Crossing the District Line: Border Mismatch and Targeted Redistribution

Allison Stashko*

University of Utah

March, 2020

Abstract

Electoral district borders regularly cross the borders of local governments. At the same time, legislatures allocate resources using transfers to local governments. Political parties may try to target these transfers in order to win elections, but can only do so imperfectly because of border mismatch. This border mismatch creates inequality: otherwise similar local governments receive different transfers depending on the district map. To show this, I incorporate border mismatch into a model of political competition and test the predictions using data on transfers from U.S. states to counties. The results demonstrate a novel link between redistricting and voter welfare.

JEL Classification: D72, H77, K16

Keywords: Targeted Redistribution, Redistricting, Gerrymandering

*David Eccles School of Business, University of Utah. Email: allison.stashko@eccles.utah.edu. I would like to thank Laurent Bouton, Garance Genicot, and Arik Levinson for their guidance. I also thank Gaurav Bagwe, Jennifer Brown, John Londregan, Juan Margitic, Adam Meirowitz, JJ Naddeo, Nicola Persico, and Dario Sansone for their helpful comments. This paper benefited from the feedback of seminar participants at Cornell University, London School of Economics, Louisiana State University, New York University, Tulane University, Stanford GSB, University of Nottingham, University of South Carolina, University of Rochester, and University of Utah.

1 Introduction

Electoral district borders and local government borders are often mismatched. In the United States, for example, counties can span many electoral districts and districts can include many counties. In some cases, border mismatch is inevitable, as districts must have roughly the same population size and counties, cities, and towns do not.[1] In other cases, border mismatch results from gerrymandering, where political parties draw electoral districts to improve their election outcomes.

This paper shows that border mismatch between electoral districts and local governments leads to inefficient and more unequal public spending. I focus on a novel mechanism that links a district map to the distribution of public resources. There are two key ideas underlying the logic of this mechanism. First, political parties strategically direct spending to certain districts in order to win more seats in a legislature.[2] Second, elected representatives use transfers to local governments to allocate resources.[3] When local governments are split across many districts, political parties can only imperfectly target those districts. Public spending becomes distorted because elections are held at the district-level, spending occurs at the local government-level, and borders are incongruent.

To characterize the effects of border mismatch on public spending, I incorporate the concept of imperfect targeting into a model of political competition. In the model, two parties compete to win as many districts as possible, but can only do so through county-level spending. The main result is that border mismatch leads to inequality across counties. Two otherwise identical counties can receive different amounts of transfers depending on how the district lines are drawn. In particular, transfers to a county will depend on the number of districts that the county is in and the voting behavior in other counties that share those districts. Changes in public spending caused by border mismatch are inefficient since they are driven solely by political competition.

As an example, consider two counties, A and B, that are identical except for the way that they are districted. County A is also an electoral district; it has one representative. County B is split in half across two districts; it has two representatives that are shared

[1]For this reason, border mismatch occurs in many countries where districts are based on population size (e.g., Australia, Canada, Germany, India, Mexico, and the United Kingdom). In the U.S., all state legislative districts are required to have roughly the same population per representative (*Baker v. Carr*, 369 U.S. 186 (1962) and *Reynolds v. Sims*, 377 U.S. 533 (1964)). Border mismatch also occurs between federal congressional districts and counties, though less frequently.

[2]This idea is central to a large literature on distributive politics, including models of legislative bargaining and of political competition between two parties (for example, Enelow and Hinich 1982, Lindbeck and Weibull 1987, Cox and McCubbins 1986, Dixit and Londregan 1996, Baron and Ferejohn 1989, Baron 1991, Banks and Duggan 2005, Persson and Tabellini 2000), as well as empirical studies (see Golden and Min 2013).

[3]State to local and federal to local government transfers account for 35% and 5% of local government revenue, respectively. Local government spending accounts for 33% of all government spending in the U.S. (Census of Governments, 2012).

with neighboring counties. This example raises the question, is a county better off with one 'whole representative' or two 'half representatives'? The answer is not obvious. On one hand, county B has the advantage of having more representatives in total. Political parties may spend more in county B because they can can influence two elections at once. On the other hand, county A has the advantage of not sharing its representative with any neighbors. By comparison, county B has only a limited influence in each of its district elections. In fact, political parties could ignore county B in both district elections in favor of neighboring counties, especially if the vote shares in neighboring counties are more sensitive to public spending. The model captures this tradeoff and predicts precisely when one effect dominates the other.

Intuitively, in each district, political competition concentrates in the counties where vote shares are most sensitive to transfers. I focus on two factors that make vote shares more sensitive: voting eligibility and turnout rates. This means that a county receives more transfers if it has a higher share of eligible voters or a higher turnout rate *relative* to all the other counties in the same districts. Thus, a county with high turnout relative to its neighbors is better off split across many districts. A county with low turnout relative to its neighbors is better off in its own district.

From the model, I derive a set of empirical predictions that can be tested using data on intergovernmental transfers from U.S. states to counties. Testing the predictions of the model requires highly disaggregated measures of voting eligibility and turnout. I compile precinct-level election returns and block-level census data in order to compute these measures at the county-district level. Data on intergovernmental transfers come from the Census of Governments from 2007 and 2012. To address the fact that district maps are not randomly drawn, I exploit plausibly exogenous variation in the number of districts per county. This variation comes from the rule that districts must have equal population size (similar to the rule for classroom size in Angrist and Lavy, 1999).

Empirical evidence is consistent with the model and suggests that border mismatch leads to sizable distortions in the allocation of public resources. Additionally, because of the way voters are distributed geographically, border mismatch creates more losers than winners. That is, most individuals would benefit if their county were split across fewer districts. For the median county, changing the district map so that a county has its own representatives would increase transfers from the state by 9%, or $126 per person.

Additional empirical analyses provide evidence that district borders are driving these results. As a placebo test, I measure voting eligibility and turnout in a county relative to neighboring counties that are not in any of the same districts. This measure is uncorrelated with transfers, suggesting that general geographic patterns of turnout cannot explain the results. I also show that regressing transfers on county-level turnout leads to misleading null results. This suggests that future empirical studies of distributive politics should take electoral district borders into account when considering local government outcomes.[4]

The results contribute to ongoing debates about redistricting and partisan gerrymandering. Political parties today are able to draw maps to their advantage with increasing accuracy. And yet, little is known about how district maps affect individual welfare. Instead, most studies of redistricting focus on electoral outcomes such as the partisan (Cox and Katz 1999, McCarty, Poole, and Rosenthal 2009), ideological (Caughey, Tausanovitch and Warshaw 2017), or racial (Cameron, Epstein and O'Halloran 1996) composition of a legislature. This paper offers a more direct link between redistricting and voter welfare by studying public spending outcomes. I also find that gerrymandered district maps have more instances of border mismatch (consistent with Edwards et al., 2017). Thus, the paper demonstrates a new way in which redistricting in general, and gerrymandering in particular, affects voters.

While this paper asks a positive question about district borders and public spending, normative implications for redistricting follow. The most straightforward implication is that district borders should cross local government borders as infrequently as possible. Most states already require that districts "respect political borders", but the priority placed on doing so varies.[5] These laws intend to keep voters with similar interests grouped into the same districts, and evidence that voters sort into local jurisdictions supports this rationale (Epple Romer and Sieg, 2001). This paper provides an additional rationale. Respect for political borders limits inefficient spending since local government borders are integral to public finance. In contrast, other districting conventions, like having compact district shapes, may be less important since they have no direct ties to policymaking.

At the extreme, border mismatch would be eliminated if counties were used as state legislative districts. Before several Supreme Court decisions in the 1960's, this is exactly what many states did. However, these state legislative districts were malapportioned, in that districts varied too much in population size. The Supreme Court decided that small counties had too much power in state legislatures and required states to redraw maps with an equal number of representatives per person in every

[4]Other empirical studies have had to confront the differences between local government and district borders. They either exclude counties that span multiple districts (Martin 2003, Ansolabehere and Snyder 2006) or create county-level measures using population-weighted averages of districts (Ansolabehere, Gerber, and Snyder 2002). This is different from the relative measure used here. For example, a county split evenly across two districts with turnout rates of 50% has an average measure of turnout of 50%. But, the same county may represent only a small share of the turnout in both districts, say 10%.

[5]Other than having roughly equal population size across districts and abiding by the Voting Rights Act of 1965, states have full discretion over redistricting laws. Seven states have no law or guideline regarding border mismatch.

district. Ansolabehere, Gerber and Snyder (2002) study the effects of the ensuing decrease in malapportionment. They find that transfers from states to counties increase in the number of representatives per person. Importantly, I find that the effects of border mismatch on transfers are of the same order of magnitude as the effects of malapportionment. This paper therefore highlights a new tradeoff. When drawing a district map, a planner must balance equal representation per person against equal representation per county.

The rest of the paper is organized as follows. Section 2 discusses related theoretical and empirical studies of distributive politics. Section 3 presents the model. Section 4 derives empirical specifications and predictions. Section 5 describes the data and Section 6 presents the empirical results. Section 7 discusses implications for redistricting and the relationship between gerrymandering and border mismatch. Section 8 concludes.

2 Related Literature

In this paper I present a probabilistic voting model with imperfect targeting. Probabilistic voting models are tractable models with many empirical applications to public spending and campaign efforts (see review by Golden and Min 2013). The basic intuition is that political parties compete over voters who are on the margin of choosing one party over the other. These are known as 'swing voters'. As in previous work (Cascio and Washington 2014, Martin 2003, Stromberg 2004a), I argue that political parties also use information about exogenous variation in voting eligibility and turnout when targeting swing voters. Since political parties also make efforts to mobilize turnout (Ansolabhere and Snyder 2006, Hall and Thompson 2018), I will address the possibility that turnout is endogenous to transfers in the empirical section. Of course, there are other ways to conceptualize how political parties compete for votes. Political parties may target swing voters (Lindbeck and Weibull 1987), core voters (those who have strong ideological preferences; Cox and McCubbins 1986), or both (Dixit and Londregan 1996). None of the theoretical results in this paper require variation in the distribution of swing or core voters, though I explore them as alternative explanations in the empirical section.

The novel feature of the probabilistic voting model in this paper is that political parties are constrained in their ability to target districts. I consider a targeting constraint that is flexible in order to reflect the reality of border mismatch: a district may include many targetable groups and a targetable group may span many districts. Existing assumptions about the ability to target special interest groups in the literature are less general and can be interpreted as a special cases of the assumption that I make.

For instance, in most models of district elections, targetable groups and districts are the same (Lindbeck and Weibull 1987, Persson and Tabellini 2000, Stromberg 2008). In others, a targetable group can contain many districts (i.e., media markets, Fletcher and Slutsky, 2011) or districts can contain many targetable groups (Battaglini 2014, Bouton, Genicot, and Castanheira 2019).[6] Martin (2017) studies yet another targeting constraint in a legislative bargaining setting. There, a legislature must choose a policy that has fewer dimensions than there are districts. This captures the idea that some policies are formulaic and depend on only a handful of demographic characteristics. Formally, this constraint is analogous to the special case where there are fewer counties than districts.[7] Finally, Milesi-Ferreti, Perotti and Rostagno (2002) assume that political parties can target social groups that reside in multiple districts and can also target the districts themselves. Their aim is to study public good provision under different electoral systems.

While I focus on political competition between parties, border mismatch could also affect the behavior of individual representatives. In abstracting from the behavior of representatives, I rule out targeted spending that results from a common pool problem (Baron 1991, Baqir 2002, Berry 2008) and the possibility that representatives free-ride when they share counties (Besley 2006). However, Snyder and Ueda (2007) find no evidence of free-riding among multi-member state legislative districts in the United States. There is, however, evidence that splitting a local government across multiple districts is associated with weaker government accountability (Bowen 2014, Winburn and Wagner 2010, Gulzar and Pasquale 2017). Although this is an interesting mechanism that implies a relationship between the number of districts per county and transfers, the mechanism would fail to explain the important role of neighboring counties in determining transfers.

The model further abstracts from the behavior of local governments. If local governments also engage in targeted spending, then political parties at the state level would predict where transfers would go within a county and adjust their strategies accordingly (Dixit and Londregan 1998). However, in other cases a county government does not have discretion to target transfers (e.g., funding for road repairs). Thus, the imperfect targeting constraint characterized in this paper may bind for some types of spending and not for others. For this reason, I test the predictions of the model

[6]The case where groups are smaller than districts yields results that are present under border mismatch. In particular, when political parties can target groups that are smaller than districts, then transfers are increasing in a group's share of swing voters in the district. Bouton, Genicot, and Castanheira (2019) call this the 'relative sensitivity' effect and compare distributive outcomes under majoritarian and proportional electoral systems. Here, the model generalizes the relative sensitivity effect in majoritarian elections to the case where targetable groups can be in multiple districts.

[7]Also related to Martin (2017), I discuss the possibility that formulaic spending may limit the discretion of political parties to target counties in the empirical section.

separately for different types of transfers. Overall, the empirical evidence in this paper does not refute theories of free riding, accountability, or local government politics. The evidence does, however, suggest that the imperfect targeting created by border mismatch determines transfers to local governments in some part.

Last, this paper relates to empirical studies of distributive politics.[8] Several papers find that a local government benefits from having a greater number of representatives in the central legislature (Hirano 2006, Hirano 2011, Halse 2013, Hirano and Ting 2015, Yuan 2018). Similarly, small states have an advantage over large states in the U.S. Senate because they have more representatives per person (Atlas et al., 1995, Lee 1998 and 2000, Ansolabehere, Snyder, and Ting 2003, Hoover and Pecorino 2005, Knight 2008). In the context of state to local government transfers in the U.S., however, border mismatch complicates the relationship between representation and spending. More representatives may be good or bad, as the neighbors who share those representatives matter too. The policy implications of the empirical results also differ from many studies of representation. Unlike longstanding political institutions, variation in representation caused by border mismatch can be addressed every ten years through the redistricting process.

3 Model

In this section I introduce a model of distributive politics with border mismatch. Two parties compete to win as many electoral districts as possible, but can only do so through a proposal of county-level spending, where counties and districts do not coincide.

A continuum of citizens is partitioned into C counties, $c = 1, ..., C$, and is also partitioned into D electoral districts, $d = 1, ...D$. Let $n(c, d)$ be the size of the population in the intersection of county c and district d. Each district has the same population size, normalized to 1 ($\sum_c n(c, d) = 1$), while counties vary in population size. Let n_c denote the population of county c ($\sum_d n(c, d) = n_c$). Border mismatch occurs whenever a county intersects more than one district. Formally, border mismatch occurs unless the partition of counties is a refinement of the partition of districts.

[8] There is also a set of empirical papers that address border mismatch between electoral districts and political or media areas. There, border mismatch is part of an identification strategy rather than the object of study itself (Snyder and Stromberg 2010, Jennes and Persyn 2015, Spenkuch and Toniatti 2018, Gulzar and Pasquale 2017).

3.1 Public Consumption

An individual in county c benefits from government transfers to their county, q_c. I impose the following functional form on the utility in transfers:

$$u_c(q_c) = \frac{a_c}{1-\frac{1}{\rho}}\left(\frac{q_c}{n_c^{\alpha}}\right)^{1-\frac{1}{\rho}}$$

where $a_c > 0$ and $0 < \rho < 1$, so that $u_c(.)$ is increasing and concave. The parameter $\alpha \in [0, 1]$ gives flexibility in interpreting spending in county c. If $\alpha = 0$ then q_c is a local public good, if $\alpha > 0$ then there is some degree of congestion, and if $\alpha = 1$ then q_c is a private transfer. Individuals may also differ in how they benefit from a dollar of spending depending on county-specific factors captured by the parameter a_c.

I introduce this functional form for two reasons. First, it helps to clarify that the transfers under consideration can be private goods or local public goods. This is especially important in the empirical setting, since the nature of goods funded through intergovernmental transfers is not obvious. Second, the parameterized utility function will help to interpret empirical specifications in Section 4. All of the theoretical results to follow, however, require only that $u_c(.)$ is increasing and concave in q_c (see Appendix A for proofs).

3.2 Political Competition

Two political parties, A and B, compete to win as many seats in the state legislature as possible.[9] Each district elects one representative by majority rule. Parties simultaneously propose a policy of county-level transfers, $\mathbf{q^j} \in \mathbb{R}_+^C$, subject to a balanced budget constraint, $\sum_c q_c^j \leq y$, where y is the total budget to be distributed and $j = A, B$. Political parties face uncertainty over voting behavior at the time when they commit to their policy. After uncertainty resolves and voters vote, the party that wins the most district elections implements its proposed policy.

[9] Political parties may alternatively aim to win a majority or supermajority of seats in a legislature. In the Online Appendix, I solve the problem where political parties simultaneously try to maximize the probability of winning a majority of seats and incorporate the predictions into empirical analysis. As in Lindbeck and Weibull (1987), Snyder (1989), and Stromberg (2008), this objective creates an advantage for counties that are in districts that favor the leading party. Such districts are more likely to be pivotal in deciding who wins a majority of seats. I find only weak empirical evidence that the probability of being in a pivotal district affects transfers from states to counties. This is consistent with some empirical evidence that parties aim to maximize the number of seats won in the U.S. Congress (Incerti 2015).

3.3 Voting

Individuals vary in whether or not they are eligible to vote, the likelihood that they turn out to vote, and their ideological preferences over parties A and B. Let $e(c,d) \in (0,1]$ be the share of the population that is eligible to vote in county c and district d. An eligible voter in county c and district d votes with probability $t(c,d) \in (0,1]$.

Each eligible voter has ideological preferences over parties A and B that are independent from their utility in transfers and can not be manipulated by either party. Let ν_i be individual i's preference for party B over party A. Voters' ideological preferences are subject to an aggregate uncertainty shock, η, which resolves after the policy is announced and before voting. The aggregate uncertainty can be interpreted as any state-wide election news that realizes after platforms have been announced and before election day.

When voting, individual i in county c chooses party A if and only if the differential in utility from transfers exceeds their ideological preference for party B:

$$u_c(\mathbf{q}^A) - u_c(\mathbf{q}^B) > \nu_i + \eta.$$

Political parties know the distribution of ideological preferences ν_i in each county-district intersection. They form correct expectations over turnout rates and the aggregate shock to ideological preferences. For simplicity, I use uniform distributions to represent the distribution of ideological preferences:[10]

$$\begin{aligned} \nu_i &\sim \text{Unif}\left(\beta(c,d) - \frac{1}{2\phi}, \beta(c,d) + \frac{1}{2\phi}\right) \\ \eta &\sim \text{Unif}\left(-\frac{1}{2\gamma}, \frac{1}{2\gamma}\right) \end{aligned}$$

where $\beta(c,d) \in \mathbb{R}$ represents the bias toward party B in county c and district d. The parameter $\phi > 0$ determines the variance in voters' ideological preferences. Given that preferences are uniformly distributed, ϕ is also proportional to the margin of voters who are indifferent between parties A and B (if such a margin exists). For this reason, ϕ is sometimes referred to as a measure of 'swingness' (Lindbeck and Weibull 1987). For now, I suppress variation in ϕ. Since variation in swingness is difficult to measure empirically, it is an advantage that none of the results require variation in swingness.

[10] As in other probabilistic voting models, this uncertainty is critical for ensuring the existence of an equilibrium in pure strategies. Uniform distributions simplify the analysis and ensure that each party's objective function is concave in their own policy. Persson and Tabellini (2000) and Bouton, Genicot, and Castanheira (2019) also use uniform distributions for these reasons. Alternatively, one could make assumptions on the shape of the distribution to ensure the existence of an equilibrium. However, these assumptions are difficult to verify. For example, if ν_i and η are normally distributed, additional assumptions are required to guarantee existence of an equilibrium (see Lindbeck and Weibull, 1987).

Further, allowing for variation in swingness at the county or county-district level is tractable and does not change the interpretation of the theoretical results. I show this in the Online Appendix and also use existing measures of swingness in some extensions to the empirical analysis.

3.4 The probability of winning a district

Party A wins district d with probability p^d. To simplify notation, let $v(c,d)$ be the expected share of the population that votes in county c and district d, $v(c,d) = n(c,d)e(c,d)t(c,d)$.

$$\begin{aligned} p^d =& Pr\left(\sum_c v(c,d)\ Pr(i \in c \cap d \text{ votes for } A) \geq \frac{1}{2}\right) \\ =& Pr\left(\sum_c v(c,d)\left[\frac{1}{2} + \phi\left(u_c(\mathbf{q}^A) - u_c(\mathbf{q}^B) - \eta - \beta(c,d)\right)\right] \geq \frac{1}{2}\right). \end{aligned}$$

Taking probability over η, the probability that A wins district d simplifies to

$$p^d = \gamma\ \frac{\sum_c v(c,d)\ (u_c(\mathbf{q}^A) - u_c(\mathbf{q}^B) - \beta(c,d))}{\sum_c v(c,d)} + \frac{1}{2}.$$

Here I assume that there are voters at the margin of choosing A or B in every county-district intersection, for every realization of η and every policy combination. This is ensured by constraints on the parameters of the distributions of ideological preferences. I also assume that every district is contestable, meaning that it could be won by either A or B. These two assumptions are summarized below and ensure that every county receives positive transfers in equilibrium.[11]

Assumption 1 (Interior Solution). *For all $\mathbf{q}^A$, $\mathbf{q}^B$ and for all η,*

1. *All counties have swing voters in any district that they intersect:*

$$u_c(\mathbf{q}^A) - u(\mathbf{q}^B) - \beta(c,d) - \eta \in \left[\frac{-1}{2\phi}, \frac{1}{2\phi}\right] \text{ for each } d \text{ such that } n(c,d) > 0$$

2. *All districts are contestable:*

$$\frac{\sum_c v(c,d)(u_c(\mathbf{q}^A) - u_c(\mathbf{q}^B) - \beta(c,d))}{\sum_c v(c,d)} \in \left[\frac{-1}{2\gamma}, \frac{1}{2\gamma}\right]$$

To see how county-level spending influences a district election, consider the change

[11] I could alternatively allow for incontestable districts and for areas with no swing voters. This would lead to corner solutions where counties in no contestable districts and counties with no swing voters receive zero transfers.

in the probability that party A wins district d as spending in county $\hat{c}$ increases:

$$\frac{\partial p^d}{\partial q^A_{\hat{c}}} = \gamma \, \frac{v(\hat{c}, d)}{\sum_c \, v(c, d)} \, \frac{\partial u_{\hat{c}}}{\partial q^A_{\hat{c}}}. \tag{1}$$

For every additional dollar spent in county $\hat{c}$, party A increases the marginal utility of all individuals living there. This additional spending convinces a mass of voters at the margin to vote for A instead of B in county $\hat{c}$. However, it is possible that some of district d voters are outside of county $\hat{c}$. The overall change in the probability of winning a majority of votes in the district is proportionate to the share of district d voters in county $\hat{c}$.[12] Next, with an expression for p^d, it is simple to write each political party's objective function.

3.5 Political equilibrium

Party A chooses $\mathbf{q^A}$ to maximize the expected number of districts won, subject to the balanced budget constraint:

$$\max_{\mathbf{q}^A} \sum_d p^d \quad \text{s.t.} \quad \sum_c q^A_c \leq y.$$

Party B maximizes $\sum_d (1 - p^d)$ subject to its budget constraint.

Given uniform uncertainty over ideological preferences, the objective functions for parties A and B are continuous and concave in the choice of policy (see Appendix A). A unique pure strategy Nash equilibrium exists and can be characterized using first order conditions.

Proposition 1. *There is a unique pure strategy Nash equilibrium* $(\mathbf{q}^{A*}, \mathbf{q}^{B*})$ *in which* $\mathbf{q}^{A*} = \mathbf{q}^{B*} = \mathbf{q}^*$*, and there exists* $\lambda > 0$ *such that for each county* $\hat{c}$

$$\sum_d \frac{v(\hat{c}, d)}{\sum_c v(c, d)} \frac{\partial u_{\hat{c}}}{\partial q_{\hat{c}}} = \frac{\lambda}{\gamma}. \tag{2}$$

In equilibrium, Parties A and B each allocate resources in order to equate across counties the marginal change in the probability of winning another seat. The problem is symmetric, so both parties converge to the same policy. The change in the probability of winning an additional seat from a marginal increase in spending in county $\hat{c}$ can be broken into two parts: a) the change in marginal utility for all voters in that county,

[12] In a more general model where ideological preferences are not uniformly distributed, the mass of voters on the margin of choosing A or B would also depend on the degree of ideological bias in the county-district. Although this effect is absent from the baseline model, bias plays a similar role in the model where parties maximize the probability of winning a majority of districts. In the Online Appendix, both the theoretical and empirical models allow for ideological bias in a county-district to affect equilibrium transfers.

and b) the share of a district's voters that live in that county, for every district that the county intersects. I call this second term the total share of voters.

Corollary 1. *Transfers to county $\hat{c}$ are increasing in the total share of voters, where*

$$\textit{Total share of voters} = \sum_d \frac{v(\hat{c}, d)}{\sum_c v(c, d)}.$$

The expression for the total share of voters makes clear that the amount of transfers to a given county depends on the voting population of every county that shares one of the same districts (from the denominator of each term in the summation). If two counties intersect with the same district, I will refer to them as 'political neighbors' and say that they 'share' a representative. A county has a 'whole' representative when there are no political neighbors for that district.

Political neighbors are key in understanding the equilibrium allocation of transfers. Intuitively, in each district, political competition concentrates in the county that has the highest share of voters. Thus, two identical counties can receive different spending depending on the political neighbors. Similarly, it is not always the case that a county will receive more money by having an additional representative in the legislature. Having more representatives may increase spending in a county because the county's voters influence more elections. On the other hand, as a county gets divided into more and more districts, it has a smaller influence on each election. Whether one effect dominates the other will depend on how a county relates to its political neighbors. This can be illustrated with a two-county example.

3.6 When are two 'half representatives' better than one?

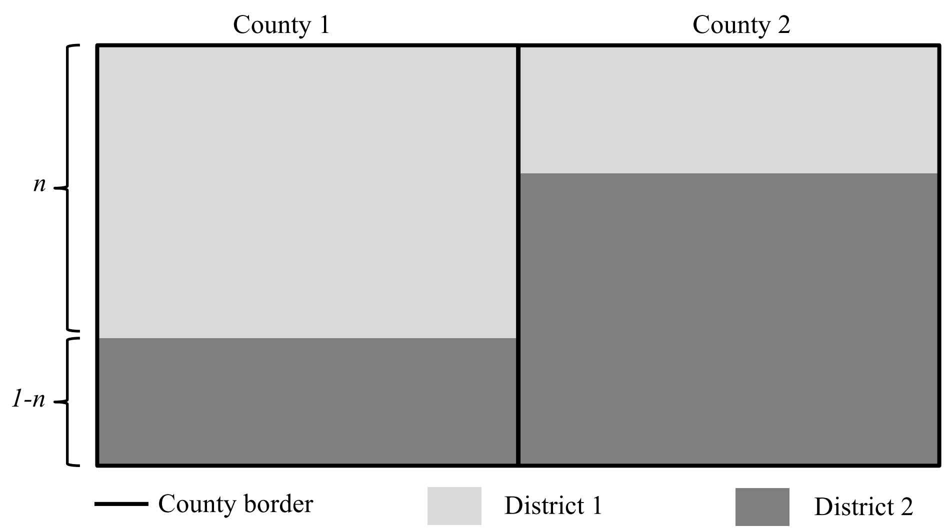
Suppose there are two districts and two counties (See Figure 1). In this setup, the effect of border mismatch can be assessed with a single parameter, the share of county 1 in district 1. Let $n(1,1) = n(2,2) = n$ (then $n(1,2) = n(2,1) = 1-n$). Borders are mismatched if $n \in (0,1)$. The counties are identical except the turnout rate in county 1 is higher than in county 2: $e(1,1)t(1,1) = e(1,2)t(1,2) = \bar{t} > \underline{t} = e(2,1)t(2,1) = e(2,2)t(2,2)$. The first order conditions simplify to:

$$\underbrace{\left(\frac{n\bar{t}}{n\bar{t} + (1-n)\underline{t}} + \frac{(1-n)\bar{t}}{(1-n)\bar{t} + n\underline{t}}\right)}_{\geq 1} \frac{\partial u_1}{\partial q_1} = \frac{\lambda}{\gamma} \tag{3}$$

$$\underbrace{\left(\frac{(1-n)\underline{t}}{n\bar{t} + (1-n)\underline{t}} + \frac{(n)\underline{t}}{(1-n)\bar{t} + n\underline{t}}\right)}_{\leq 1} \frac{\partial u_2}{\partial q_2} = \frac{\lambda}{\gamma}. \tag{4}$$

When borders are congruent ($n = 0$ or $n = 1$), counties 1 and 2 receive the same transfers in equilibrium, as the total share of voters equals one in both counties and the marginal utility in transfers is also the same. With border mismatch, however, spending in county 1 is strictly higher than spending in county 2. County 1 attracts more political competition because it represents a disproportionately large share of voters in each district. Conversely, county 2 is better off with one whole representative. The disparity in spending across counties peaks at $n = 1/2$, when residents of county 1 reach their maximum utility and residents of county 2 reach their minimum utility in the political equilibrium.

Figure 1: Two counties and two districts



Note: Two counties (solid black lines) and two districts (shaded areas), where n is the share of county 1 in district 1 and the share of county 2 in district 2. When $n = 0$ or $n = 1$, transfers are the same for each county, all else equal. When $n \in (0, 1)$, borders are mismatched. Transfers are higher in the relatively high turnout county, all else equal.

In general, a relatively high turnout county is better off ‘cracked’ across many districts rather than in its own district. That way, in each district the county has a larger share of turnout compared to its share of the district population. A relatively low swing county is better off ‘packed’ into fewer districts, or with only whole representatives, so that it can not be ignored for political neighbors. Note that any inequality in transfers in the two county example stems from asymmetry in voting behavior, $v(c, d)$, or preferences, a_c. However, in general, two counties can be identical except for in the way that they are districted and still receive different transfers in equilibrium. That is, two counties with the same population size, preferences, voting eligibility, and turnout rates can have different total shares of voters and thus receive different amounts of transfers in equilibrium.[13] This is a form of horizontal inequality, in that two counties

[13]To see this, simply compare County 1 in the case where $n = 0$ to the same County 1 in the case where

with the same characteristics are treated differently. There is no reason to expect that a social planner would treat identical counties differently. In this sense, horizontal inequality represents a form of inefficiency. Next, I compare the political equilibrium to an efficient benchmark to make this point explicit.

3.7 An efficient benchmark

The political equilibrium is closely related to the solution of a utilitarian welfare maximization problem. A social planner that maximizes the utility of all individuals would weight the marginal utility in q_c by the population of the county, n_c.[14] In the political equilibrium, the marginal utility of q_c is instead weighted by how spending in county c increases the probability of winning an additional seat. Thus, the political equilibrium is efficient in the special cases where

$$\sum_d \frac{v(\hat{c}, d)}{\sum_c v(c, d)} = n_{\hat{c}}.$$

This would be the case if counties and districts were the same *and* if counties were all the same population size. On the other hand, if counties that vary in population size are used as districts, then the model captures the typical concern over malapportionment: small counties would receive more money in the political equilibrium than is socially optimal.[15] In another special case, the political equilibrium is efficient if the expected turnout rate is the same everywhere (i.e., if $e(c, d)v(c, d)$ is equal for all c, d). Then, in each district, a county contributes a share of the turnout equal to its population share. Apart from these special cases, heterogeneous voting behavior together with border mismatch leads to inefficient spending, relative to the utilitarian benchmark. It is also worth noting that horizontal inefficiency arises even in the case where there is no border mismatch, but where counties are nested within districts (as in Bouton, Genicot, and Castanheira 2019).

4 Bringing the Theory to Data

The aim of this section is to bridge the gap between the theoretical model and empirical analysis. Taking the natural logarithm of the first order condition for spending in county $\hat{c}$ (Equation 2) and using the functional form assumption for utility yields a

$n \in (0, 1)$ in the above example (Figure 1).

[14]The social planner solves $\max_{\mathbf{q}_c} \sum_c n_c u_c(q_c)$, subject to $\sum_c q_c \leq y$. Optimal spending in each county c satisfies $n_c \frac{\partial u_c}{\partial q_c} = \lambda$.

[15]If counties were districts, then the first order conditions for the political equilibrium would satisfy $\frac{\partial u_c}{\partial q_c} = \frac{\lambda}{\gamma}$.

deterministic function for the log of transfers to county $\hat{c}$ in state s and year y:

$$\ln(q_{\hat{c}sy}) = \rho \ln\left(\sum_d \frac{v(\hat{c},d)_{sy}}{\sum_c v(c,d)_{sy}}\right) + \alpha(1-\rho)\ln(n_{\hat{c}sy}) + \rho\ln(a_{\hat{c}sy}) - \rho\ln\left(\frac{\lambda_{sy}}{\gamma_{sy}}\right) \quad (5)$$

Equation 5 can be reinterpreted as an empirical specification with the introduction of statistics. Suppose that the intensity in preferences for public consumption, $\ln(a_{\hat{c}sy})$, is a linear function of observable and unobservable characteristics in a county (as in Stromberg 2004a):

$$\ln(a_{\hat{c}sy}) = \beta_a' x_{\hat{c}sy} + \epsilon_{\hat{c}sy}$$

where x_c is a vector of observable variables, and $\epsilon_{\hat{c}sy}$ is drawn from a distribution with mean zero and variance σ_ϵ^2. Plugging this into Equation 5 yields an estimable equation for transfers to a county.

Empirical Prediction 1. *Transfers to county $\hat{c}$ are increasing in the total share of voters across districts and weakly increasing in population size. That is:*

$$\ln(q_{\hat{c}sy}) = \beta_1 \ln\left(\sum_d \frac{v(\hat{c},d)_{sy}}{\sum_c v(c,d)_{sy}}\right) + \beta_2\ln(n_{\hat{c}sy}) + \beta_a' x_{\hat{c}sy} + \theta_{sy} + \epsilon_{\hat{c}sy} \quad (6)$$

Where

$$\beta_1 = \rho \in (0,1)$$
$$\beta_2 = \alpha(1-\rho) \in [0,1)$$

These predictions can be tested in the data by computing the main variable of interest, the total share of voters, and estimating β_1 and β_2. In a second specification, I decompose the total share of voters into two terms: the number of representatives per county and the average share of voters per district. This gives a way to easily compare counties that have the same number of representatives, generalizing the two county example in Section 3.6.

Let $R_{\hat{c}}$ be the number of representatives with constituents in county $\hat{c}$, so that $R_{\hat{c}} = |\{d \mid n(\hat{c},d) > 0\}|$. The total share of voters for county $\hat{c}$ can be rewritten:

$$\begin{aligned}\sum_d \frac{v(\hat{c},d)}{\sum_c v(c,d)} &= \sum_d \mathbb{1}_{v(\hat{c},d)>0}\left[1 + \frac{v(\hat{c},d)}{\sum_c v(c,d)} - 1\right] \\ &= R_{\hat{c}} - \sum_d \psi_{\hat{c}}^d \quad (7)\end{aligned}$$

where $\psi_{\hat{c}}^d = 1 - \frac{v(\hat{c},d)}{\sum_c v(c,d)} > 0$ when $v(\hat{c},d) > 0$ and $\psi_{\hat{c}}^d = 0$ otherwise. The parameter

$\psi_{\hat{c}}^d$ is the share of turnout in district d that resides outside of county $\hat{c}$.[16] This expression helps to reinterpret the total share of voters in terms of the number of whole and shared representatives. In particular, there is a penalty, measured by ψ_c^d for sharing the representative of district d with other counties. The model gives a sufficient statistic, $\sum_d \psi_c^d$, that allows us to compare counties that have the same number of representatives.

Plugging Equation 7 into Equation 6 yields the second empirical specification and set of predictions.

Empirical Prediction 2. *Transfers to county c are increasing in the number of representatives and the average share of voters, by the same magnitude. Transfers are weakly increasing in population size. That is:*

$$\ln(q_{\hat{c}sy}) = \pi_1 \ln(R_{\hat{c}sy}) + \pi_2 \ln(1 - \bar{\psi}_{\hat{c}sy}) + \pi_3 \ln(n_{\hat{c}sy}) + \pi_a' x_{\hat{c}sy} + \theta_{sy} + \epsilon_{\hat{c}sy} \tag{8}$$

Where

$$\bar{\psi}_{\hat{c}sy} = \frac{1}{R_{\hat{c}sy}} \sum_d \psi_{\hat{c}sy}^d$$

$$\pi_1 = \pi_2 = \rho \in (0,1)$$

$$\pi_3 = \alpha(1-\rho) \in [0,1)$$

Decomposing the total share of voters in this way has two additional benefits. First, I introduce a more demanding set of empirical predictions. The model predicts that $\pi_1 = \pi_2$, but the estimation of these coefficients is unrestricted. In contrast, the first specification imposes that $\beta_1 = \pi_1 = \pi_2$. Second, R_c is measured without error, whereas measurement error in $v(c,d)$ may lead to attenuation bias for the coefficients on the total and average share of voters. The identification and estimation of these specifications follow a brief description of the data.

5 Data

To test the predictions of the model, I merge data on intergovernmental transfers, voting eligibility, turnout, political candidate characteristics, and other county covariates. I use data that is as disaggregated as possible to construct variables for the total share of voters and the average share of voters. This is necessary since legislative district borders cross most statistical boundaries used by the U.S. Census. I use block-level voting

[16]For example, in the the two county example of Section 3.6, $\psi_1^1 = 1 - \frac{n\bar{t}}{n\bar{t}+(1-n)\underline{t}} = \frac{(1-n)\underline{t}}{n\bar{t}+(1-n)\underline{t}}$. This is the share of district 1 voters that are in county 2.

eligibility data and precinct-level election returns to avoid the need to disaggregate data to county-district intersections. This limits the dataset to district maps drawn after the 2000 census. The timing of data collection follows the timing in the model: I first observe a district map, then an election, then public spending outcomes (Table 1).[17] A full description of how the data were compiled, including state-specific exceptions and notes, is available in the online appendix. Summary statistics are reported in Appendix B.

Table 1: Timing in the model compared to timing of data collection

Time period	Data element	Years of data collection
T=0	District map	2001 and 2006
T=1	Policy Proposals	(unobserved)
T=2	Voting	2004 and 2010
T=3	Public spending	2006-07 and 2011-12

Data for intergovernmental transfers comes from the Census of Governments, which collects public finance data every five years. Ideally, I would be able to identify all transfers that go from the state government to the geographic area defined by county borders. A simple starting point would be to use transfers from state governments to county governments only. This would be misleading, however, since counties vary in their responsibilities and often delegate decisions to school district and special purpose governments (for services such as fire, utilities, transportation, and others).[18] Counties have a median of 6 special purpose governments, 2 school districts, and 5 municipalities. Transfers to these governments within counties make up an important set of transfers to county areas. I therefore construct q_c as the sum of intergovernmental transfers from the state government to the county government and to all other governments within that county. By this measure, transfers from the state account for a significant share (36%) of total revenue in a county. The majority of these transfers (75%) are for education purposes. A downside of this measurement of q_c is that it includes some transfers that don't fit the model well because they target geographical areas smaller than a county. I return to this issue in Section 6.3.

[17]For example, when states report public spending outcomes for the 2006-07 fiscal year, the budget was determined by officials who were elected in 2004, in most states. The relevant district maps were drawn following the 2000 census in 2001. However, some states changed their district maps in between redistricting years and elections, and some states have elections in odd years. See the Data Appendix for more details.

[18]The Census of Government assigns each special government to a unique county. However, it is difficult to gauge the extent of border mismatch between special governments and counties. The Census Bureau states, "Governments that cross county boundaries are enumerated in a single county" (Census of Governments, 2012 Organization Table 13). This data limitation also makes it difficult to study transfers to towns or municipalities, which use special governments but are not linked to them in the Census of Governments.

To determine county and district borders, I use data collected by the Census Redistricting Program. State legislative districts must be redrawn every ten years. Before and after the redistricting process, states report to the Census Redistricting Program how census blocks are assigned to state legislative districts.[19] Blocks are the smallest geographic entity used by the Census Bureau, and they are rarely split across multiple districts (only 23 blocks out of over 11 million spanned multiple districts). Each block is also assigned to a unique county so that I can aggregate block-level Census data to the county-district level. In the main analyses, I focus on districts for the lower houses of state legislatures.[20]

I use two different measures of the voting population, $v(c, d)$, to compute the total share of voters and the average share of voters in each county. First, I use variation in the voting eligible population. This is equivalent to assuming that turnout is constant across counties in expectation; any variation in turnout in the data is only noise. The voting age population (VAP, population over age 18) from the U.S. Census serves as a proxy for voting eligibility and is available at the block-level.[21] For the second measure, I use turnout to measure the number of voters in a county-district. Here, the assumption is that variation in turnout is informative to political parties. I measure turnout using precinct-level state legislative election results (Ansolabehere, Palmer, and Lee, 2014). I include all 21 states that report results for state legislative elections in the relevant years and that have a county identifier.[22] Some precincts are split across state legislative districts. However, since votes are reported for each district election, turnout can be assigned to county-district intersections even in these cases.[23]

The two ways of measuring the share of voters are complementary. An advantage

[19] I use information from the 2000 Redistricting Program Phase 3 (collected post-2000 Census in 2001) and from the 2010 Redistricting Program Phase 1 (collected pre-2010 Census in 2004-2006). The block assignments represent the legislative districts that were used in the most recent elections. Participation in the Census Redistricting Program is voluntary. In 2000, the following states did not participate: AK, CA, FL, HI, KY, ME, MD, MN, MT, NE, NH, TX.

[20] All states except for Nebraska have a bicameral legislature. I focus on the lower house for two reasons. First, in some states, the lower house districts are nested within upper house districts. The total share of voters in the lower house is the relevant measure in these states. Second, some states use staggered elections for the upper house of the state legislature. These states don't fit the assumptions of the model well and also have more limited turnout data. When I incorporate the upper house into the empirical analysis I find that results are qualitatively the same as the main analyses to follow (Online Appendix).

[21] VAP is a common proxy for voting eligibility. However, VAP differs from the voting eligible population for various reasons, depending on the state (e.g., VAP includes non-citizens and persons with felony convictions).

[22] They are: AK, AL, CA, DE, FL, IA, KS, LA, MN, MO, NJ, NC, OH, OK, OR, PA, SC, TN, TX, WI, and WY. Merging these data required revising or adding new information about districts (See Data Appendix). After this cleaning, 99.2% of county-districts from the precinct data merge with the county-district level dataset that I construct from Census data.

[23] This would not be possible, for instance, if we want to assign turnout to intersections of state legislative districts and school districts, towns, or municipalities. Although the geographical overlap is known, assigning turnout to these smaller governments would require disaggregating turnout data. Disaggregation would also be required to use turnout data for other elections (e.g., governor, president).

of using VAP is that the data are available for all states. The share of the population that is under 18 is also not easy for politicians to manipulate within an election cycle. On the other hand, the assumption that turnout is uninformative to political parties is unsatisfactory. However, using turnout data limits the analysis to the subset of states with precinct-level election returns for state legislative elections. Endogeneity is also a greater concern in this case, since political parties may try to mobilize turnout. Given the tradeoffs between identifying assumptions, endogeneity concerns, and data availability, I present results using VAP and turnout side by side.

Finally, I include additional covariates of public spending from multiple sources. At the county-level, I include demographic covariates from the American Community Survey 5-year samples, centered around 2007 and 2012. I treat county-level VAP as a covariate of public spending because the school-age population (not eligible to vote) is an important determinant of education spending. However, I do not include county-level turnout as a separate regressor, because it is a political variable and the model implies how it should enter the regression. I add political covariates using a candidate-level dataset (Klarner et al., 2013) and information on who draws legislative districts, whether or not there was a court-drawn district, and whether or not a county required pre-clearance under Section 5 of the Voting Rights Act (Levitt, 2010).[24] This allows me to include political covariates that are not in the model but may correlate with voting behavior and transfers: incumbency status and party affiliation of representatives in a county, and alignment of a county's representatives with the party in control of the legislature.

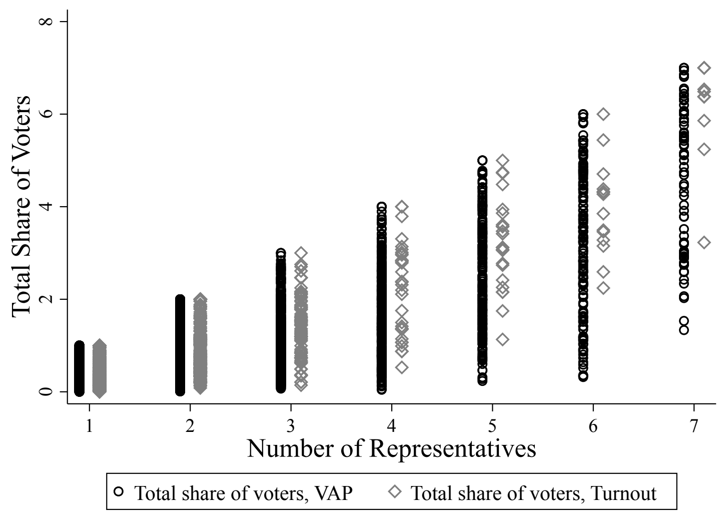
6 Empirical Results

The data confirm that border mismatch between counties and state legislative districts is ubiquitous in the United States: 64% of counties in the sample are split across two or more districts. This is in spite of the fact that most counties (68%) are smaller in population than the size of a district for the lower house of the legislature in their state. Even among these smaller counties, 49% are split across multiple districts. In the whole sample, 52% percent of counties have more representatives than their population size would imply. As a result, there is significant variation in the total share of votes for counties with the same number of representatives, whether measured using VAP or turnout (Figure 2). For instance, some counties have two representatives because they are approximately twice the size of a district (total share of voters equals 2). Other counties are in two districts but represent a very small share of voters in each

[24] For those counties and municipalities covered by Section 5 of the Voting Rights Act of 1965, states had to receive clearance from the U.S. Department of Justice before making any districting changes.

district (total share of voters is close to 0). This variation is necessary to identify the coefficients for the number of representatives and the average share of voters in a county.

Figure 2: Total share of voters by number of representatives



Note: This figure plots the total share of voters for each county, measured using VAP (circles) or turnout (diamonds), by the number of representatives per county. This plot excludes counties in the the top 5% of number of representatives.

Table 2 reports OLS estimates of Equations 6 and 8. In all regressions, I include state-year fixed effects.[25] I also adjust standard errors for clustering at the state-year level, since assignment of the number of representatives and political neighbors is determined by state-level redistricting in each year.[26] In line with Prediction 1, the OLS estimate for the total share of voters is positive and statistically significant, whether measured by VAP (Column 1, $\hat{\beta}_1 = 0.25$, SE= 0.06) or turnout (Column 3, $\hat{\beta}_1 = 0.17$, SE= 0.07). In line with Prediction 2, the coefficients for the number of representatives and average share of voters are positive, statistically significant, and similar in magnitude using VAP (Column 2, $\hat{\pi}_1 = 0.27$, SE= 0.06 and $\hat{\pi}_2 = 0.24$, SE= 0.06) and turnout (Column 4, $\hat{\pi}_1 = 0.21$, SE= 0.07 and $\hat{\pi}_2 = 0.16$, SE= 0.07). Although

[25] I also estimate Equations 6 and 8 using county fixed effects in the Online Appendix. Most states use the same district map for the entire time period of the sample. However, even for these states there is some within-county variation over time due to changes in the age composition (share of VAP) and voting behavior (share of turnout). Using VAP to measure voters, the estimated coefficients are statistically significant and imply a value of ρ between 0.12 and 0.13. Using turnout to measure voters, the sample size is smaller and the estimates of ρ are not statistically significantly different from zero.

[26] I choose to allow for clustering at the state-year level in the main specifications since the problem is modeled at the state-year level. However, there could be correlations over time within a state. If I allow for clustering at the state-level instead, standard errors increase only slightly: for both the coefficients on total and average share of voters, SE=0.06 in columns 1 and 2 and SE=0.07 in columns 3 and 4 of Table 2.

the coefficients for the average share of voters and the number of representatives are similar, the coefficient for the number of representatives is statistically significantly larger. This is likely due to the fact that the number of representatives is measured without error while VAP and turnout are measured with error.

The estimated coefficient on the log of population is positive and statistically significant in all four regressions, ranging from 0.68 (SE= 0.06) to 0.73 (SE=0.06). These point estimates imply a value of α between 0.87 and 0.91, suggesting that the transfers are closer in nature to private transfers than local public goods. The estimated coefficients for log of population are consistent with the prior that a portion of intergovernmental transfers are allocated based on population. However, the fact that the total share of voters helps to explain variation in intergovernmental transfers suggests that politicians also have some degree of discretion over transfers.[27]

The estimated coefficients for the number of representatives, average share of voters, and total share of voters are fairly stable across both specifications and both ways of measuring vote shares. This is consistent with the theoretical model, since the three coefficients all equal the same underlying parameter, ρ. The parameter ρ captures the intensity of preferences for public consumption and determines the magnitude of the effect of border mismatch. For instance, a value of $\rho = 0.15$ ($\hat{\beta}_1$ from Column 3) implies that if the total share of voters increases by 100%, then transfers increase by 15%. Although a 100% increase in the total share of voters is an unlikely event, large swings in the total share of voters are feasible before and after redistricting. To get a sense of the magnitude of changes to transfers under alternative district maps, I do several counterfactual exercises in Section 6.2.

Next, I use a placebo test as a first step in ruling out the possibility that the OLS results are driven by county-level unobservables. For the placebo measure, I construct the average share of voters in a county, relative to all neighboring counties that are *not* in any of the same political districts (i.e., 'non-political neighbors'). By pairing counties with geographical neighbors in different districts, this placebo test is essentially a way to construct counterfactual districts for each county. In Table 3, I show that the average share of voters relative to 'non-political neighbors' are uncorrelated with state-level transfers (using either VAP or turnout to measure voters). This suggests the OLS results are driven by existing district borders rather than general geographical patterns of covariates related to both voting and spending.

To further demonstrate the importance of district borders for county outcomes, I use only county-level regressors in Table 4. Consistent with other studies of representation (Ansolabehere, Snyder, and Ting 2003, Hirano 2006 and 2011, Knight 2008, Fiva

[27]In fact, a standard deviation change in the total share of voters is associated with a change in transfers that is similar in magnitude to a standard deviation change in other factors that could affect formulaic spending like median household income (see online appendix for full set of coefficient estimates).

Table 2: OLS estimates using either Voting Age Population (VAP) or turnout to measure total share of voters: y=log of intergovernmental transfers from state to county

	(1)	(2)	(3)	(4)
ln(Tot. share of voters, VAP)	0.26***			
	(0.05)			
ln(Avg. share of voters, VAP)		0.25***		
		(0.05)		
ln(Tot. share of voters, Turnout)			0.15**	
			(0.07)	
ln(Avg. share of voters, Turnout)				0.14**
				(0.06)
ln(Number of representatives)		0.28***		0.21***
		(0.05)		(0.07)
ln(Population)	0.67***	0.67***	0.75***	0.74***
	(0.05)	(0.05)	(0.07)	(0.06)
Percent incumbent representatives	0.04**	0.04**	0.08***	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)
Percent Democrat representatives	0.05***	0.05***	0.08**	0.07**
	(0.01)	(0.01)	(0.03)	(0.03)
Percent Dem. reps × Dem. control	-0.01	-0.01	-0.06	-0.04
	(0.03)	(0.03)	(0.04)	(0.04)
N	5401	5401	1203	1203
Dependent variable mean	17.58	17.58	17.52	17.52
Adjusted R^2	0.959	0.959	0.968	0.969
p-value for $H_0 : \pi_1 = \pi_2$		0.013		0.008

Note: Standard errors clustered at the state-year level in parentheses. All regressions include state-year fixed effects and the following controls: log of population, percent urban, poverty rate, median household income, percent black, percent hispanic, percent female, percent unemployed, percent without a high school diploma, and an indicator for whether or not the county is covered by the Voting Rights Act Section 5.

Table 3: OLS estimates of placebo test (using non-political neighbors to measure share of voters): y=log of intergovernmental transfers from state to county

	(1)	(2)
ln(Avg. share of voters, among non-political neighbors, VAP)	-0.008	
	(0.007)	
ln(Avg. share of voters, among non-political neighbors, Turnout)		0.004
		(0.007)
ln(Number of representatives)	0.048***	0.079***
	(0.014)	(0.020)
ln(Population)	0.895***	0.871***
	(0.014)	(0.016)
Percent incumbent representatives	0.035*	0.050*
	(0.019)	(0.026)
Percent Democrat representatives	0.053***	0.083***
	(0.016)	(0.023)
Percent Dem. reps × Dem. control	-0.016	-0.033
	(0.027)	(0.034)
N	4978	1754
Dependent variable mean	17.62	17.77
Adjusted R^2	0.959	0.972

Note: 'Non-political neighbors' refer to any counties that are geographically adjacent to the county of interest, but that are not in any of the same districts. Standard errors clustered at the state-year level in parentheses. All regressions include state-year fixed effects and the following controls: percent voting age population, percent urban, poverty rate, median household income, percent black, percent hispanic, percent female, percent unemployed, percent without a high school diploma, and an indicator for whether or not the county is covered by the Voting Rights Act Section 5.

and Halse 2016, Hirano and Ting 2015, Yuan 2018), there is a positive and statistically significant relationship between the number of representatives and transfers, though much smaller than in Table 2. The relationship between turnout and VAP is negative and statistically significant. This is because a large VAP implies a small school-age population, and a majority of transfers are spent on education. Finally, there is an economically and statistically insignificant relationship between transfers and turnout. However, in light of the theoretical model, the estimates in Table 4 are misleading. In particular, the regressions suffer from omitted variable bias because they do not take the average share of voters into account. Importantly, there is only a weak correlation between county-level turnout and the average share of turnout (the coefficient of correlation is 0.07). This exercise demonstrates the importance of measuring local governments *relative* to electoral districts when studying political determinants of public spending.

So far, the OLS results are consistent with predictions of the model, yield plausible parameter values, and are stable across different specifications. There is also reason to believe that OLS estimates are robust to some of the typical threats to identification in this setting. For instance, county-level turnout may be related to unobservable factors that influence transfers to that county. However, this does not imply omitted variable bias in this setting. Rather, an unobservable variable causes bias only if it is correlated to the total *share* of turnout as well as transfers. To make this concrete, consider a county that has a high turnout rate. The county's healthy political participation may correspond with a greater ability to organize as a community and lobby the state government for the county's needs. This would create biased estimates of the effects of county-level turnout on transfers. However, the turnout in neighboring counties is unlikely to be correlated with the given county's ability to lobby. Then, the total share of voters is also unlikely to be correlated with the proposed omitted variable.

Given the weak correlation between the total share of turnout and county-level turnout as well as the results of the placebo test, it is difficult to think of county-level unobservables that would reverse the OLS results. However, the fact that district maps are not randomly drawn does create some concern over endogeneity. For instance, gerrymandering may lead to systematic variation in the total share of voters depending on the partisanship of a county.[28] A variable that is correlated with both partisanship and transfers could cause the OLS estimates to be biased. Many states also have provisions to keep communities of interest together in the same districts. A community

[28] The theoretical relationship between partisanship and the total share of voters is unclear. For example, suppose a party in control of drawing the map chooses to put their opponent's voters into as few districts as possible. This is a well-known gerrymandering tactic called packing. If the opponent's voters are a high turnout group, then the total vote share of the group decreases due to packing. However, if the opponent's voters are a relatively low turnout group, then the total vote share of the group increases due to packing.

Table 4: OLS estimates with county-level information only: y=log of intergovernmental transfers from state to county

	(1)	(2)	(3)	(4)
ln(Number of representatives)	0.07**			0.07***
	(0.03)			(0.02)
ln(Voting age population)		-2.37***		-2.46***
		(0.31)		(0.34)
ln(Turnout)			-0.03	0.04
			(0.03)	(0.04)
ln(Population)	0.88***	3.27***	0.93***	3.29***
	(0.01)	(0.30)	(0.03)	(0.33)
Percent incumbent representatives	0.06**	0.07***	0.06**	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)
Percent Democrat representatives	0.05**	0.09**	0.06**	0.09**
	(0.02)	(0.04)	(0.02)	(0.03)
Percent Dem. reps × Dem. control	-0.03	-0.06	-0.05	-0.05
	(0.03)	(0.04)	(0.03)	(0.04)
N	1203	1203	1203	1203
Dependent variable mean	17.52	17.52	17.52	17.52
Adjusted R^2	0.965	0.968	0.965	0.968

Note: Standard errors clustered at the state-year level in parentheses. All regressions include state-year fixed effects and the following controls: percent urban, poverty rate, median household income, percent black, percent hispanic, percent female, percent unemployed, percent without a high school diploma, and an indicator for whether or not the county is covered by the Voting Rights Act Section 5.

of interest has shared cultural or economic interests but is only vaguely defined and therefore difficult to identify in the data (Levitt, 2010). If communities of interest also benefit from specific state policies, then the presence of communities of interest would be another source of bias. To address these concerns, I introduce an instrument to exploit plausibly exogenous variation in the way that a county is districted.

6.1 Instrumental Variables Approach

I construct an instrument for the number of representatives per county based on the fact that districts must have roughly the same population size within a state.[29] The ideal district population, n_d, is equal to the total state population divided by the number of seats. If the number of representatives were determined by population size alone, then a county would receive $int(n_c/n_d)$ whole representatives (where $int(.)$ gives the integer value), and an additional shared representative for any remainder of the county population. Formally, we have the following population-based rule for the total number of representatives:

$$r_c = int\left(\frac{n_c}{n_d}\right) + \mathbb{1}_{\{int\left(\frac{n_c}{n_d}\right) \neq \frac{n_c}{n_d}\}}$$

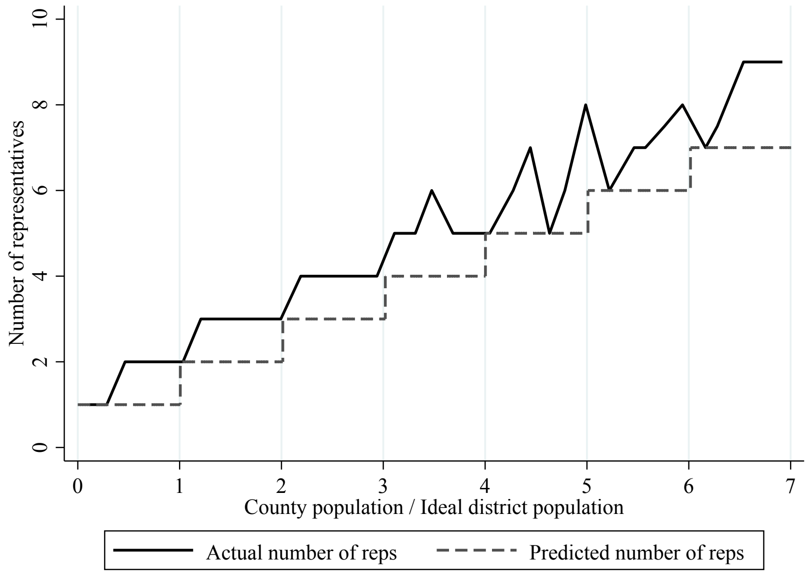
In practice, many constraints complicate the redistricting problem such that $r_c(.)$ cannot be implemented. However, identification requires only that the population-based rule has some predictive power over the actual number of representatives, and that it does not have an independent effect on spending. Figure 3 plots the rule r_c against the observed number of representatives in the data. The predicted and actual number of representatives are highly correlated (with a coefficient of correlation 0.92), though the actual number of representatives is systematically higher than r_c. Importantly, the number of representatives in the data also has discontinuities near integer multiples of the ideal district population size. This is especially true for counties with three or fewer representatives, which accounts for 78% of observations. Since population and state-year fixed effects are already included as regressors, the discontinuities create the variation used to identify the coefficient for the number of representatives. In some specifications, I also instrument for the average share of turnout using previous state legislative election data to construct a lagged value of the average share of turnout.[30] This significantly reduces sample size, due to data availability.

I report IV estimates in Table 5. In Column 1, the number representatives is

[29] In states with multi-member districts, districts must have roughly the same number of representatives per person. See Online Appendix for the theoretical extension to multimember districts.

[30] Lagged variables have been used to instrument for turnout (Stromberg 2004a), but there is some concern over their validity. For instance, if previous turnout had an independent effect on the budget, and the previous budget decisions have some persistence, then the instrument is not valid.

Figure 3: Number of representatives: Predicted vs. actual



Note: The dashed line shows the predicted number of representatives according to county population alone. The solid line shows the median band plot for the number of representatives in the data. The plot excludes counties in the top 5% of population as percent of the ideal district population.

treated as endogenous in an estimation of Equation 8. The coefficients for the average share of voters and number of representatives are positive, statistically significant, and larger in magnitude than OLS estimates. These estimates are consistent if the average share of voters, measured using turnout, is exogenous. This could be the case even if the number of representatives is endogenous. For instance, gerrymanderers may pay attention to the number of districts that a county is split into more so than the relative turnout rates of the potential political neighbors surrounding the district. Indeed, in Column 2, I treat the average share of voters as endogenous and fail to reject that the OLS coefficient on the average share of turnout is consistent (the Durbin-Wu-Hausman test p-value$= 0.24$). In Column 3, I treat the total share of voters as endogenous and use both instrumental variables in an over-identified estimation of Equation 6. The coefficient for the total share of voters again suggests a larger effect than implied by OLS estimates.[31]

It is important to note that IV estimates use a different source of variation to

[31] The F-statistics (corrected for clustering) are large for Columns 1 and 2 but relatively small for Column 3. The IV coefficient in column 3 may have up to 25% of the bias of the OLS coefficient, rather than the standard 10% threshold (Stock and Yogo 2005). The Hansen J Statistic also rejects the null hypothesis that the over-identifying restrictions hold. However, this test does not rule out the case that the instruments are valid but incoherent (Parente and Silva, 2012). I rule out the possibility that that the IV results are driven by any single state-year cluster in a leave-one-out cross validation exercise (see Online Appendix).

estimate a local effect. In particular, the IV estimates in Table 5, Columns 1 and 3 rely on variation among counties close to a discontinuity in the ideal number of representatives. These counties are larger and more urban than the overall sample. In Columns 4 and 5 of Table 5, I show that the OLS point estimates are very similar to IV estimates when using a subsample of counties with population size close to a discontinuity in the instrument r_c.[32] Thus, the differences in IV and OLS point estimates likely stem from heterogeneous effects rather than biased OLS estimates. To interpret the magnitude of the effects of border mismatch, I use the OLS point estimates in Column 3 ($\hat{\rho} = 0.15$), both because this is a more conservative approach and because of evidence that the local effect estimated with instrumental variables is larger than the average effect.

6.2 Interpreting the coefficients

To interpret the economic significance of the empirical results, it helps to interpret coefficients through examples. Consider what happens to a county with one whole representative that gets split evenly into two districts under a new district map. Suppose that the county of interest has the sample median turnout rate of 30%. If the political neighbors under the new map have the 25th percentile turnout rate of 20%, then the share of total voters would increase by 20%. If we take the parameter value of $\rho = 0.15$, then transfers would increase by 3.2%, or $43 per person.[33] If neighboring counties have the 75th percentile turnout rate of 40%, then the total share of voters would decrease by 15%, so per capita spending would decrease by 2.5%, or $33 per person. The large differences in turnout rates in these examples are not unusual for geographically neighboring counties in the dataset. In fact, 47% of counties are adjacent to at least one other county in a different quantile for turnout.

In a second back of the envelope calculation, I consider what would happen to the distribution of transfers if there were no political neighbors. Due to the uneven geographical distribution of turnout rates, a majority of counties (80%) have a lower turnout rate than their political neighbors. These counties are less densely populated than the average county, accounting for 62% of the population. This suggests that most individuals would benefit from having fewer political neighbors. To see this, I compute the change in spending if each county had their own representatives, and with a number of representatives equal to the nearest integer multiple of an ideal district population size: $\bar{\psi}_c = 0$, $R_c = int(n_c/n_d + 0.5)$, and $\rho = 0.15$. This eliminates the distortionary

[32] I compute the deviation from the nearest integer multiple of an ideal district size and include counties in the bottom quartile of this measure (they are within 28% deviation of an ideal district size).

[33] Recall that the value of the total share of voters equals one for any county with its own representative. Under the new map, the county has a total share of voters equal to $2 * 0.3/(0.3 + 0.2) = 1.2$. The total share of voters increases from 1 to 1.2, or by 20%. Median per capita spending is $1,352.

Table 5: IV estimates: y=log of intergovernmental transfers from state to county

	IV			OLS	
	(1)	(2)	(3)	(4)	(5)
ln(Avg. share of voters, Turnout)	0.45***	0.35**		0.31	
	(0.11)	(0.15)		(0.26)	
ln(Number of representatives)	0.61***	0.42**		0.40	
	(0.13)	(0.15)		(0.30)	
ln(Tot. share of voters, Turnout)			0.49***		0.29
			(0.12)		(0.25)
ln(Population)	0.42***	0.54***	0.43***	0.55	0.64**
	(0.11)	(0.13)	(0.12)	(0.30)	(0.27)
Percent incumbent representatives	0.08***	0.09**	0.10**	0.09*	0.10*
	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)
Percent Democrat representatives	0.04	0.03	0.03	0.08	0.08
	(0.03)	(0.04)	(0.05)	(0.07)	(0.07)
Percent Dem. reps × Dem. control	-0.01	0.04	-0.01	-0.02	-0.10
	(0.04)	(0.04)	(0.05)	(0.17)	(0.12)
Instrument for Number of reps	X		X		
Instrument for Turnout		X	X		
N	1203	545	545	138	138
Adjusted R^2	0.835	0.849	0.845	0.963	0.963
p-value for $H_0 : \pi_1 = \pi_2$	0.000	0.065		0.204	
First stage F statistic	22.713	32.911	9.255		
Durbin-Wu-Hausman test: p-value	0.000	0.239	0.206		
Hansen J Statistic: p-value			0.020		

Note: Standard errors clustered at the state-year level in parentheses. All regressions include state-year fixed effects and the following controls: percent voting age population, percent urban, poverty rate, median household income, percent black, percent hispanic, percent female, percent unemployed, percent without a high school diploma, and an indicator for whether or not the county is covered by the Voting Rights Act Section 5.

effects of border mismatch, while also limiting malapportionment to some degree.[34] Under this counterfactual redistricting, the median county would have a 9% increase in per capita spending, equal to \$126. The total money reallocated across states would amount to \$11.8 billion dollars in 2012, or 2% of all state to local government transfers.

6.3 Extensions and Robustness Checks

Sub-sample analyses. Another way to limit concern that gerrymandering causes biased OLS estimates is to consider several sub-samples of the data (Appendix C). First, I limit the sample to counties in states where an independent commission or court drew the district map.[35] I also run the regressions separately for the 2012 fiscal year. This was the last fiscal year in which representatives were elected using the maps that followed the 2000 census. The longer time between redistricting and budgeting may limit the possible bias created by redistricting and gerrymandering. I estimate Equation 8 for these sub-samples and in other robustness checks since the coefficient on the number of representatives can be estimated more precisely in smaller samples. The point estimates for ρ in these sub-samples are similar to those in Table 2 and therefore do not provide any evidence of large OLS bias due to redistricting strategies.

I also restrict the sample to include only counties that are in at least one district with a close margin of victory (within 30%, Appendix C). These counties are more likely to satisfy the assumption that their districts are contestable. In this sub-sample, point estimates of ρ are larger than in Table 2 when using VAP to measure the average share if voters (the estimates using turnout are imprecise). This is consistent with the idea that the share of voters includes some districts that are incontestable and thus would not attract political competition. In this sense, the estimates in Table 2 likely represent a lower bound.

Alternative ways to measure a county's share of voters. The literature suggests several additional factors that may affect targeted spending: the extent to which voters are informed about the election, the variability of partisan preferences (swing voters), and the degree of partisan bias (core voters). Variation in election-relevant information and swingness can be easily incorporated into the probabilistic voting framework (see Online Appendix). The theoretical intuition is only slightly changed: political competition concentrates in areas with a high total share of *informed* or *swing* voters. These predictions can be tested with some additional data. I measure

[34]Under this plan, fifteen percent of state legislative districts in the sample would have population size that U.S. Courts would consider major deviations ($>10\%$) from the district population (Cox v. Larios, 542 U.S. 947 (2004)). Note that this counterfactual redistricting could require a different number of representatives in the state legislature.

[35]This includes 20 states: AK, AZ, CA, CO, FL, GA, ID, IA, ME, MD, MN, MT, NH, NM, SC, SD, TX, WI, WA.

information using two proxies: the percent of the population with a bachelor's degree or higher and the overlap of districts with media markets (as in Snyder and Stromberg 2010). I use a measure of swingness that is common in the literature: the standard deviation in Democratic votes in previous elections (Ansolabehere and Snyder 2006, Wright 1974).[36] I multiply each of these proxies by turnout at the county-district level and construct the measures of the average share of educated, informed, and swing voters.

Next, I consider the possibility that political parties target core voters with spending. While incentives to target core voters are more difficult to incorporate in a tractable probabilistic voting model (see Dixit and Londregan 1996), it may be that a county's total share of voters *for the winning party* influences spending.[37] To test this possibility, I construct a measure of a county's share of votes for the winning party in each district.

I estimate Equation 8 using each of these alternative measures for the average share of voters in the Online Appendix. Standard errors are large for the coefficients for the average share of educated, informed, and swing voters. However, the coefficients for the number of representatives are statistically significant, ranging from 0.09 to 0.14. This suggests that variation in information and swingness may play a role in determining transfers, but that they don't add much explanatory power on top of voting eligibility and turnout. There is also little evidence that support for the winning party is an important factor. There is no statistically significant relationship between the average share of voters for the winning party and transfers once the average share of turnout has been accounted for.

Alternative assumptions about targeting. A critical assumption in the model is that state legislatures make spending decisions at the county level. There are two potential issues with this assumption in the context of local governance in the United States. First, it is not obvious that legislatures have sufficient discretion to target counties. Second, it may be that legislatures do have the discretion to target resources, but to different levels of governments other than counties. I approach these issues one at a time.

The discretion to target funds may be limited by the use of formulas in state bud-

[36] The term swingness is sometimes used to refer to the amount of voters that self-report as independent from any political party. In this context, however, swingness refers to the share of voters that are on the margin of choosing one party or another. These marginal voters could conceptually have strong preferences for either party. Therefore, the variation in vote shares better captures the concept of swingness in the theoretical model.

[37] In Table 2, there is no strong evidence that money flows towards counties which supported the winning party (from the coefficient on the interaction of the percent of representatives that are democrats and an indicator for if democrats have a majority of seats). However, this is different from testing whether the total share of winning votes influences transfers.

gets. Indeed, there is evidence at the federal level that formulaic spending constrains the ability of Congress to target spending (Martin 2017). On the other hand, even when part of a state's budget is determined by formulaic spending, the budget is disbursed through transfers to local governments. For example, politicians determine school funding formulas knowing how students are distributed across the state and knowing that the education budget will be transferred to local school districts. Still, if formulaic spending were based on very few dimensions, it could leave little room for targeting counties once the formulas are in place. I explore this possibility further by estimating OLS coefficients separately for different types of transfers (Appendix D).[38] I find that the main results are driven by transfers for non-education purposes, in line with the fact that education spending is determined largely by formulas (Chingos and Blagg, 2017). Overall, empirical evidence is consistent with some degree of discretion over county-level spending, though this assumption may not be appropriate for each type of transfer.

Next, legislatures likely have the ability to target spending at a finer level than counties. In the model, we can easily consider the case where political parties have access to more precise targeting, even at the household level. However, as long as political parties must spend some part of the budget on counties, then the total share of voters in a county will still be a determinant of transfers.[39] Thus, county-district mismatch is still relevant even though legislatures may have access to more precise targeting. The more problematic issue is if legislatures target a different local government. In particular, if states make transfers to school districts and school districts differ from counties, then the total share of voters in a school district would be a determinant of public spending. This would create an omitted variable problem if the share of voters in a school district is correlated with the share of voters in the county. School district border mismatch is the main concern, since the largest share of transfers to local governments is for education spending. States fall into one of two cases when it comes to school districts. In one case, all or almost all school districts are also counties. In the other case, there are typically multiple school districts per county. I identify counties in states where the median number of school districts per county is equal to one and compare them to all other counties.[40]

In states where school districts and counties generally do not coincide, the predic-

[38]Unlike aggregate transfers from states to counties, transfers for specific purposes may equal zero in the sample. I take the inverse hyperbolic sine of transfers to construct the outcome variable.

[39]To see this, consider an exogenous budget y for county-level transfers and an exogenous budget h for household-level transfers. If a party could use the whole budget for household level transfers they would be better off, but since they must allocate y to counties, they will do so to satisfy the same first order conditions in the baseline model.

[40]The states where there is typically one school district per county are: FL, GA, KY, LA, MD, NV, NC, RI, SC, TN, UT, VA, WV (Census of Governments 2012 Organization Tables).

tions hold for non-education transfers only (Appendix E). In contrast, in states where counties generally do coincide with school districts, the predictions hold for both education and non-education transfers (point estimates are statistically significant when using VAP, but noisy when using turnout to measures the average share of voters). In summary, the extent to which county borders matter over other types of borders likely varies from state to state.[41] However, the data are consistent with a binding county-level targeting constraint. In the next section, I discuss policy implications. From that standpoint, a focus on county borders makes sense since it is easier to avoid border mismatch between smaller local governments and districts than it is to avoid border mismatch between counties and districts.

7 Implications for Redistricting

The model and empirical analyses together imply that border mismatch distorts public spending, hurting some individuals and benefiting others. What do these results imply for redistricting? An immediate implication is that splitting counties into multiple districts should be done only when necessary in order to limit inequality amongst similar counties.

Eliminating border mismatch completely, however, is not possible unless districts are allowed to vary in population size. Currently, federal guidelines limit the maximum total percent deviation from the ideal district population size to 10%, and some states place stronger constraints on deviations in district population size (Levitt, 2010). These tight restrictions stem from the 'one person one vote' rationale and they are further backed by evidence that malapportionment matters for public spending. Ansolabehere, Gerber and Snyder (2002) demonstrate that malapportionment prior to redistricting reform in the 1960s led to an unequal distribution of spending per person across counties.[42] In particular, they find that a 100% increase in the number of state representatives *per person* leads to a 20% increase in transfers from the state to a county.[43] Importantly, the effects of border mismatch are of the same order of magnitude: a 100% increase in representation *per county*, measured by the total share of voters, leads to a 15% increase in transfers, by the more conservative OLS estimates in Table 2. Thus, a social planner who wants to ensure equal representation must

[41]In the Online Appendix, I also estimate results separately for each region of the United States. The average results in Table 2 are driven by counties in the South and West, where county governments typically have more responsibilities than counties in the Northeast and Midwest. However, it is difficult to draw conclusions since there is large variation in the sample size across regions.

[42]The redistricting reforms could also serve as a source of variation in total share of voters. However, the disaggregated data necessary to construct the total share of voters is not available for these years.

[43]Elis, Malhotra and Meredith (2009) study the effect of malapportionment for federal outlays to states. They find that outlays to states increase by roughly 7% if representatives per person increase by 100%.

weigh the effects of malapportionment against the effects of border mismatch. In future rounds of redistricting, relaxing the constraint on malapportionment could prevent border mismatch and limit inefficient spending. The optimal amounts of malapportionment and border mismatch will ultimately depend on a government's preferences over vertical inequality (unequal treatment of counties depending on population size) and horizontal inequality (unequal treatment of similar counties).

The results also speak to partisan gerrymandering. In particular, I find suggestive evidence that border mismatch is more common in states that are more vulnerable to gerrymandering or with maps that exhibit greater degrees gerrymandering. To do so, I regress the number of representatives per county on variables that capture the population-based redistricting constraints (the ideal number of representatives based on population, the total number of counties in a state, and the total number of districts in a state) as well as an indicator for whether or not the map was drawn by an independent commission (instead of a state legislature or political commission) and a state-wide measure of gerrymandering called the efficiency gap (Appendix F).[44] I find that, among counties that should have the same number of representatives based on their population size, those counties in states with independent commissions are split into fewer districts.[45] Gerrymandering, as measured by the efficiency gap, is also associated with a greater number of districts per county. These results suggest that partisan control over redistricting exacerbates inequality among similar counties by increasing the frequency of border mismatch.

8 Conclusion

This paper shows that border mismatch between electoral districts and local governments distorts the distribution of public funds. I characterize the effects of border mismatch on public spending with a probabilistic voting model in which parties compete to win legislative districts but can only allocate resources at the county level. Coun-

[44]The efficiency gap is the most well-known, though controversial, measure of gerrymandering is the Efficiency Gap (Stephanopoulos and McGhee 2015, Chambers, Miller, and Sobel 2017, Cover 2018). The efficiency gap is a measure of how efficiently a party wins seats given its state-wide vote share. If a party can win many seats with fewer votes, then the state is considered more gerrymandered. Empirically, the Efficiency Gap is associated with changes in the ideology of the median legislator and of the policies implemented (Caughy Tausanovitch and Warshaw, 2017).

[45]These findings are consistent with Edwards et al. (2017), who document that electoral districts for federal and state legislatures are more likely to cross local government borders if they are drawn by politicians than if they are drawn by independent commissions or courts. There are some differences between Edwards et al. (2017) and the analysis in this section. The unit of observation in Edwards et al. (2017) is a congressional or state legislative district. Their outcome variable is the number of counties that the district splits. Instead, I try to take into account state-level population constraints by using counties as the unit of analysis and predicting whether the county is in more districts than its population size implies.

ties can be in multiple districts and districts can contain multiple counties. Border mismatch of this nature occurs in every U.S. state. More generally, border mismatch is relevant to public spending in any country where administratively important local governments differ from electoral districts.

In the theoretical model, the amount of money that a county receives from the state depends on its share of voters in each district that the county intersects. Thus, a county's own outcome depends on the voting eligibility and turnout rate in any neighboring counties that share the same districts. A county that has high turnout relative to its neighbors is better off split across many districts. That way, the county represents a disproportionate share of voters in each district, relative to its population size. Conversely, a low turnout county is better off in its own district so that it won't be ignored by competing parties in favor of high turnout neighbors.

Transfers from U.S. states to county governments are consistent with specific empirical predictions implied by the model. Having a higher total share of voters leads to a statistically and economically significant increase in transfers to a given county. In particular, the effect size in on par with the changes to spending that occur due to variation in the number of representatives per person, suggesting a tradeoff between malapportionment and border mismatch.

The redistricting process underwent a revolution in the 1960s with the Supreme Court's "one person one vote" rulings. A second revolution seems to be underway, with increasing public support for more fair and transparent redistricting. However, the Supreme Court recently decided in a five to four vote that the court does not have authority to review partisan gerrymandering.[46] Nonetheless, ongoing legal battles in state courts call for better ways to measure gerrymandering and its harms to individual voters. This paper suggests that a simple metric, a county's total share of voters, has sizable effects on public spending outcomes and should be taken seriously in the redistricting process.

[46] *Lamone v. Benisek*, 588 U.S. (2019) and *Rucho v. Common Cause*, 588 U.S. (2019). Most recently, a North Carolina state court decided that partisan gerrymandering did in fact violate the state's constitution (*Common Cause v. Lewis*, 18-CVS-14001 N.C. Super Ct., Wake County (2019))

Appendices

A Proofs

Proof of Proposition 1. The probabilistic voting model presented here is a constant-sum game that fits the framework of Banks and Duggan (2005). By their Theorem 3, the game has a unique equilibrium which is in pure strategies if (i) the strategy space is compact and convex, (ii) each party's objective function is continuous in both parties' strategies, and (iii) each party's objective function is strictly concave in its own strategy. Each of these conditions are straightforward to demonstrate given the uniformly distributed ideological preferences.

From the closed form expression for p^d, it is immediate that each objective function ($\sum_d p^d$ for party A and $\sum_d (1-p^d)$ for party B) is continuous in $\mathbf{q}^A$ and $\mathbf{q}^B$ for any continuous $u_c(.)$. The strategy space of feasible balanced budget policies $Q = \{\mathbf{q} \in \mathbb{R}_+^C | \sum_c q_c \leq y\}$ is compact and convex. The objective function for party A is strictly concave in $\mathbf{q^A}$:

$$\frac{\partial^2}{\partial (q_{\hat{c}}^A)^2} \sum_d p^d = \gamma \sum_d \frac{v(\hat{c},d)}{\sum_c v(c,d)} \frac{\partial^2 u_{\hat{c}}(q_{\hat{c}}^A)}{\partial (q_{\hat{c}}^A)^2} < 0$$

and all off-diagonal elements of the Hessian matrix equal zero.

The best response for party A satisfies the following necessary first order conditions:

$$\sum_d \frac{v(\hat{c},d)}{\sum_c v(c,d)} \frac{\partial u_{\hat{c}}(q_{\hat{c}}^{*A})}{\partial q_{\hat{c}}^{*A}} = \frac{\lambda}{\gamma}.$$

The game is symmetric. Repeating the same steps for party B, $\mathbf{q}^{*A} = \mathbf{q}^{*B}$. □

The following comparative static result corresponds to the empirical predictions.

Proof of Corollary 1. Let $\kappa_{\hat{c}}$ equal the total share of voters:

$$\kappa_{\hat{c}} = \sum_d \frac{v(\hat{c},d)}{\sum_c v(c,d)}.$$

By the implicit function theorem,

$$\frac{\partial q_c^*}{\partial \kappa_c} = \frac{-u_c'(q_c^*)}{\kappa_c u_c''(q_c)} > 0$$

□

B Summary Statistics

Table 6: Summary Statistics

	Mean	St. Dev.	Min	Max
Outcome				
Transfers from state to county (millions USD)	132.05	420.88	0.15	6032.98
Per capita transfers from state to county (USD)	1478.03	649.54	174.68	7404.16
Districting and Voting Behavior				
Number of Representatives	2	2	1	28
Percent Voting Age Population (VAP)	0.771	0.030	0.667	0.919
Turnout rate	0.429	0.117	0.047	0.841
Total share of voters, VAP	1.035	1.875	0.001	25.381
Total share of voters, Turnout	1.060	1.883	0.002	25.189
County population / Ideal district population	1.042	1.973	0.001	24.613
Political Covariates				
Percent incumbent representatives	0.79	0.35	0.00	1.00
Percent Democrat representatives	0.32	0.42	0.00	1.00
Dem. majority in house	0.25	0.43	0.00	1.00
VRA Section 5 County	0.07	0.25	0.00	1.00
Other Covariates				
Population	82216	203530	89	3183143
Percent urban	0.37	0.29	0.00	1.03
Percent black	0.08	0.14	0.00	0.82
Percent Hispanic	0.08	0.13	0.00	0.91
Percent female	0.50	0.02	0.31	0.55
Percent without a high school diploma	0.11	0.05	0.02	0.30
Percent unemployed	0.04	0.01	0.00	0.12
Percent below Poverty Rate	0.16	0.05	0.01	0.39
Median household income (1,000 USD)	44.13	10.04	21.50	94.54
Observations	1203			

Note: These summary statistics are for the sample of counties which have non-missing turnout data.

C Sub-sample analyses

Table 7: Sub-sample analyses: y=log of intergovernmental transfers from state to county

	Using VAP to Measure Voters		
	Committee Drawn Maps	Fiscal Year 2012	Close Elections
ln(Avg. share of voters, VAP)	0.25***	0.36***	0.36***
	(0.09)	(0.08)	(0.12)
ln(Number of representatives)	0.28***	0.38***	0.43***
	(0.09)	(0.09)	(0.11)
ln(Population)	0.63***	0.56***	0.54***
	(0.09)	(0.08)	(0.12)
N	2002	2952	412
Dependent variable mean	17.37	17.57	18.26
Adjusted R^2	0.944	0.957	0.983
p-value for $H_0 : \pi_1 = \pi_2$	0.197	0.074	0.045
	Using Turnout to Measure Voters		
	Committee Drawn Maps	Fiscal Year 2012	Close Elections
ln(Avg. share of voters, Turnout)	0.29**	0.20**	0.11
	(0.11)	(0.09)	(0.12)
ln(Number of representatives)	0.28*	0.25***	0.19
	(0.12)	(0.09)	(0.13)
ln(Population)	0.61***	0.70***	0.72***
	(0.10)	(0.08)	(0.12)
N	489	843	255
Dependent variable mean	16.96	17.37	17.93
Adjusted R^2	0.948	0.964	0.978
p-value for $H_0 : \pi_1 = \pi_2$	0.760	0.078	0.156

Note: Standard errors clustered at the state-year level in parentheses. All regressions include state-year fixed effects and the following controls: percent incumbent representatives, percent Democrat representatives (also interacted with Democratic majority in the lower house), percent voting age population, percent urban, poverty rate, median household income, percent black, percent hispanic, percent female, percent unemployed, percent without a high school diploma, and an indicator for whether or not the county is covered by the Voting Rights Act Section 5.

D Variation by type of transfers

Table 8: Types of transfers: y=inverse hyperbolic sine of intergovernmental transfers

	Using VAP to Measure Voters				
	Edu.	Not Edu.	Roads	Public Welfare	Health and Hospitals
ln(Avg. share of voters, VAP)	0.19*	0.48***	0.05	2.19***	1.49
	(0.10)	(0.10)	(0.39)	(0.64)	(0.98)
ln(Number of representatives)	0.12	0.52***	0.12	2.56***	1.68*
	(0.11)	(0.09)	(0.38)	(0.67)	(0.96)
ln(Population)	0.86***	0.37***	0.65	-0.62	0.21
	(0.09)	(0.09)	(0.47)	(0.65)	(0.93)
N	5401	5401	5401	5401	5401
Dependent variable mean	17.91	16.67	14.17	7.84	9.59
Mean share of transfers	0.75	0.25	0.07	0.03	0.02
Adjusted R^2	0.801	0.835	0.587	0.625	0.426
p-value for $H_0 : \pi_1 = \pi_2$	0.021	0.289	0.318	0.023	0.330
	Using Turnout to Measure Voters				
	Edu.	Not Edu.	Roads	Public Welfare	Health and Hospitals
ln(Avg. share of voters, Turnout)	-0.08	0.22**	0.48**	1.44*	1.68**
	(0.10)	(0.09)	(0.21)	(0.83)	(0.79)
ln(Number of representatives)	-0.15	0.37***	0.60**	2.36***	2.62***
	(0.14)	(0.08)	(0.22)	(0.81)	(0.92)
ln(Population)	1.13***	0.58***	0.23	-0.57	0.17
	(0.17)	(0.06)	(0.23)	(0.75)	(0.77)
N	1203	1203	1203	1203	1203
Dependent variable mean	17.89	16.57	14.53	8.19	10.54
Mean share of transfers	0.76	0.24	0.06	0.04	0.02
Adjusted R^2	0.926	0.842	0.601	0.603	0.409
p-value for $H_0 : \pi_1 = \pi_2$	0.242	0.076	0.191	0.030	0.056

Note: Standard errors clustered at the state-year level in parentheses. All regressions include state-year fixed effects and the following controls: percent incumbent representatives, percent Democrat representatives (also interacted with Democratic majority in the lower house), percent voting age population, percent urban, poverty rate, median household income, percent black, percent hispanic, percent female, percent unemployed, percent without a high school diploma, and an indicator for whether or not the county is covered by the Voting Rights Act Section 5.

E School districts and counties

Table 9: Subsamples according to school district overlap: y=log of intergovernmental transfers from state to county

	One School District		Many School Districts	
	Edu.	Not Edu.	Edu.	Not Edu.
ln(Avg. share of voters, VAP)	0.16**	0.43***	0.16	0.53***
	(0.07)	(0.12)	(0.18)	(0.15)
ln(Number of representatives)	0.17**	0.55***	0.06	0.53***
	(0.06)	(0.14)	(0.20)	(0.12)
ln(Population)	0.81***	0.42***	0.92***	0.33**
	(0.06)	(0.13)	(0.17)	(0.13)
N	1712	1712	3689	3689
Dependent variable mean	18.01	16.37	17.87	16.81
Adjusted R^2	0.967	0.764	0.775	0.857
p-value for $H_0 : \pi_1 = \pi_2$	0.278	0.039	0.024	0.989
	One School District		Many School Districts	
	Edu.	Not Edu.	Edu.	Not Edu.
ln(Avg. share of voters, Turnout)	0.08	0.02	-0.10	0.29**
	(0.06)	(0.11)	(0.11)	(0.12)
ln(Number of representatives)	0.11	0.36*	-0.14	0.36***
	(0.07)	(0.17)	(0.13)	(0.10)
ln(Population)	0.81***	0.66***	1.18***	0.54***
	(0.07)	(0.12)	(0.17)	(0.08)
N	425	425	778	778
Dependent variable mean	17.51	16.62	17.05	16.54
Adjusted R^2	0.979	0.777	0.921	0.854
p-value for $H_0 : \pi_1 = \pi_2$	0.319	0.015	0.316	0.392

Note: States are divided into two categories: One School District indicates that the median number of school districts per county is less than or equal to one; Many School Districts indicates that the median number of school districts per county is greater than one. Standard errors clustered at the state-year level in parentheses. All regressions include state-year fixed effects and the following controls: percent incumbent representatives, percent Democrat representatives (also interacted with Democratic majority in the lower house), percent voting age population, percent urban, poverty rate, median household income, percent black, percent hispanic, percent female, percent unemployed, percent without a high school diploma, and an indicator for whether or not the county is covered by the Voting Rights Act Section 5.

F Border mismatch, independent redistricting commissions, and gerrymandering

Table 10: The extent of border mismatch: y= number of representatives (R_c)

	(1)	(2)	(3)
Independent commission	-1.19***	-0.61***	-0.53***
	(0.31)	(0.17)	(0.17)
Gubernatorial veto	-0.55**	-0.15	-0.06
	(0.24)	(0.16)	(0.15)
Efficiency Gap		1.27***	1.10***
		(0.32)	(0.29)
Ideal number of representatives	0.67***	1.11***	1.08***
	(0.13)	(0.02)	(0.03)
Total counties in the state	-0.00	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)
Total districts in the state	-0.01**	0.00	0.00
	(0.01)	(0.00)	(0.00)
VRA-covered county	0.10	-0.07	0.01
	(0.17)	(0.05)	(0.11)
Additional Controls			X
N	5485	2116	2116
Dependent variable mean	2.95	2.81	2.81
Adjusted R^2	0.659	0.897	0.899

Note: Standard errors clustered at the state-year level in parentheses. All regressions include year fixed effects. Independent commission equals 1 if an independent commission draws the electoral district map (AK, AZ, ID, IA, MT, WA). Advisory Commission equals 1 if there is a commission that provides suggestions when there is disagreement in the legislature. Political commission equals 1 if a committee made up of legislators and/or political appointees draws the map. The comparison group is states where the legislature itself draws the map. VRA-covered county indicates if the county is covered under Section 5 of the Voting Rights Act. The Additional Controls are: percent urban, percent black, percent democrat representatives, percent incumbent representatives, and whether or not the state has democratic control (interacted with percent democratic representatives in the state).

References

Angrist, Joshua D. and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." The Quarterly Journal of Economics 114(2): 533-575.

Ansolabehere, Stephen, Alan Gerber and James M. Snyder. 2002. "Equal Votes, Equal Money: Court-Ordered Redistricting and Public Expenditures in the American States." The American Political Science Review 96(4): 767-777.

Ansolabehere, Stephen, James M. Snyder, and Michael M. Ting. 2003. "Bargaining in Bicameral Legislatures: When and Why does Malapportionment Matter?" American Political Science Review 97(3): 471-481.

Ansolabehere, Stephen and James M. Snyder. 2006. "Party Control of State Government and the Distribution of Public Expenditures." Scandinavian Journal of Economics 108(4): 547-569.

Atlas, Cary M., Thomas W. Gilligan, Robert J. Hendershott and Mark A. Zupan. 1995. "Slicing the Federal Government Net Spending Pie: Who Wins, Who Loses, and Why." The American Economic Review 85(3): 624-629.

Banks, Jeffrey S. and John Duggan. 2005. "Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-motivated Candidates." In *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks,* ed. David Austen-Smith and John Duggan, 15-56. New York: Springer.

Baron, David P. and John A. Ferejohn. 1989. "Bargaining in Legislatures." American Political Science Review 83(4):1181-1206.

Baron, David P. 1991. "Majoritarian Incentives, Pork Barrel Programs, and Procedural Control." American Journal of Political Science 35(1): 57-90.

Baqir, Reza. 2002. "Districting and Government Overspending." Journal of Political Economy 110(6): 1318-1354.

Battaglini, Marco. 2014. "A dynamic theory of electoral competition." Theoretical Economics 9(2014): 515-554.

Berry, Christopher. 2008. "Piling On: Multilevel Government and the Fiscal Common-Pool." American Journal of Political Science 54(4): 802-820.

Besley, Timothy. 2006. "Principled Agents?: The Political Economy of Good Government." Oxford University Press.

Bouton, Laurent ⓡ Garance Genicot ⓡ Micael Castanheira. 2018. "Electoral Systems and Inequalities in Government Interventions." Working Paper.

Bowen, Daniel C. 2014. "Boundaries, Redistricting Criteria, and Representation in the U.S. House of Representatives." American Politics Research 42(5): 856-895.

Cameron, Charles, David Epstein and Sharyn O'Halloran. 1996. "Do Majority-Minority Districts Maximize Substantive Black Representation in Congress." American Political

Science Review 90(4):794-812.

Cascio, Elizabeth U. and Ebonya Washington. 2014. "Valuing the Vote: The Redistribution of Voting Rights and State Funds Following the Voting Rights Act of 1965." The Quarterly Journal of Economics 129(1): 379-433.

Caughey, Devin, Chris Tausanovitch and Christopher Warshaw. 2017. "Partisan Gerrymandering and the Political Process: Effects on Roll-Call Voting and State Policies." Election Law Journal 16(4): 453-469.

Chambers, Christopher P., Alan D. Miller, and Joel Sobel. 2017. "Flaws in the Efficiency Gap." Journal of Law and Politics 33(1): 1-33.

Chingos, Matthew M. and Kristin Blagg. 2017. "Making Sense of State School Funding Policy." Urban Institute Research Report.

Cover, Benjamin Plener. 2018. "Quantifying Partisan Gerrymandering: An Evaluation of the Efficiency Gap Proposal." Stanford Law Review 70: 1131-1233.

Cox, Gary W. and Jonathan N. Katz. 1999. "The Reapportionment Revolution and Bias in U.S. Congressional Elections." American Journal of Political Science 43(3): 812-841.

Cox, Gary W. and Mathew D. McCubbins. 1986. "Electoral Politics as a Redistributive Game." The Journal of Politics 48(2): 370-389.

Dixit, Avinash and John Londregan. 1996. "The Determinants of Success of Special Interests in Redistributive Politics." The Journal of Politics 58(4): 1132-1155.

Dixit, Avinash and John Londregan. 1998. "Fiscal Federalism and Redistributive Politics." Journal of Public Economics 68: 153-180.

Edwards, Barry, Michael Crespin, Ryan D. Williamson, and Maxwell Palmer. 2017. "Institutional Control of Redistricting and the Geography of Representation." The Journal of Politics 79(2): 722-726.

Elis, Roy, Neil Malhotra and Marc Meredith. 2009. "Apportionment Cycles as Natural Experiments." Political Analysis 17(4): 358-376.

Enelow, James M. and Melvin J. Hinich. 1982. "Nonspatial Candidate Characteristics and Electoral Competition." The Journal of Politics 44(1): 115-130.

Epple, Dennis, Thomas Romer and Holger Sieg. 2001. "Interjurisdictional Sorting and Majority Rule: An Empirical Analysis." Econometrica 69(6): 1437-1465.

Fletcher, Deborah and Steven Slutsky. 2011. "Campaign Allocations Under Probabilistic Voting." Public Choice 146(3-4):469-499.

Golden, Miriam and Brian Min. 2013. "Distributive Politics Around the World." Annual Review of Political Science 16: 73-99.

Gulzar, Saad and Benjamin J. Pasquale. 2017. "Politicians, Bureaucrats, and Development: Evidence from India." American Political Science Review 111(1): 162-183.

Hall, Andrew B. and Daniel M. Thompson. 2018. "Who Punishes Extremist Nominees? Candidate Ideology and Turning Out the Base in US Elections." American Political

Science Review 112(3): 509-524.

Fiva, Jon H. and Askill H. Halse. 2016. “Local Favoritism in At-Large Proportional Representation Systems.” Journal of Public Economics 143: 15-26.

Hirano, Shigeo. 2006. “Electoral Institutions, Hometowns, and Favored Minorities.” World Politics 58: 51-82.

Hirano, Shigeo. 2011. “Do Individual Representatives Influence Government Transfers? Evidence from Japan.” The Journal of Politics 73(4): 1081-1094.

Hirano, Shigeo and Michael M. Ting. 2015. “Direct and Indirect Representation.” British Journal of Political Science 45(3): 609-634.

Hoover, Gary A. and Paul Pecorino. “The Political Determinants of Federal Expenditure at the State Level.” Public Choice 123: 95-113.

Incerti, Devin. 2015. “The Optimal Allocation of Campaign Funds in House Elections.” Working Paper.

Jennes, Geert and Damiaan Persyn. 2015. The Effect of Political Representation on the Geographic Distribution of Income: Evidence using Belgian Data.” European Journal of Political Economy 37: 178-194.

Knight, Brian. 2008. “Legislative Representation, Bargaining Power and The Distribution of Federal Funds: Evidence from the US Congress.” The Economic Journal 118: 1785-1803.

Lee, Frances E. 1998. “Representation and Public Policy: The Consequences of Senate Apportionment for the Geographic Distribution of Federal Funds.” Journal of Politics 60: 34-62.

Lee, Frances E. 2000. “Senate Representation and Coalition Building in Distributive Politics.” The American Political Science Review 94: 59-72.

Levitt, Justin. 2010. “A Citizen’s Guide to Redistricting.” The Brennan Center for Justice at New York University School of Law.

Lindbeck, Assar and Jorgen Weibull. 1987. “Balanced-budget redistribution as the outcome of political competition.” Public Choice 52: 273-297.

McCarty, Nolan, Keith T. Poole and Howard Rosenthal. 2009. “Does Gerrymandering Cause Polarization?” American Journal of Political Science 53(3):666-680.

Martin, Gregory J. 2017. “Dividing the Dollar with Formulas.” Journal of Politics 80(2): 479-493.

Martin, Paul S. 2003. “Voting’s Rewards: Voter Turnout, Attentive Publics, and Congressional Allocation of Federal Money.” American Journal of Political Science 47(1): 110-127.

Milesi-Ferretti, Gian Maria, Roberto Perotti and Massimo Rostagno. 2002. “Electoral Systems and Public Spending.” The Quarterly Journal of Economics 117(2): 609-657.

Parente, Paulo M.D.C. and J.M.C. Santos Silva. 2012. “A cautionary note on tests of overidentifying restrictions." Economics Letters 115: 314-317.

Persson, Torsten and Guido Tabellini. 2000. "Political Economics Explaining Economic Policy." MIT Press. Cambridge, MA. Chapters 6,7, and 8.

Snyder, James M. 1989. "Election Goals and the Allocation of Campaign Resources." Econometrica 57(3): 637-660.

Snyder, James M. and David Stromberg. 2010. "Press Coverage and Political Accountability." Journal of Political Economy 118(2): 355-408.

Synder, James M. and Michiko Ueda. 2007. "Do Multimember Districts Lead to Free-Riding?" Legislative Studies Quarterly 32: 649-679.

Spenkuch, Jorg and David Toniatti. 2018. "Political Advertising and Election Results." The Quarterly Journal of Economics

Stephanopoulos, Nicholas O. and Eric M. McGhee. 2015. "Partisan Gerrymandering and the Efficiency Gap." University of Chicago Law Review 82(2): 831-900.

Stock, James and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In: Andrews DWK Identification and Inference for Econometric Models. New York: Cambridge University Press; pp. 80-108.

Stromberg, David. 2004a. "Radio's Impact on Public Spending." The Quarterly Journal of Economics 199(1): 189-221.

Stromberg, David. 2004b. "Mass Media Competition, Political Competition, and Public Policy." The Review of Economic Studies 71(1):265-284.

Stromberg, David. 2008. "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida." American Economic Review 98(3): 769-807.

Winburn, Jonathan and Michael W. Wagner. 2010. "Carving Voters Out: Redistricting's Influence on Political Information, Turnout, and Voting Behavior." Political Research Quarterly 63(2): 373-386.

Young, Alywn. 2018. "Consistency Without Inference: Instrumental Variables in Practical Application." Working Paper.

Yuan, Haishan. 2018. "National Representation and Local Public Expenditure: A Natural Experiment from Japan." Working Paper.