essential components of such a standard.

Voting and elections—and particularly redistricting—are areas where the interplay between social science and the law is extremely rich and developments on both sides have been important. In recent years there have been developments both on the theoretical side and on the empirical side, from economists, political scientists, and legal scholars, that have enriched our understanding of voting and elections. The purpose of this article is to highlight some of the work that has been associated with those developments. I focus relatively heavily on issues surrounding redistricting and political geography—as this is the area where new social science techniques have had the most pronounced impact to date—but I also discuss voter turnout and how modern campaigns are conducted and briefly comment on some of the issues these raise for election law, especially in the United States.

Theoretical models in the social sciences are always wrong, in the sense that they are abstractions designed to highlight one or two important forces rather than to be a complete description of reality, like models in particle physics. As the great economist Joan Robinson (1962) famously said, “A model which took account of all the variegation of reality would be of no more use than a map at the scale of one to one.” Having said that, models that are highly sensitive to assumptions that are made for convenience are undesirable because their conclusions are not robust to potentially small deviations from those ad hoc assumptions. Moreover, models that are not microfounded in the sense that they take individual voting decisions as the primitive object of the

¹For an outstanding analytic treatment of majority-minority districting, see Shotts (2002). Shotts develops a model to assess the claim that majority-minority districting can lead to the somewhat perverse outcome that more conservative legislators are elected, thereby shifting policy outcomes to the right, despite the more left-leaning political preference of minority voters.
models of elections and, especially, redistricting have moved in a more robust and microfounded direction.

On the empirical side, there has been a massive increase in the amount, and richness, of publicly available data. At the same time, there has been a much greater emphasis on trying to identify the true causal effect of various phenomena, rather than being satisfied with mere correlations. This, in many ways, parallels the identification revolution that began in labor economics and has now spread across the social sciences.\(^2\)

Combining both theory and empirics, there has been an increased use of computer science techniques in the study of elections—especially redistricting. In fact, many electoral problems are inherently combinatoric in nature. One prominent example is constructing districting plans, where enumerating the set of all districts (even subject to the constraint that districts contain equal numbers of voters) is NP-hard for even a modest-size jurisdiction. Or, to put it differently, the number of feasible congressional districts for the state of California is (much) larger than the number of atoms in the observable universe.

The remainder of the article proceeds as follows. The next section deals with redistricting, including partisan gerrymandering, incumbency advantages, and political geography. The third section discusses issues around voter registration and turnout.\(^3\) The fourth draws together some of the implications of the previous sections for election law. The final section contains some concluding remarks, but perhaps most importantly highlights five broad and important open questions in the field that I hope will serve as partial impetus for future research.

Two final points are worth emphasizing. First, although this article, like much of the literature, is somewhat US-centric, several of the issues are germane to elections in other jurisdictions. Second, I have been deliberately selective in the papers I have cited here, and this is not meant to be an encyclopedic exercise, nor have I sought to compile an exhaustive list of references. I hope, however, that it will serve as a useful summary and entry point into the recent literature.

**REDISTRICTING**

**Partisan Gerrymandering**

Because of uneven population growth, it is commonplace in all democracies for electoral boundaries to be redrawn from time to time. There is significant heterogeneity in how that process is carried out, with the United States being perhaps the starkest illustration of politicians themselves being in control of the redistricting process.

Allowing partisan political actors such power predictably leads to abuse of that power. The term gerrymander dates to 1812, when Governor of Massachusetts (and later Vice President of the United States) Elbridge Gerry signed a reapportionment bill that created a district in Essex County that was so oddly shaped it was said to resemble a salamander. Political cartoons of the day depicted it as such, giving rise to the portmanteau gerrymander.

From a social scientist’s perspective, understanding the optimal strategy for a gerrymanderer is important for at least two reasons. First, it allows one to understand how large the advantages

---

\(^2\) See Angrist & Pischke (2009) for an elegant account of the tools and techniques that permit causal inference even when randomized controlled trials are not possible.

\(^3\) This is one section that is particularly US-centric because many of the issues covered in it do not arise in jurisdictions with compulsory voting and quasi-automatic voter registration.
of drawing electoral boundaries are, and hence the resources that will be optimally expended in pursuit of that advantage. Second, if one hopes to regulate the practice one must consider the strategic interactions between regulator and regulated.

A series of important papers consider properties of districting plans as a whole by using an analytic framework known at the seats-votes curve. The seats-votes curve is a mapping from the proportion of the vote won to the proportion of seats won and is of particular interest in two-party elections. The two key parameters are the bias (the difference between the proportion of seats won with 50% of the votes and one-half) and the responsiveness (the slope of the curve at one-half).

For instance, Owen & Grofman (1988), Sherstyuk (1998), Gilligan & Matsusaka (1999), and Cox & Katz (2002) all consider the trade-off between biasedness and responsiveness for the gerrymander. These analyses, in no small part, gave rise to the now long-standing intuition for the optimal strategy for a gerrymander: the so-called pack-and-crack approach. Under this strategy, the gerrymanderer concentrates her opponents into several unwinnable districts (packing) and spreads the remaining opponents as well as her now relatively more numerous supporters over the remaining districts (cracking).

Although elegant, a drawback of the approach that these papers take is that it is not microfounded, in the sense that it analyzes properties of state-wide districting plans, rather than analyzing the placement of individual voters (or blocks of voters) into particular districts. There is no guarantee that there exists a feasible allocation of individual voters satisfying the constraints that districts be contiguous and contain an equal number of voters possessing the aggregate properties that optimize the trade-off between biasedness and responsiveness.

Gilligan & Matsusaka (1999), by contrast, do analyze a microfounded model of gerrymandering. In their model, the gerrymanderer observes the voting intention—via party affiliation—of each voter perfectly. That is, he knows for sure which party each voter intends to vote for. This leads him to the conclusion that the optimal gerrymandering strategy is to create as many districts with a bare majority of supporters in them as possible, because such districts are won by the gerrymanderer with certainty. The only limitation on this is the proportion of one’s supporters in the population. Indeed, with a bare majority in the population, the gerrymanderer wins all districts.

Despite its intuitive appeal, Friedman & Holden (2008) showed formally that this is, in fact, not the optimal strategy in general. The intuitions for pack-and-crack come from theoretical models with special assumptions—such as there being only two types of voters—that do not generalize. Friedman & Holden showed, in a model with a continuum of voter types, that instead a strategy of matching one’s most ardent supporters with a slightly smaller number of one’s most ardent opponents and continuing this process into the center of the signal distribution is, in fact, optimal in general. The following table from their paper provides a numerical example of the superiority of this matching slices strategy to the pack-and-crack approach.

Table 1 shows how to construct the optimal gerrymander for a hypothetical state with five districts and assumes (without loss of generality) that the redistricter is the right-wing (e.g., Republican in the US setting) party. Voters come in a continuum (an infinite number) of types ranging from the far left to the far right, and these preferences are drawn from a normal distribution. There is an aggregate shock to voter preferences so that although the redistricter receives a signal of voter preferences, she is not certain of the eventual voting outcome (in the aggregate). Panel a shows the relative mass of the upper (i.e., from the right of the distribution) slice of the district to the lower (i.e., from the left of the distribution) slice of the district. District 1 is composed of 62% from the right tail and 38% from the left tail, and the redistricter has an 87.5% chance of winning that district. There are two other salient characteristics of the optimal matching slices strategy. The first is that the relative size of the upper-to-lower slice gets larger as the districts become less safe. This is because signals in the far right or far left tails of the distribution are more precise.
Table 1  Matching slices gerrymandering\(^a\)

<table>
<thead>
<tr>
<th>a. Baseline example</th>
<th>District</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Upper slice</td>
<td>0.62</td>
</tr>
<tr>
<td>Lower slice</td>
<td>0.38</td>
</tr>
<tr>
<td>Prob. (win)</td>
<td>87.5%</td>
</tr>
</tbody>
</table>

b. Signal coarseness

<table>
<thead>
<tr>
<th>Signal variance</th>
<th>E (districts won)</th>
<th>Probability of winning district</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0.50</td>
<td>3.46</td>
<td>97.4%</td>
</tr>
<tr>
<td>2.50</td>
<td>2.83</td>
<td>87.5%</td>
</tr>
<tr>
<td>4.50</td>
<td>2.53</td>
<td>68.2%</td>
</tr>
</tbody>
</table>

c. Spread of voter preferences

<table>
<thead>
<tr>
<th>Preference variance</th>
<th>E (districts won)</th>
<th>Probability of winning district</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3.0</td>
<td>2.55</td>
<td>71.0%</td>
</tr>
<tr>
<td>5.0</td>
<td>2.83</td>
<td>87.5%</td>
</tr>
<tr>
<td>25.0</td>
<td>3.78</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

d. Partisan bias of the population

<table>
<thead>
<tr>
<th>% Republican</th>
<th>E (won)</th>
<th>Value</th>
<th>Probability of winning district</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>30%</td>
<td>2.04</td>
<td>0.58</td>
<td>49.4%</td>
</tr>
<tr>
<td>40%</td>
<td>2.44</td>
<td>0.48</td>
<td>87.0%</td>
</tr>
<tr>
<td>50%</td>
<td>2.83</td>
<td>0.33</td>
<td>87.5%</td>
</tr>
<tr>
<td>60%</td>
<td>3.24</td>
<td>0.20</td>
<td>88.7%</td>
</tr>
<tr>
<td>70%</td>
<td>3.67</td>
<td>0.12</td>
<td>90.2%</td>
</tr>
</tbody>
</table>


than those in the middle.\(^4\) In a sense, the gerrymanderer finds it optimal to cut districts less finely because she is less certain about how voters will actually end up voting. The second characteristic of the optimal gerrymander is that the probability of winning districts gets lower as voters are assigned to districts more from the middle of the distribution, rather than from the tails. This is a key point to which I return shortly.

To see why pack-and-crack is dominated by matching slices, consider the cracked districts under the former strategy. Those are composed of an identical array of voters (or as nearly identical as is practicable), including those most likely to vote for the gerrymanderer. For a right-wing party those voters are used as right-of-the-median voters in multiple districts, rather than being used as the median (and hence pivotal) voters in some district. The same logic obviously applies for a left-wing party.

\(^4\)This is easy to see for the normal distribution by simply calculating the likelihood ratio and observing that it goes to positive (respectively negative) infinity as one goes to the far right (respectively left) tail of the distribution. This fact, however, is much more general and extends beyond the normal distribution.
Table 1 also highlights various comparative static properties of the optimal gerrymander. Panel b shows how the optimal gerrymander changes as the signal gets more or less precise. The middle row, with a signal variance of 2.50, is identical to the baseline example in panel a. The top row, with a signal variance of 0.50, shows that the redistricter wins more districts in expectation (3.46 compared with 2.83). It is also worth noting that one can compute the best pack-and-crack strategy for exactly this example. With a signal variance of 0.5 under pack-and-crack, the redistricter wins 2.86 districts in expectation, a full 0.6 districts less than the 3.46 under matching slices.

Friedman & Holden (2008) show that she also cuts districts more finely (has a small relative mass of upper-to-lower tail voters) when signals are more precise. Much has been written in the popular press and elsewhere (Issenberg 2012) about the sophistication of modern US presidential campaigns in terms of gathering and analyzing information about voters preferences—which corresponds to obtaining a more precise signal. Not only does this lead to the redistricter doing better in terms of the expected number of seats won, it causes her to change her optimal strategy. It is this latter point that is arguably of most interest: The voters who win and lose from gerrymandering are determined, at least in part, by the informational environment. This has significant policy implications, as I discuss in my concluding remarks.

The final comparative static property of interest is how the value of gerrymandering is affected by the general left-right leaning of the population. Panel d compares the expected number of seats won with a proportional share of the population to the expected number of seats won under the optimal gerrymander. The difference between these two numbers is the value of the gerrymander. Notice that this value is the greatest when the gerrymanderer has only 30% support in the population (the value is 0.58 expected seats), and it falls (monotonically) to 0.12 when the gerrymanderer has 70% support in the population. This fact demonstrates that being the redistricter is more valuable when one is in the minority. This has implications for competition to become the redistricter.

Finally, this gerrymandering value—the difference between the expected number of districts won and the number of expected districts won under proportional representation—is a useful and easy way to calculate a measure of the potential for mischief when it comes to drawing electoral boundaries. In the examples above, it is a function of the informational environment, as it would be in practice. However, one could also easily incorporate geographical and other constraints. Indeed, the reduction in the gerrymandering value from the imposition of a constraint (such as contiguity, communities of interest, compactness, or minority representation) is a good measure of how meaningful such constraints are. And it puts a numerical value on them in natural and useful units—i.e., the expected number of districts won.1

Incumbency Advantages

A striking fact about US congressional elections—although the phenomenon is far more widespread than that—is the very high reelection rates for incumbent representatives. Figure 1 (Friedman & Holden 2009) shows the time series of that reelection rate from 1898 to the early 2000s. Although there are dips in certain years, the rate is very high: In 2004, 97.9% of those who ran won (see the top line and left axis). Moreover, there was an upward trend over the century, indeed really in the post–World War II era.

Many have seen this as a worrying trend, even leading *The Economist* (2004) to compare the current state of democracy in America to that in North Korea. There are many reasons to be

---

1I am grateful to George Akerlof for suggesting this interpretation to me in a seminar I gave at UC Berkeley in 2006.
worried about elected officials becoming too entrenched: corruption, the ability to tilt elections in their favor even if they are not the best candidate, and effectively foreclosing entry of new candidates, among other things. Of course, there is a countervailing effect that learning on the job may lead elected officials to become more productive over time, all else equal.

Whatever the balance of these competing effects, it is natural to ask why the reelection rate is so high, and what has caused the upward trend over time. The literature has offered many alternative explanations, ranging from the benign to the nefarious. Yet trying to empirically identify the causal effect of various changes to the institutional environment—money in politics, say, or the rise of modern media—is a tricky exercise. There are powerful selection effects in terms of who runs for office. A careful empiricist attempting to tease out the causal effects of changes in the institutional environment would be rightly worried about both omitted variable bias and reverse causality in taking the most straightforward empirical approaches.

It is fair to say that for a long time the literature did not fully recognize these concerns—or certainly did not embrace them. Then, after the start of the identification revolution in labor economics, several papers looked for causal effects of the incumbency advantage. Ansolabehere et al. (2000) used the change in districts after census years to distinguish between the incumbency advantage for old voters who were previously in a representative’s district and recently added, or new voters. They show that two-thirds of the incumbency advantage comes from these old voters. Levitt et al. (1997) find that pork barrel spending in a district helps incumbents, while Levitt (1994) presents persuasive evidence that, quite surprisingly, campaign spending has little impact on the outcomes of congressional races.

---

6This overview of the literature is based heavily on that in Friedman & Holden (2009).

7One should be careful not to conflate campaign spending and campaign contributions. Having a large war chest could deter other candidates (in terms of either entry or the quality of the candidates who do enter).
Other papers have sought to rule out potential explanations that seem intuitively plausible. For example, Ansolabehere et al. (2006) argue persuasively that the increasing availability and influence of television over time cannot explain the rise in the incumbent reelection rate.

One potential source of the incumbency advantage is redistricting. Indeed, many popular commentators and even scholars have made strong claims that this is almost certainly (and obviously) the case. A series of, again, fairly empirical papers, which are fairly attentive to being able to make causal inferences, cast doubt on this.

For instance, Ansolabehere & Snyder (2002) show that the advantage to being an incumbent in settings where redistricting does not play a role (e.g., the US Senate and gubernatorial races) has risen at similar rates as that for the US House of Representatives. It was long ago noted that the time series of decline of marginal districts is not consistent with redistricting as a cause (see, for instance, Burnham 1970 and Gross & Garand 1984). Gross & Garand, in particular, consider data on marginals back to 1824.

Gelman & King (1994) adopt a more macropolitical approach by estimating the seats-votes curve for various states. Recall that the responsiveness of the seats-votes curve is the slope of the curve at one-half. Gelman & King show that redistricting leads to an increase in responsiveness, so that the share of seats won for a given party is more sensitive to their share of their voteshare.

Friedman & Holden (2009) adopt a regression discontinuity approach by observing that, with relatively few exceptions, redistricting takes place after the decennial census, so that new districts come into effect in 1962, 1972, 1982, and so on. However, the other factors that are potential explanations for the incumbency advantage (e.g., money, media, match quality) tend to evolve smoothly over time. This serves as the basis for identifying the effect of redistricting. By fitting a smooth function (a cubic spline or a high-order polynomial) to the time series of incumbent reelection rates, and then a step function that is only permitted to change in redistricting years, the authors show that for redistricting to be the culprit it would have to be the case that the step function takes steps up in those redistricting years.

In fact, they do not find this effect; if anything, they find evidence of the opposite. Figure 2 from their analysis shows the smooth and step functions. Note that the step function is basically flat, except after 1962 and in 1992. Those are years when certain landmark US Supreme Court decisions occurred (1962) and when the VRA was reauthorized with tighter provisions (1982)—after the 1982 round but before the 1992 round. Friedman & Holden conjecture that these additional constraints on redistricting actually caused the incumbent reelection rate to decline, all else equal. In short, redistricting may cause the initially high level of the incumbent reelection rate, but it cannot have caused the increase since the 1950s.

What exactly causes there to be a high incumbency advantage remains an open question, and it is fair to say that although new social science approaches have been relatively successful in ruling out various potential explanations, they have had very little success in pinning down what the real cause is.

Friedman & Holden (2009) cite two instructive quotes: “Although elections may be uncompetitive for many reasons—including money in politics and the declining prestige of political service—the role of incumbent protection through the redistricting process is undeniable…Thanks to the wizardry of computer programs that draw incumbent-safe districts with ease” (Wilmot 2004). “And it is the yawning gap between the huge problems our country faces today—Social Security reform, health care, education, climate change, energy—and the tiny, fragile mandates that our democracy seems able to generate to address these problems that is really worrying. Why is this happening? Clearly, the way voting districts have been gerrymandered in America…is a big part of the problem” (Friedman 2005).
Geography

One issue of geography that arises specifically in the context of redistricting is how to measure what has become known as the compactness of political districting plans, that is, how oddly shaped they are. It is natural to think that political districts that have particularly odd-looking boundaries have been manipulated in the redistricting process for political advantage. A classic example is the Illinois fourth congressional district, depicted in Figure 3, and commonly referred to as the ear muff district given its odd shape. Indeed, it runs up a freeway for several miles solely for the purpose of connecting voters on the north and south sides of Chicago. Because voters on the north and south sides are known to have, on average, quite different political preferences, this is suspicious.

That oddly shaped districts are suspicious is not controversial. The key question, however, is how to construct a mathematical measure that captures the idea of oddly shaped in a meaningful way and in a single number. There is a long literature on this issue, and a large number of measures have been proposed (see Fryer & Holden 2011 for a list of references too numerous to reproduce here). These measures include the area of various circumscribing figures (circles, octagons, and others), the perimeter length of a district, and the ratio of the perimeter length to the area of a circumscribing figure. Fryer & Holden (2011) point out that most of the measures that have been proposed suffer from at least one of several shortcomings. They argue that a meaningful measure of compactness must allow for comparisons of different districting plans that are not sensitive to population density, physical size, or the number of districts being drawn. They furthermore suggest that measures must apply to districting plans, not to individual districts.

In response to this, Fryer & Holden (2011) propose three axioms that they claim any reasonable districting plan should satisfy:

---

$^9$In the following description, I quote directly from Fryer & Holden (2011, p. 501).
1. Anonymity: The index does not depend on the identity of any given voter.
2. Invariance: The index does not depend on a state’s population density, physical size, or number of districts.
3. Clustering: If two states with the same number of voters, the same number of voting districts, and the same value for the minimum-partitioning problem have different total intradistrict distances, then the state with the larger value is less compact.

They then demonstrate how this can be calculated (efficiently), then rank and map the resulting districting plans using ARC GIS software. Finally, they estimate counterfactual seats-votes curves in several states based on the maximally compact districts.

Axiom 2 makes it possible to compare indices that satisfy it across states, a property that many previous measures fail to satisfy. Axiom 3 is essentially what compactness means: One should put voters who are close together in the same district and voters who are far apart in different districts. Axiom 1 ensures that all voters are equally weighted.

The Fryer-Holden (hereafter FH) measure of compactness consists of two components. Component $N(umerator)$ sums the squared distance between all pairwise combinations of voters in a district and then it sums those objects over all districts in a state. Component $D(enominator)$ is precisely the component $N$ calculation, but for the districting plan that produces the minimum such sum among all districting plans. The FH index (they call it the relative proximity index, or RPI) is component $N$ divided by component $D$, for any districting plans in a state. Notice that it necessarily applies to districting plans, not individual districts.

It is certainly not surprising that the FH index satisfies the three axioms they propose—the index was designed to do just that. What is perhaps more surprising is that they prove the following theorem: Any districting plan satisfying the three axioms ranks districting plans identically to the RPI. That is, given the axioms, the RPI is ordinally unique.
A serious computation challenge arises, however. The component \( N \) is trivial to calculate—it could be done on an iPhone for a modest-size US state. However, calculating component \( D \) (the denominator of the index) is computationally burdensome. In fact, in the language of computer science, it is a nondeterministic polynomial-time hard (NP-hard) problem. That is, the computational complexity rises exponentially with the number of voters in a state. This means, at a practical level, that even using census block data it is impossible to calculate the index for even a medium-size state. Fryer & Holden develop an algorithm based on so-called power diagrams (used in tropical geometry and string theory) that approximates the actual value of the denominator very accurately, but also very computationally efficiently. They then calculate it for the districting plans for the 106th Congress, with data from the US census.

According to that exercise, the five states with the most compact districting plans according to the FH index are Idaho, Washington, Arkansas, Mississippi, and New Hampshire, whereas the five least compact states are Tennessee, Texas, New York, Massachusetts, and New Jersey. It is also interesting to note that the FH index ranks districting plans quite differently to some other popularly used measures. For instance, the rank correlations between the RPI and the dispersion and perimeter measures are \(-0.37\) and \(-0.29\), respectively.

A final step is that Fryer & Holden are able to estimate counterfactual seats-votes curves for the maximally compact districting plans using Gelman & King’s (1994) method. In the handful of states that they calculate, the maximally compact districting plans are all more responsive than the existing ones. It would be very desirable to understand how general this conclusion is, and what might be the reason for it.

A larger issue in political geography than merely calculating what existing districting plans look like is how population shifts over time and voting behavior interact. A series of spectacular papers by Jonathan Rodden and coauthors have shed new light on a range of important issues associated with this.\(^{10}\)

Rodden (2010) begins with the observation that in societies with a high degree of geographic mobility, voters will sort into residential areas with similar demographic profiles—including political preferences. When there is heterogeneity of such mobility—perhaps due to income differences—there will naturally be a rich distribution of political preferences across districts. Rodden summarizes the recent empirical literature that uses advances in the size and richness of data sets, as well as some of the modern empirical techniques (such as those discussed in the introduction to this article), and combines these data with the theoretical literature on political competition with heterogeneous plurality districts (see Rodden’s paper for the references therein).

Putting these two rather distinct literatures together produces a rich set of new insights about how the geographic distribution of political preferences affects the policies (or at least platforms) that candidates and parties choose, and which candidates and parties compete in which elections. A key insight is that, because the distribution of political preferences is left-skewed, Democrats will tend to try and cover a greater range of ideological positions than Republicans. It also provides a rationale, taking preferences to be multidimensional, for why moral-values issues tend to matter more in presidential rather than congressional elections.

Rodden & Chen (2013) show that a significant amount of partisan bias in US legislative elections is caused by patterns of economic/political geography. Their basic observation is that, for reasons

---

\(^{10}\)One on which we can barely touch here is the importance of local economic conditions for the reelection of incumbents. Ebeid & Rodden (2006) show that the link between voter behavior and macroeconomic aggregates is weak in states where economic conditions are largely out of the control of politicians (e.g., because they are heavily natural resource dependent), but strong in other states. This is a version of the classic signal extraction problem that takes place in principal-agent theory in economics.
of urbanization and industrialization, Democrat-leaning voters tend to be concentrated in cities, leading to them winning less than 50% of the seats in elections where they win 50% of the votes. The authors refer to this as unintentional gerrymandering. Their contribution is extremely important in that it focuses our attention on the crucial link between economic and social patterns and electoral outcomes, and it does so in a way that highlights how institutional features of the electoral system interact with those facts. Moreover, the technical approach that they take breaks new ground in developing tools for further study of the links between economic/political geography and electoral outcomes.

Rodden & Chen (2013) match precinct-level voting returns from the 2000 US presidential election with the geographic shape files produced by the US census. This allows them to match to the demographic data contained in the census at the block-group level. They then perform 25 simulations of districting plans for Florida for different possible legislature sizes (2–200). These simulations reveal a pro-Republican bias in the distribution of seats in the legislature. By their estimates, given 50% of the vote, Republicans would win between 56% and 68% of seats in the legislature purely due to geographic features. This is a material effect, even relative to what a sophisticated partisan gerrymanderer could achieve in the absence of geography (see above).

In short, it is simply not possible to ignore geography in thinking sensibly about electoral returns. Moreover, the interaction between strategy gerrymandering (again, see above) and these natural geographic gerrymanders seems like a very promising area of future research.

This is an area where big data techniques and mapping technology have been very fruitfully applied. The interested reader is referred to Stanford’s Spatial Social Science Lab (https://sites.stanford.edu/sssl/) for further details.

VOTER REGISTRATION AND TURNOUT

In electoral systems, such as that in the United States, that require voters to be registered and also allow for voluntary voting, there are two important margins to consider. First, voters have to register, and second, they have to turn out. This leads to an obvious but tricky question: If people do not vote, is it because of barriers to registration or turnout? There are important costs and impediments to both. Registration is often plagued by lack of information, and historically in the United States by direct discrimination. Turnout is often complicated by work or childcare commitments, polling queues, or even the weather.

Perhaps surprisingly, this strand of the literature recognized the importance of, and challenges with, obtaining causal inferences. Field experiments (randomized controlled trials) have often been used, allowing the genuine causal effect of a particular intervention to be obtained (see, for instance, Arceneaux & Nickerson 2009; Dale & Strauss 2009; Gerber & Green 1999, 2000; Gerber et al. 2003; Michelson 2006; Nickerson 2006; and very early work by Gosnell 1927). The focus of these papers is on how to get more registered voters to turn out—i.e., the turnout margin rather than the registration margin.

On the registration margin, Nickerson (2014) randomly assigns a face-to-face registration drive across 620 streets in 6 cities. He finds a 4.4% increase in registration and that 24% of those registered as a result of the intervention turn out to vote. He also finds that the registration effect is larger on poorer streets, but the turnout effect is larger on more affluent streets.

Bhatt et al. (2015) use a natural experiment in Massachusetts in 2012 (based on a legal dispute concerning that state’s compliance with federal voter registration requirements) to estimate the causal effect of lowering voter registration costs on voter registration, turnout, and voting behavior in US presidential elections. They use a difference-in-differences and a triple-differences specification and under both find a statistically significant effect on voter registration and turnout that
is sizable in economic magnitude. Interestingly, conditional on registration, there is no material difference in turnout. There is, however, a large effect on Democratic voteshare. Because, conditional on registration, turnout is not materially different, the authors conclude that the registration margin is the key driver of overall electoral participation.

Another set of papers exploits exogenous shifts in the information set of voters to understand the information-turnout margin. Stromberg (2004) finds that areas with a higher share of radio ownership (and hence subject to more election broadcasts) had higher turnout during the 1920s–1930s. Gentzkow (2006) finds that substitution away from media outlets with higher levels of political coverage reduces turnout.

There is inherent selection bias in making inferences from observational data about who registers and turns out to vote. Controlling for observable variables is of little help. This strand of the literature—as early as 1927—recognized these problems and used field experiments to obtain causal inferences. In more recent times, techniques using natural experiments have also been fruitfully applied by political scientists and economists.

ELECTION LAW IMPLICATIONS

In this section I briefly consider some of the implications of the previous sections for election law and discuss a handful of recent cases. It is not my purpose here to provide a detailed account of any aspect of the large and important subject of election law. The interested reader is referred to Issacharoff et al. (2012) for a definitive and classic treatment of those issues.

As mentioned above, Vieth v. Jubelirer squarely raises (again) the question of whether there exist judicially manageable standards for adjudicating partisan gerrymandering claims. Social science may play an important role in providing such standards, particularly given the advent of extremely rich data at a detailed geographic level. The work of Rodden and coauthors, discussed above, seems to this author to be a significant step in the direction of articulating such a standard.

So, too, with which campaign expenditures constitute speech and which do not, social science may play an important role. At the heart of such a distinction—and the level of scrutiny the Court is likely to apply—is the question of what constitutes content-neutral regulation. Advances in our understanding of how individuals respond to certain speech (e.g., through the use of functional magnetic resonance imaging machines), as well as the ability of computer algorithms to parse speech and text, break it into components, and classify it, all point to the intriguing prospect of a principled, social science–based approach to classifying speech and providing a taxonomy of it.

With regard to redistricting, Cox & Holden (2011) take up the matching slices characterization of the optimal gerrymandering strategy and explore several of the implications of it for the VRA and the legal treatment of racial gerrymandering and race-conscious redistricting. One important question is whether requirements within the VRA have any partisan consequences. The two that Cox & Holden consider are those contained in sections 2 and 5. As mentioned in the introduction (and since their paper), the US Supreme Court effectively invalidated section 5 (the so-called “preclearance provisions”) in Shelby County v. Holder. Section 2, which prohibits states from using any voting practice “in a manner which results in a denial or abridgement of “ minority voting rights [42 U.S.C. §2 (1973)], remains constitutional. As noted above, in some instances section 2 has led to a requirement to create so-called majority-minority districts, i.e., districts in which a racial minority constitutes a majority of the voters.

Cox & Holden point out that majority-minority districting requirements affect Democratic and Republican gerrymanders differently. This is because the VRA imposes different constraints on Democratic and Republican redistricting authorities. The requirements of the VRA are consistent...
with the optimal strategy for partisan gerrymandering by Democrats, because they seek to draw districts with small majorities of African American voters.

By contrast, partisan redistricting Republicans would not want to combine African American voters as a majority of any single district but would rather concentrate them into districts where conservative Republicans constitute (small) majorities. In other words, the VRA has differential partisan impact, given the optimal gerrymandering strategies of Democrats and Republicans. It is beyond the scope of this article to go much further into the issue, but one naturally wonders whether a Supreme Court that has ruled section 5 of the VRA unconstitutional might find an implicit partisan slant (under the optimal gerrymander) in section 2 unappealing on certain constitutional grounds as well.

A second question is whether, if a widely accepted and sensible measure of “oddly shaped” or “not compact enough” were adopted, it would have an important impact on districting practices. For gerrymandering, the question really hinges on patterns of residential segregation discussed above. One thing to note is that geography is more important for the matching slices strategy than pack-and-crack, because matching slices requires more nuanced district creation. Instead of just putting Democrats with Republicans, it requires putting particular Democrats with Republicans (e.g., the most ardent Democrats with slightly smaller slices of ardent Republicans in district 1 of a Democrat matching slices gerrymander).

For the creation of majority-minority districts it hinges on this, too. Yet the evidence of how oddly shaped the individual districts giving effect to race-conscious redistricting are suggests that it may, in fact, constrain this enterprise more (see, for instance, the majority opinion in Shaw v. Reno).

A third election law question concerns voter identification laws. Several recent cases consider such laws in various states (for a notable example, see Frank v. Walker). In the above section on voter registration and turnout, we point to a recent paper by Bhatt et al. (2015) looking at the effect of the costs of registration on turnout and how the treatment effects of this differ by partisanship. Voter identification laws, in economic terms, are an increase in registration costs for voters. They have to spend the time and money to obtain a valid government-issued ID. Some voters have it already, some do not. And as the aforementioned case highlights, those who do not are not a representative cross-section of the electorate. The empirical evidence offered by Bhatt et al. (2015) suggests that, aside from what other effects they may have, such laws materially reduce actual voting, and that there is a partisan skew (away from Democrats) to this reduction.

CONCLUDING REMARKS

Like many subfields in economics and political science, in recent years the study of voting and elections has been fortunate to have much richer data available, and significantly more computing power with which to analyze those data. I conclude by suggesting five important open questions in the field.

Is gerrymandering self-correcting (at all)? One might wonder whether gerrymandering is, at least to a degree, self-correcting in the sense that by cutting it too finely a redistricter risks a very bad outcome. A big statewide swing could do more damage to a party that has heavily gerrymandered the state than to one that has not. It seems unlikely that gerrymandering is completely self-correcting, because so many resources are devoted to doing it and attempting to be in the position to do it. Yet the O’Connor conjecture (as I refer to it) has some resonance—a risk-averse gerrymanderer should not cut things too finely. Indeed, we saw above in Table 1 that the optimal gerrymander involves cutting individual districts less finely as the signal becomes less precise. The key question is how to quantify the self-correcting component. Friedman-Holden matching
slices tell us how to evaluate this in terms of expected districts won, but it seems essential to know something more about the preferences of the redistricter to fully address the question.

How do social connections and networks affect voting behavior? It has been increasingly realized in recent years that much economics activity takes place between people or institutions that are already connected in some manner. Pioneering work in the economics of networks by Jackson & Wolinsky (1996) has sparked a large literature. Indeed, that literature draws heavily on concepts (e.g., Katz-Bonacich centrality) from other parts of the sciences and social sciences. It seems clear, if only from casual inspection of how US primary elections work, that who one is socially connected to is an important determinant of how one votes. Certainly, campaigns that make use of social network data (such as Facebook profiles) and use social networking tools themselves seem to think so. Yet, to the best of my knowledge, there is no canonical model of voting on networks. Using graph-theoretic techniques to study how voting on social networks works and changes over time is an important and enticing research agenda.

How large an impact do big donors have on electoral outcomes? The Citizens United decision by the United States Supreme Court was viewed in many corners as a boon to big-money donors who already have significant influence over the political process. For instance, The New York Times’ (2010) editorial page put it this way: “With a single, disastrous 5-to-4 ruling, the Supreme Court has thrust politics back to the robber-baron era of the nineteenth century.” It even rated a (controversial) mention in President Obama’s subsequent State of the Union address. There is certainly anecdotal evidence that political candidates pander in certain ways to big donors—for instance, the way in which current US Republican presidential hopefuls court casino magnate Sheldon Adelson’s endorsement is known colloquially as “the Las Vegas primary.” However, there is, to the best of my knowledge, little systematic evidence about the impact of big-money donors on electoral outcomes, let alone policy outcomes. It is the latter that should be of primary concern for those who are not politicians themselves. In principle, relatively unfettered big-money access could have a major impact on policy outcomes, or very little. The forces going in the former direction are clear, but there are countervailing forces. First, competition among donors may lead to a lot of money being spent but policy differences cancelling each other out, much like a consumer product advertising war. Second, because politicians cannot make binding policy commitments, donors may not get what they ultimately want. It may be possible to sustain some kind of relational contract, but politicians may not be sufficiently long-run players to sustain this. Ultimately, this is an empirical question, and it would be highly desirable to see systematic evidence on the matter.

What effect on districting composition does natural geographic sorting have? The work by Rodden and coauthors mentioned previously demonstrates that geographic sorting that occurs for reasons not directly to do with elections may nonetheless have very important effects on elections. Justice Scalia has referred to “natural packing effects” (Vieth v. Jubelirer) that come from geographic sorting, hinting that the pack-and-crack strategy of a gerrymanderer may occur naturally. Because, as we discussed above, that strategy is not the unconstrained optimal one, this particular argument does not appear compelling. But as the Rodden work shows, there is a rich set of other implications of geographic sorting that is only beginning to be explored.

To what extent can politicians be removed from the design of political institutions? Going back to where this paper began, with partisan gerrymandering, it is a very peculiar feature of (at least) American democracy that politicians themselves largely draw the boundaries of the electoral districts they end up representing. But this involvement of politicians in the design of the process goes well beyond redistricting. All manner of electoral rules are influenced, if not determined, by political actors. There is, in principle, no reason why there should not be a clear demarcation between those who design electoral institutions and those who participate in them. A significant question is, can there be such a clear demarcation? One currently topical area seeking to “get
politicians out of the design of the political process” is through algorithmic redistricting (Fifield et al. 2015, Holden 2015). It remains to be seen how successful this can be as a practical matter, and whether such approaches can be extended to other aspects of the design of electoral institutions.

There is much work to do to answer these questions, but some of the new social sciences methods and approaches mentioned in this paper provide a useful starting point and set of tools for that work. Moreover, the answers to these and other questions in the field will enrich our understanding of how institutional arrangements affect democratic outcomes. The prospect of even richer data, including perhaps those used by campaign organizations themselves, and plentiful computational power for analyzing and representing analyses of those data is as enticing as it is important.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I thank Rosalind Dixon, Martin Krygier, and an anonymous editor for helpful discussions and comments, and acknowledge the Australian Research Council for support under Future Fellowship FT130101159. I also thank Adam Cox, John Friedman, and Roland Fryer, with whom I have coauthored several papers over the years on topics discussed here.

LITERATURE CITED


Avery v. Midland County, 390 U.S. 474 (1968)


Cedengez v. Green, 328 U.S. 549 (1946)


Frank v. Walker, No. 14-2058 (7th Cir. 2014)


Shelby County v. Holder, 570 U.S. ____ (2013)


Wesberry v. Sanders, 376 U.S. 1 (1964)

Contents

“Be Operational, or Disappear”: Thoughts on a Present Discontent
  Christopher Tomlins ................................................................. 1

A Comparison of Sociopolitical Legal Studies
  Mauricio García-Villegas .......................................................... 25

Comparative Legal Research and Legal Culture: Facts, Approaches, and Values
  David Nelken ................................................................. 45

Corporate Complicity in International Human Rights Violations
  Leigh A. Payne and Gabriel Pereira ........................................... 63

Field Research on Law in Conflict Zones and Authoritarian States
  Mark Fathi Massoud .............................................................. 85

Law, Settler Colonialism, and “the Forgotten Space” of Maritime Worlds
  Renisa Mawani ................................................................. 107

Lawyers, Globalization, and Transnational Governance Regimes
  Milton C. Regan, Jr. .............................................................. 133

Possibilities and Contestation in Twenty-First-Century US Criminal Justice Downsizing
  Michelle S. Phelps .............................................................. 153

Takings as a Sociolegal Concept: An Interdisciplinary Examination of Involuntary Property Loss
  Bernadette Atuahene ............................................................ 171

The Rule of Law: Pasts, Presents, and Two Possible Futures
  Martin Krygier ................................................................. 199

Theorizing Transnational Legal Ordering
  Gregory Shaffer .............................................................. 231

Voting and Elections: New Social Science Perspectives
  Richard Holden ................................................................. 255
Next-Generation Environmental Regulation: Law, Regulation, and Governance

Neil Gunningham and Cameron Holley .................................................. 273

Randomized Control Trials in the United States Legal Profession

D. James Greiner and Andrea Matthews ............................................. 295

Neoliberalism, Post-Communism, and the Law

Bojan Bugarić ....................................................................................... 313

The Regulation of Witchcraft and Sorcery Practices and Beliefs

Miranda Forsyth .................................................................................. 331

Impact of Judicial Elections on Judicial Decisions

Herbert M. Kritzer ............................................................................. 353

Overlooked and Undervalued: Women in Private Law Practice

Joyce S. Sterling and Nancy Reichman .................................................. 373

Legal Discrimination: Empirical Sociolegal and Critical Race Perspectives on Antidiscrimination Law

Lauren B. Edelman, Aaron C. Smyth, and Asad Rabim ............................. 395

Legal Socialization: Coercion versus Consent in an Era of Mistrust

Rick Trinkner and Tom R. Tyler ......................................................... 417

Corporate Lawyers in Emerging Markets

Bryant G. Garth ................................................................................... 441

The Firm as a Nexus of Organizational Theories: Sociological Perspectives on the Modern Law Firm

Alan James Kluegel ........................................................................... 459

Sociolegal Approaches to the Study of Guilty Pleas and Prosecution

Brian D. Johnson, Ryan D. King, and Cassia Spohn ............................... 479

Indexes

Cumulative Index of Contributing Authors, Volumes 2–12 ....................... 497

Cumulative Index of Article Titles, Volumes 2–12 .................................. 500

Errata

An online log of corrections to Annual Review of Law and Social Science articles may be found at http://www.annualreviews.org/errata/lawsocsci