Advance Access publication July 23, 2009

Political Analysis (2009) 17:358–376
doi:10.1093/pan/mpp012

Apportionment Cycles as Natural Experiments

Roy Elis
Department of Political Science, Stanford University, 616 Serra Street, Stanford, CA 94305
e-mail: elis@stanford.edu (corresponding author)

Neil Malhotra
Graduate School of Business, Stanford University, 518 Memorial Way, Stanford, CA 94305

Marc Meredith
Department of Political Science, University of Pennsylvania, 208 S. 37th Street, Philadelphia, PA 19104

Although there are compelling theoretical reasons to believe that unequal political representation in a legislature leads to an unequal distribution of funds, testing such theories empirically is challenging because it is difficult to separate the effects of representation from the effects of either population levels or changes. We leverage the natural experiment generated by infrequent and discrete census apportionment cycles to estimate the distributional effects of malapportionment in the U.S. House of Representatives. We find that changes in representation cause changes in the distribution of federal outlays to the states. Our method is exportable to any democratic system in which reapportionments are regular, infrequent, and nonstrategic.

1 Introduction

Following Census 2000, the state of Utah was allocated three seats in the U.S. House of Representatives, falling short of the threshold for a fourth seat by 857 people—less than 0.04% of its population. Utah vociferously challenged the Census count, taking its case all the way to the Supreme Court (*Utah v. Evans* 2001, 2002) where it contested both the imputation methods used by the Census Bureau to estimate apportionment populations and also the statute that counts overseas military personnel but not Mormon missionaries residing abroad. Utah lost both cases. As a result, the average district in Utah in 2003 represented 746,000 residents instead of 560,000 residents as would have been the case if the state had been awarded four seats. The contested seat went to North Carolina, which increased its representation in the House from 12 to 13 members.

Authors' note: We would like to thank Jowei Chen, Vincenzo Galasso, Saar Golde, Karen Jusko, Brian Knight, Frances Lee, Jonathan Rodden, Jonathan Wand, and the participants of Stanford University's Workshop in Statistical Modeling and American Empirical Seminar for valuable comments and suggestions. We would also like to thank Fred Wollens for his excellent research assistance. A previous version of this article was presented at the 2008 Annual Meeting of the Midwest Political Science Association in Chicago.

© The Author 2009. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

Although the U.S. House follows the principle of ''one person, one vote,'' significant disparities in representation are unavoidable in practice as demonstrated by the fact that the 2000 apportionment left Utah with 15% less representation than the average state on a per capita basis. The extent to which Utah contested the results of Census 2000 suggests that differences in relative representation have important policy consequences, not the least of which is the distribution of federal funds to the states. The fundamental problem of confirming this intuition empirically is that both representation and population may directly affect distributional outcomes. Hence, observed associations between representation and distributional outcomes are likely to jointly capture the effect of representation and any unmeasured direct effects of population.

To illustrate the identification problem, consider the well-documented positive relationship between *changes* in states' net receipts from federal expenditures and changes in representation per capita in the Senate (Atlas et al. 1995; Hoover and Pecorino 2005; Larcinese et al. 2007). Because each state has a fixed number of senators, per capita representation is directly proportional to the inverse of state population, and hence changes in Senate representation are driven solely by changes in state population. In an observational study, disentangling the effect of population changes from the effect of representational changes necessitates making strong functional form restrictions on the effects of changes in population. It is also likely that there are important unobservables (e.g., economic development) that drive both the changes in state population and also a state's share of the distribution of federal dollars. For instance, states with below-average population growth gain relative representation in the context of the Senate by definition. If, for example, the federal government directs financial assistance toward economically depressed areas, regressions predicting changes in federal spending with changes in relative representation will pick up a spurious correlation due to the economic processes driving both population trends and federal spending trends. Thus, although these studies are instructive at pointing to a correlation between representation and distribution, they do not provide compelling evidence of the causal effect.

Other studies examine the correlations between *levels* of representation and levels of expenditures (Lee 1998, 2000; Larcinese et al. 2007; Knight 2008). Yet, as Alesina and Spolaore (1997) argue, geographically large political units exhibit diseconomies of scale and therefore require more public monies than geographically small units to obtain the same level of services. Hence, the fact that large, sparsely populated states such as Alaska and Montana receive higher per capita federal funds may mainly be driven by greater need.

In this paper, we use a natural experiment in the U.S. House of Representatives to separate the distributional effects of population changes, population levels, and potential unobservables from the distributional effects of changes in political representation. Specifically, we use the infrequent and discrete nature of reapportionment in the U.S. House of Representatives as our identifying source of variation. Previous work has overlooked the fact that the very mechanisms designed to reapportion the House constitute a structural source of variation in representational equality. The House is reapportioned only once every 10 years despite continuous population changes among the states. This implies a structural break in relative representation following each census-based reapportionment. Our analysis leverages the discrete shocks created by the census cycles, allowing us to compare outlays received by states immediately before and after reapportionment. The effect is identified by ''switchers,'' or states that gain or lose seats. By focusing on a narrow window of time around each reapportionment, we are able to develop a difference-in-first-differences technique to estimate the effects of representation without imposing strict functional form restrictions.

Changes in House representation significantly increase federal outlays at the state level. Our point estimates imply that the elasticity of total outlays with respect to representation is

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

about 0.07. In other words, a 10% increase in a state's share of representatives equates to a 0.7% increase in a state's share of the federal budget, suggesting that Utah's outlays were about 2% lower in 2004 than they would have been had Utah been awarded a fourth Congressional seat, a loss of about $100 per capita. We also find that the effect of representation is contextual; large states do not benefit as much from the same percentage change in the number of representatives as small states. We find no evidence of asymmetric effects between gains and losses in representation.

The article proceeds as follows. In the second section, we review recent literature linking representation to distribution. Section 3 describes the complex apportionment process used by the Census Bureau to assign congressional seats to states. Section 4 presents two identification strategies stemming from census-based apportionment in the U.S. House of Representatives. We describe the data and measures in Section 5. In the final section, we present the results and discuss their implications.

2 Recent Advances in the Study of Representation and Distribution

A fundamental concept in the study of politics is *representation*, or the degree to which elected officials with political power reflect the preferences of citizens (Mansbridge 2003). Our paper focuses on estimating the distributive consequences of unequal representation in a legislature. We assume that legislators respond to reelection incentives and thus compete to procure targetable goods and benefits for their districts (Mayhew 1974; Fiorina 1989; Weingast et al. 1981). Moreover, we assume a positive correlation between voting weights and bargaining power, as is usually the case in distributive models of legislative bargaining (Baron and Ferejohn 1989).[1] Thus, we expect that overrepresented political units will receive a disproportionately large share of distributive goods.[2]

Although the theoretical bases for the relationship are sufficiently compelling, a general method for estimating its magnitude has proved elusive. A number of recent studies have addressed the identification problem in innovative ways. Ansolabehere, Gerber, and Snyder (2002) use the natural experiment generated by the Supreme Court's decision in *Baker v. Carr* (1962), which mandated one person, one vote in U.S. state legislatures as an exogenous shock to state-level apportionment. They estimate that the court-ordered political equalization resulted in a transfer of about $7 billion from counties that were overrepresented to counties that were underrepresented. Knight (2008) and Hauk and Wacziarg (2007) examine differences in earmarks that originate in the Senate versus the House and find that relatively more earmarks are targeted to small states in the Senate's version of the legislation. Hirano (2006) and Hirano and Ting (2008) use legislator deaths in the Japanese Diet as an exogenous shock to estimate how targeted spending changes once a constituency loses representation. The authors find a significant effect of representation among ''weak'' constituency groups. Finally, in a study closest to our own, Falk (2006) uses discontinuities created by integer constraints in the apportionment formula in the House of Representative as an instrumental variable (IV) for representation. His point estimates suggest that the returns to gaining a second member of Congress are about $1000 per capita.

[1] It should be noted that examples can be constructed in a Baron and Ferejohn (1989) framework where expected payoffs decrease from an increase in voting weights.

[2] The Baron-Ferejohn model has been extended in numerous ways to explain how the strength of a constituency's voting weight significantly increases its ability to bargain and procure resources from a common pool (Banks and Duggan 2000; Snyder et al. 2005; Chen and Malhotra 2007), as well as the conditions under which malapportionment matters (Ansolabehere, Snyder, and Ting 2003).

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

Even these innovative approaches, however, have certain limitations. Ansolabehere, Gerber, and Snyder (2002) use a one-time experiment, which is specific to the United States in the 1960s and therefore their research design does not generalize to other eras in the U.S. or to comparative contexts. Knight's (2008) and Hauk and Wacziarg's (2007) method can only be applied to a subset of bills and rests on the assumption that legislators do not strategically anticipate how the bills will be resolved in conference when voting within their respective chambers. Hirano's (2006) and Hirano and Ting's (2008) approach would be difficult to apply outside the Japanese context due to the paucity of deaths and the speed with which members are replaced. Falk's (2006) approach limits him to estimating a local average treatment effect (LATE) of the effects of representation in small states. This brings into question whether his results also apply to larger states. Moreover, the necessity of looking at small states restricts the sample size from which he gains identification to a smaller number of observations, which raises concerns about small sample IV bias (see Ullah 2004 for an overview).

Using the timing of census apportionment has several advantages over previous methods. First, the method can be applied to most any legislative setting and to a broad set of outcomes so long as reapportionment is relatively infrequent and nonstrategic, making it especially useful for the comparative study of representation. Second, the method requires no assumptions about strategic interactions between the upper and lower chambers of the legislature. Third, it leverages the full sample of U.S. states and avoids potential bias of small-sample IVs.

3 The Apportionment Mechanism as a Natural Experiment

3.1 *Population-Based Apportionment*

Article 1, Section 2, of the Constitution established that the apportionment of the House of Representatives shall be based upon a national decennial census. The House has been regularly apportioned according to Census population estimates since 1790, with the exception of the 1920 Census after which there was no reapportionment. The number of seats in the House has been fixed at 435 since 1911, with the exception of a temporary increase to 437 seats following the admission of Hawaii and Alaska in 1959.[3] Since 1941, the House has been reapportioned with perfect regularity—decennially, for a fixed size of 435, by the Method of Equal Proportions (also called the Hill Method)—allowing us to consider reapportionment as a regularly recurring natural experiment.[4]

3.2 *The Method of Equal Proportions*

Every census a reapportionment constant is constructed that is approximately equal to the U.S. apportionment population divided by 435. Each state's apportionment population is divided by the reapportionment constant to produce its deserved quotient of seats with an integer part Q and fractional part q.[5] A state receives Q representatives if $Q.q \leqslant \sqrt{Q(Q+1)}$

[3]The temporary addition of seats for Alaska and Hawaii did not affect either the 1950 or the 1960 reapportionment, both of which were based on 435 seats and the apportionment populations of all relevant states.

[4]See http://www.census.gov/population/www/censusdata/apportionment/history.html and Balinski and Young (2001) for a history. It is less clear that reapportionments before 1940 can be treated as a natural experiment because Congress endogenously selected a new apportionment rule for each apportionment. For example, Balinski and Young (2001) note that the 1900 and 1910 apportionment rules were designed so that no state lost a representative in the House.

[5]In practice, the common divisor must be modified slightly to ensure that the size of the House remains exactly 435 after rounding.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

Table 1 Method of equal proportions cutpoints for the first 10 seats

State seat	*Cutpoint*	*Population needed (2000)*
1	$\sqrt{0 \cdot 1} = 0$	
2	$\sqrt{1 \cdot 2} \approx 1.414$	913,309
3	$\sqrt{2 \cdot 3} \approx 2.449$	1,581,898
4	$\sqrt{3 \cdot 4} \approx 3.464$	2,237,142
5	$\sqrt{4 \cdot 5} \approx 4.472$	2,888,138
6	$\sqrt{5 \cdot 6} \approx 5.477$	3,537,232
7	$\sqrt{6 \cdot 7} \approx 6.481$	4,185,309
8	$\sqrt{7 \cdot 8} \approx 7.483$	4,832,779
9	$\sqrt{8 \cdot 9} \approx 8.485$	5,479,856
10	$\sqrt{9 \cdot 10} \approx 9.487$	6,126,666

and $Q + 1$ representatives if $Q.q > \sqrt{Q(Q+1)}$. For example, a state with a quotient of 3.48 receives four representatives, while a state with a quotient of 3.45 receives three representatives, since $\sqrt{3 \cdot 4} \approx 3.464$. Notice that the cutpoints are seat specific and that the fractional part converges to 0.5 as Q gets large (Table 1). Also notice that the Hill method ensures the constitutionally mandated minimum of one seat per state.

3.3 *Two Natural Experiments*

Census-based apportionment creates two distinct natural experiments. The first occurs because the Method of Equal Proportions does not award seats on a fractional basis. Seat shares are therefore discontinuous in population shares near the seat-specific thresholds that determine whether deserved fractions of a seat are rounded up or down. The thresholds create variation in relative representation, defined as the ratio of seat shares to population shares, which we explain in detail below (see Fig. 1 for data from the 2000 apportionment cycle). Although we typically think of the House of Representatives as a well-apportioned legislature, it is not perfectly apportioned even immediately following reapportionment. In the 2000 cycle, for example, states such as Montana, Utah, and Mississippi barely missed the population thresholds necessary for gaining an additional seat, while states such as Rhode Island, Nebraska, and Iowa barely cleared their respective thresholds, leading to imbalanced representation.

The second natural experiment is embedded in the timing of apportionment, which induces a regular 10-year cycle in the relative equality of states' representation. States are most equally represented immediately following reapportionment, but the House grows

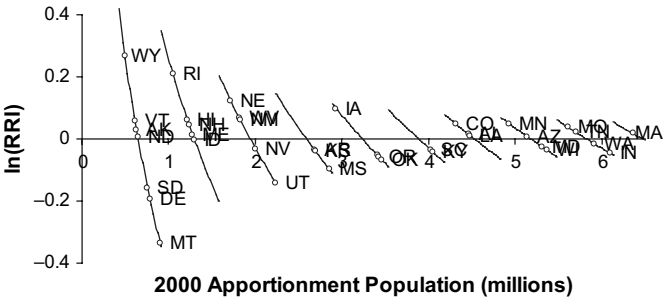


Fig. 1 Representation per capita index for the 2000 apportionment cycle (states with 10 or fewer representatives).

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

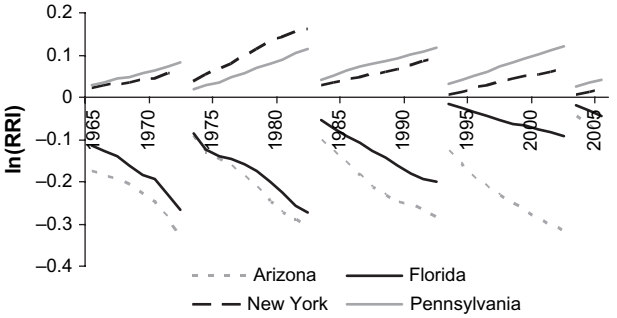


Fig. 2 Representation per capita index by year for a selection of U.S. states.

malapportioned over the years due to differential population growth rates across states. Assuming these differences are relatively constant, the House reaches maximum malapportionment by the end of each cycle. The next census then leads to the next reapportionment, equalizing representation in the first year of the subsequent cycle and so on.

The apportionment cycle begins with the census in year 0 of each decade. Official apportionment populations are reported to the President by the end of the year and the assignment of seats to states is certified soon after. The first elections following reapportionment occur in November of year 2 and representatives are seated in January of year 3. The newly elected representatives pass their first budget in October of that year, which determines federal spending in the states for year 4. For example:

- April–December 2000: census performed and results certified
- November 2002: elections in newly apportioned districts
- January 2003: Reapportioned Congress begins session
- October 2003: Budget is enacted for the 2004 fiscal year.

The most important feature of the cycle for our purposes is that states receive the last round of federal transfers distributed under maximum malapportionment in fiscal years ending with 3, and receive the first round of federal transfers legislated by the newly apportioned House in year 4 of each decade.

Figure 2 illustrates how the discrete changes in representation create sudden changes in relative representation. Fast-growing states such as Arizona and Florida are often left with substantially less than average representation per capita. However, this disadvantage tends to be significantly lower in the years immediately following reapportionment than in the years immediately before. For example, even though Arizona's population grew by similar amounts in both 2002 and 2003, Arizona was substantially less underrepresented in 2003 following its sudden increase from six to eight representatives. Conversely, states with below-average growth, such as New York and Pennsylvania, tend to be substantially less overrepresented in years immediately following a reapportionment (see Figure A1 in the Appendix).[6]

[6]Trends in population and seat shares from 1935 to 2007 are provided in the Appendix for all states except Alaska, Vermont, and Wyoming. The graphs make clear that although seat shares are correlated with population shares, the shifts in seat shares at the time of each reapportionment are abrupt, whereas year-to-year population trends are smooth. Additionally, Fig. 4 below demonstrates that our assumption of smooth population trends is empirically justified.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

4 Methodological Overview

In this section, we discuss in detail the identification problem involved in estimating the effect of House representation on distribution, which we operationalize as federal outlays to the states. We then explain our methodological strategy based on testing whether the structural breaks in representation between years 2 and 3 of each decade correspond to changes in shares of outlays between years 2 and 4.

Figure 3 provides the intuition behind our estimation strategy. The top panel of Fig. 3 highlights that between 1992 and 1994, there was a sudden 4.3 percentage point shift in relative representation from states that lost seats to states that gained seats (hereafter,

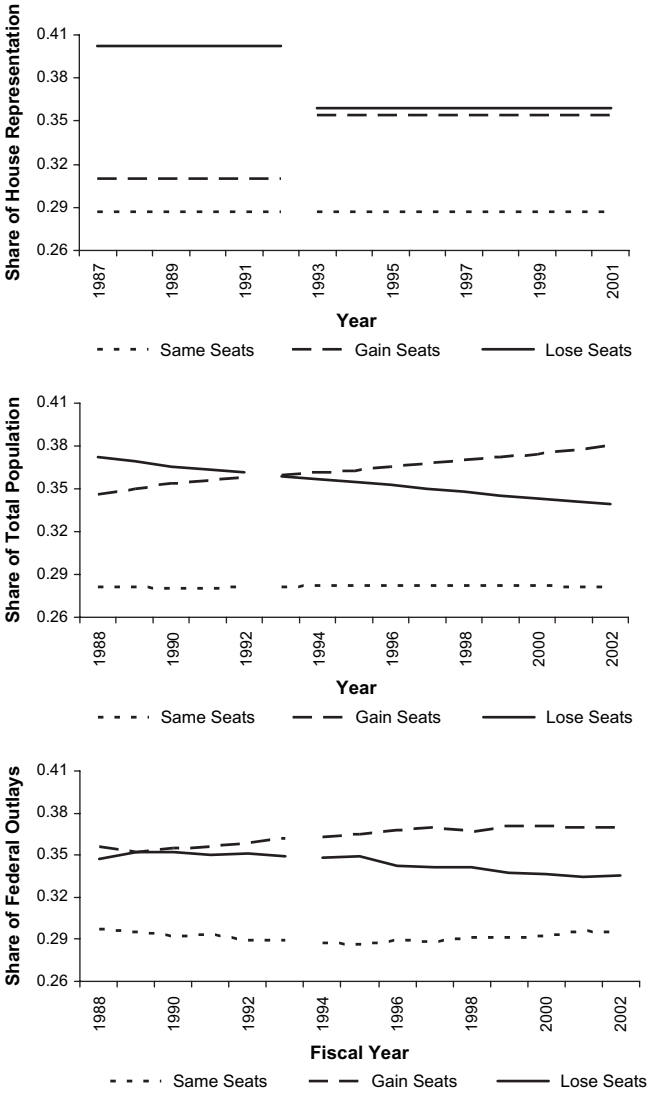


Fig. 3 Trends in House representation, population, and outlays by 1990 apportionment outcome.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

switchers). At the same time, there was a relatively linear population trend among both groups of states (Fig. 3, middle panel). If representation per se affects distributive outcomes, then we should see a greater increase in the share of federal outlays going to states gaining seats after the sudden shift in representation in 1993 (Fig. 3, bottom panel).

We model the share of federal outlays going to state s at time t as:

$$\frac{\text{outlays}_{s,t}}{\text{outlays}_{\bullet,t}} = f\left(\frac{\text{rep}_{s,t}}{435}, \text{pop}_{s,t}, \Delta\text{pop}_{s,t}, \theta_s, \epsilon_{s,t}\right), \tag{1}$$

where $\text{outlays}_{s,t}$ is the amount of federal dollars allocated to state s at time t, $\text{outlays}_{\bullet,t}$ is the total amount of federal dollars allocated to states at time t, $\text{rep}_{s,t}$ is the number of representatives in state s at time t, $\text{pop}_{s,t}$ is the population of state s at time t, $\Delta\text{pop}_{s,t}$ is the change in population of state s between time $t - k$ and time t, θ_s is an idiosyncratic state-specific effect, and $\epsilon_{s,t}$ is a stochastic error. We expect that $\frac{\partial f()_{s,t}}{\partial \text{rep}_{s,t}} > 0, \frac{f()_{s,t}}{\partial \text{pop}_{s,t}} > 0, \frac{\partial f()_{s,t}}{\partial \Delta\text{pop}_{s,t}} < 0$.

Assume that $f()$ is multiplicably separable such that we can rewrite equation (1) as:

$$\frac{\text{outlays}_{s,t}}{\text{outlays}_{\bullet,t}} = \left(\frac{\text{rep}_{s,t}}{435}\right)^{\beta} f(1, \text{pop}_{s,t}, \Delta\text{pop}_{s,t}, \theta_s, 1)\epsilon_{s,t}. \tag{2}$$

Log both sides of equation (2) such that:

$$\ln\left(\frac{\text{outlays}_{s,t}}{\text{outlays}_{\bullet,t}}\right) = \beta\ln\left(\frac{\text{rep}_{s,t}}{435}\right) + \ln(f(1, \text{pop}_{s,t}, \Delta\text{pop}_{s,t}, \theta_s, 1)) + \ln(\epsilon_{s,t}). \tag{3}$$

In this equation, β captures the elasticity of outlays with respect to representation; that is, β represents the percent change in outlay share with respect to a 1% change in representation share. There are two complications that make it difficult to estimate β via direct estimation of equation (3). First, the functional form of $g()_{s,t} = f(1, \text{pop}_{s,t}, \Delta\text{pop}_{s,t}, \theta_s, 1)$ is unknown to the researcher. Second, representation in the House is itself a function of lagged population. Therefore, if the researcher does not correctly specify the direct effects of levels and changes in population on outlays, the estimated effects of representation are likely to be contaminated by omitted variable bias.

There are two solutions to the identification problem. The first is to find an IV that affects representation but is independent of the measurement error in the direct effect of population and $\epsilon_{s\,t}$. As described above, the indivisibility of seats inherent in the Method of Equal Proportions can serve as an instrument to identify the effects of representation (Falk 2006; Angrist and Lavy 1999). The principle advantage of the IV approach is that the instrument is likely to satisfy the exclusion restriction of being independent of the unobservables, including any measurement error in the effect of population change and growth. However, the IV approach also has several limitations. For states with seven or more representatives, changes in population explain about 93% of the variation in the number of representatives within states across apportionment cycles compared to only about 63% of the variation in states with six or fewer representatives.[7] Hence, the first stage of the IV approach lacks sufficient explanatory power to be used in a substantial portion of the population of states. This is problematic for two reasons. First, it severely limits the sample size

[7]These values come from the *R*-squared of a regression of actual changes in the number of representatives on deserved changes in the number of representatives based on a simulated Congress that awards fractional seats.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

from which the effect of apportionment can be estimated. We find that one can only incorporate fewer than 20 changes in the number of representatives into the IV model, raising concerns about small sample IV bias. Second, because IV can only provide us with a LATE of representation among the subpopulation of states affected by changes in the instruments, we can only use the IV model to assess the effects of representation in small states (Imbens and Angrist 1994).

The second method, which we adopt in this paper, is to leverage the discrete timing of apportionment as the experiment. We develop an estimation framework that separates the effects of representation from the direct effects of population—as well as unobserved long-run processes of economic growth or decay associated with population changes—by using periods without changes in representation to control for $f()_{s,t}$ without having to know the exact functional form. Our framework compares states just before and just after gaining representation, using a difference-in-first-differences technique to control for preexisting trends in those states gaining or losing representation. The effect is identified by switchers, or states that experience a change in the number of representatives.

The benefits of using the difference-in-first-differences approach is that we are able to incorporate a significant amount of information that is lost in the IV approach. Instead of using less than 20 switches as in the IV model, our method incorporates 53 switches. Moreover, we are not limited to estimating a local treatment effect for small states and do not have the same concern about small sample bias.

Suppose that Congress is reapportioned between time periods $t-1$ and t. Taking first differences in $f()_{s,t}$ and $f()_{s,t-2}$,[8] gives us that:

$$f()_{s,t} - f()_{s,t-2} = \beta(\ln(\text{rep}_{s,t}) - \ln(\text{rep}_{s,t-2})) + \Delta g()_{s,t} + \Delta\epsilon()_{s,t}, \tag{4}$$

where $\Delta g()_{s,t} = (g(\text{pop}_{s,t}, \Delta\text{pop}_{s,t}, \theta_s) - g(\text{pop}_{s,t-2}, \Delta\text{pop}_{s,t-2}, \theta_s))$ and $\Delta\epsilon_{s,t} = \epsilon_{s,t} - \epsilon_{s,t-2}$. Because representation only changes before the seating of every fifth Congress, we know that $\text{rep}_{s,t-2} = \text{rep}_{s,t-4}$. Therefore, if we take difference-in-first-differences, we obtain:

$$(f()_{s,t} - f()_{s,t-2}) - (f()_{s,t-2} - f()_{s,t-4}) = \beta(\ln(\text{rep}_{s,t}) - \ln(\text{rep}_{s,t-2})) + \eta_t, \tag{5}$$

where $\eta_t = (\Delta g()_{s,t} - \Delta g()_{s,t-2}) + (\Delta\epsilon_{s,t} - \Delta\epsilon_{s,t-2})$.

If we assume that $E[\Delta g()_{s,t} | \text{rep}_{s,t}, \text{rep}_{s,t-2}] = \Delta g()_{s,t-2}$ and $E[\Delta\epsilon_{s,t+2} - \Delta\epsilon_{s,t} | \text{rep}_{s,t}, \text{rep}_{s,t-2}] = 0$, then we are able to obtain an unbiased estimate of β via ordinary least squares (OLS). We believe this is a reasonable assumption. In our sample, growth rates within states are very persistent across time: $\Delta\text{pop}_{s,t}$ and $\Delta\text{pop}_{s,t-2}$ are both on average 2.5 percentage points larger than the national average in states where $\text{rep}_{s,t} > \text{rep}_{s,t-2}$ and are both on average 1.1 percentage points less than the national average in states where $\text{rep}_{s,t} < \text{rep}_{s,t-2}$. We also plot the full distribution of $\Delta\text{pop}_{s,t}$ and $\Delta\text{pop}_{s,t-2}$ for switchers (Fig. 4). Most observations lie near the 45 degree line, suggesting that population trends tend to be smooth even in the states that experienced a change in their seat allocation. The slope of the regression line of the relationship between population shifts around the census is 1.004 (SE = 0.043). We fail to reject the null hypothesis that the slope is equal to the 45

[8]We take first difference between t and $t-2$ rather than t and $t-1$ for two reasons. Substantively, the lags in the budget cycle create concerns that reapportionment may begin having some affect on the budget in time period $t-1$. More practically, we can only observe our primary dependent variable in 2-year increments following the 1970 apportionment cycle. The estimated effects of representation are substantially smaller when we estimate $(f()_{s,t} - f()_{s,t-2}) - (f()_{s,t-1} - f()_{s,t-3})$ for the years we do have data.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

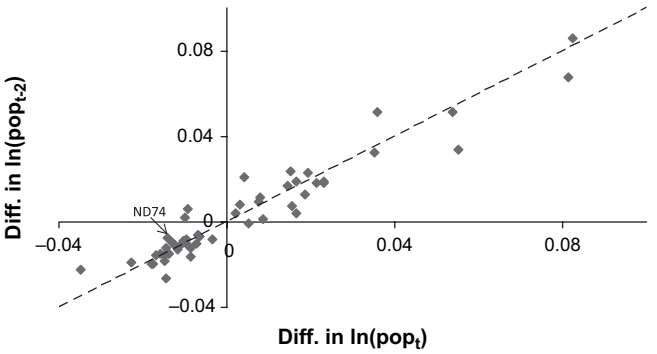


Fig. 4 Population changes and 2-year lagged population changes for switchers.

degree line ($p = .93$). These patterns increase our confidence in the assumptions about the similarities between $g(\text{pop}_{s,t-2}, \Delta\text{pop}_{s,t-2}, \theta_s)$ and $g(\text{pop}_{s,t}, \Delta\text{pop}_{s,t}, \theta_s)$.

5 Data and Measures

5.1 *Measuring Representation*

Our primary measure of representation is seat shares in the House of Representatives. Data on the seat apportionment for each state and on apportionment populations are from Balinski and Young (2001). We prefer this specification because it ensures that our representation measure is independent of population. In our estimations, we transform representation shares as follows: we take the difference in the natural log of representation share between time $t - 2$ and time period t. Note that because we are taking the difference in logs, this is approximately equal to the percent change in the share of representatives in the U.S. House.

A population-normalized measure of representation, however, allows us to describe the data in an intuitive manner and run alternative specifications of our main model. In addition to seat shares, we adopt the relative representation index (RRI) (David and Eisenberg 1961; Ansolabehere, Gerber, and Snyder 2002). The RRI for state i at time t measures the number of seats per person in the state relative to the number of seats per person in the entire House of Representatives. Thus, $\text{RRI}_{s,t} = \frac{\text{rep}_{s,t}/\text{pop}_{s,t}}{435/\text{pop}_{\bullet,t}}$. If a state's per capita representation is equal to average district size in the House, RRI = 1. Values of RRI less than 1 indicate underrepresentation, whereas values of RRI greater than 1 indicate overrepresentation. Following Ansolabehere, Gerber, and Snyder (2002), Woon (2007), and Lee (1998), we take the natural log of RRI to make gains and losses of equal representation symmetric and so that coefficients can be interpreted as elasticities. Table 2 shows the distribution of logged RRI immediately before and immediately after each of the past four apportionments—the points in each cycle when representation is most unequal and equal, respectively.

There is substantial variation in the RRI both across states and within states over time. Following the 2000 census, the best-represented state, Rhode Island, enjoyed 70% more representation per capita than Montana, the worst represented state. Montana was not, however, always poorly represented. In fact, Montana was among the best-represented states

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

Table 2 Ln(RRI) before and after each reapportionment

Year	*Median*	*Range*	*10th percentile*	*90th percentile*
1972	0.01	0.75	−0.15	0.29
1973	0.00	0.65	−0.14	0.20
1982	0.01	0.94	−0.24	0.12
1983	0.02	0.53	−0.08	0.11
1992	0.04	0.64	−0.15	0.15
1993	0.00	0.58	−0.10	0.09
2002	0.03	0.78	−0.16	0.13
2003	0.00	0.60	−0.09	0.09

just before the reapportionment based on the 1990 census with a ln(RRI) of 0.35. After losing its second seat in the 1991 reapportionment, Montana's ln(RRI) fell to −0.34.

5.2 *Measuring Distribution*

Our measure of distributive spending is shares of total federal outlays. These data comprise a panel of repeated observations at the state level over 33 years (1970, 1972, 1974, 1976, 1977, and 1979–2006). Outlays to the states were published as *Federal Expenditures by State* starting in 1981, after which the reports were folded into the *Consolidated Federal Funds Report* series, both published by the U.S. Census Bureau (1995–1997, 1985–2007, respectively). We could not locate the original data for 1981 and 1982 and used the figures published in the *Statistical Abstract of the United States* (U.S. Census Bureau 1970–2007) for those years. Earlier data were published by the Office of Economic Opportunity (1970–1980) in an annual series, *Geographic Distribution of Federal Funds* (GDFF). We use these data when possible but supplemented them with data published in *The Almanac of American Politics* (Barone and Cohen 1972–1980) and *CQ Weekly Reports* (1970–1980) when the GDFF was unavailable. The data in the *Almanac*/*CQ* match those in the GDFF in years where we could locate both sources.

We use federal outlays to states for a number of reasons. First, previous analyses by Atlas et al. (1995) and Falk (2006) use this measure, allowing comparability. Second, although there are good theoretical reasons for analyzing more specific categories of federal funds (Lee 2003 offers a compelling argument for analyzing earmarks in particular), there is little agreement about the types of funding over which legislators have distributive discretion. For example, Lee (1998) finds that funds distributed on the basis of congressionally mandated formulas rather than discretionary funds are most affected by malapportionment. Levitt and Snyder (1995), on the other hand, classify as discretionary only those programs with relatively high coefficients of variation across districts, leading them to classify programs such as food stamps as discretionary but Medicare as nondiscretionary. Analyzing total federal outlays allows us to remain agnostic about the theoretical mechanism for what renders certain types of programs more discretionary than others. It also makes it less likely that we find a significant result since nondiscretionary portions of total outlays are likely to add significant noise to the outlays data. Finally, data on total federal outlays have been reported over multiple apportionment cycles allowing us to conduct analyses starting with the 1970 cycle.

Total federal outlays in 2004 amounted to $2.04 trillion. Total outlays include federal grants-in-aid to state and local governments, wages and salaries of military and civilian government employees, direct payments to individuals (e.g. Social Security, Medicare,

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

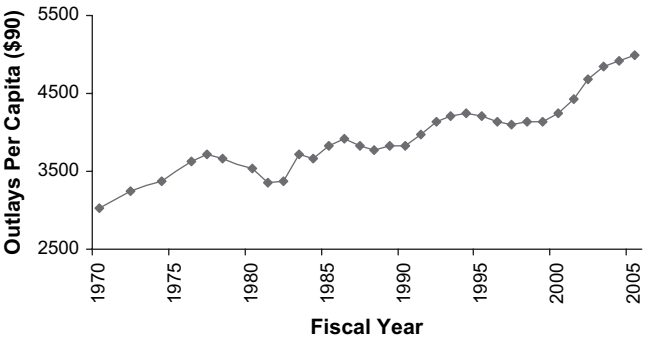


Fig. 5 Total federal outlays per capita (1990 dollars) by fiscal year.

Supplemental Security Income, and Food Stamps), federal procurement contracts, as well as support for other programs ranging from the National Endowment for the Arts to the Corporation for National Service. Direct payments accounted for 54% of the $2.04 trillion disbursed in 2004, by far the largest category. Grants-in-aid accounted for 19%, procurements for 15%, salaries and wages for 10%, and the remaining 2% were spent on other programs. The mean level of per capita outlays was $7,529 with an SD of $1633. At the bottom of the list, Nevada, Utah, and Minnesota each received less than $6,000 per capita, while Maryland, Virginia and Alaska topped the list, each receiving over $11,000 per capita.

Total outlays per capita increased in real terms by roughly 50% between 1970 and 2006 (Fig. 5). The time series also indicates a sharp kink in total outlays around the 1980 reapportionment. Since the outlays data for these years come from the *Statistical Abstract of the United States*, we were concerned that they may not be consistent with the data from other sources. Figure 6 illustrates this point by showing that the within-state correlations of per capita outlays between fiscal year t and fiscal year $t - 2$ are much lower for 1982 and 1983 than for any other set of years. Because comparable data across time for these years are essential for identifying the effects of the 1980 apportionment using the difference-in-first-differences technique, we restrict our analysis to the 1970, 1990, and 2000 reapportionments.[9]

In our estimations, we transform total federal outlays as follows: we take the first difference in difference in the natural log of a state's share of distributive spending. Note that this is approximately equal to the difference in the percent change in the share of distributive goods going to a state between time period t and time period $t - 2$ compared to analogous changes between time periods $t - 2$ and $t - 4$.

5.3 *Population*

All population data (other than apportionment populations) are from midyear estimates reported by the Bureau of Economic Analysis.

[9]The inconsistent data source for the 1980 apportionment cycle makes its inclusion problematic. Including data from the 1980 cycle produced dramatically different results depending on small changes in specification. For example, when $(f()_{s,1984} - f()_{s,1982}) - (f()_{s,1982} - f()_{s,1980})$ is used as the dependent variable, marginally significant *negative* effects of representation are estimated. In contrast, when using $(f()_{s,1985} - f()_{s,1983}) - (f()_{s,1983} - f()_{s,1981})$ as the dependent variable, large and significant positive effects of representation are estimated.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

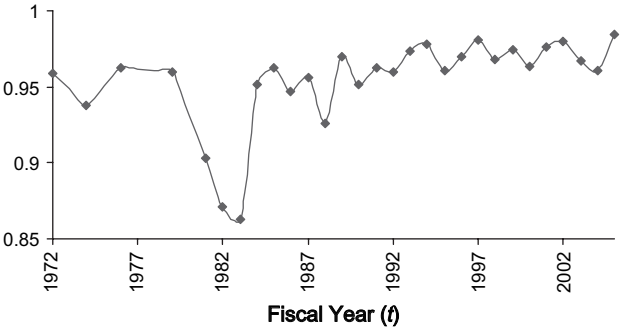


Fig. 6 Within-state correlation in per capita outlays at time t and $t - 2$.

6 Results

We find that gains in representation increase states' shares of federal outlays. Table 3 presents the estimates of equation (5).[10] As shown in column 1 of Table 3, the coefficient estimate of β in equation (1) is .067, which is statistically significant at the $p < .10$ level (two-tailed). This implies that a 10% increase in the number of representatives causes a 0.67% increase in a state's share of federal outlays. To put this effect size in perspective, Utah's representation in the House would have increased by approximately 30% if its apportionment population had been just slightly larger in 2000. Our point estimate suggests that receiving a fourth representative would have increased Utah's receipt of federal outlays by roughly $100 per capita in 2004.

One concern about equation (5) is that our difference-in-first-differences specification may not sufficiently account for changes in population. To alleviate these concerns, we estimate a variant of equation (5) that replaces difference in outlay shares with difference in population shares as the right-hand side variable. Our estimate of $\beta = -0.002$ is small and statistically insignificant ($p = .773$). This is unsurprising considering the smooth population shifts illustrated in Fig. 4.

To alleviate the concern that a few dramatic changes in outlays or representation in the sample may be driving the results, we plot the 53 switches that produce our estimate (Fig. 7). One particularly interesting observation is North Dakota in the 1970 cycle. This observation represents the largest decrease in representation, and consistent with our theoretical expectations, the largest decrease in per capita distribution. North Dakota did not experience a discontinuous shock in its population trend (Fig. 4: ''ND74''), increasing our confidence that its loss was not population driven.[11] We also estimate equation (5) using a least absolute deviations (LAD) regression, which limits the influence of high-leverage observations.[12] We continue to find a positive coefficient (see column 2 of Table 3),

[10] Standard errors are clustered by state.

[11] According to Fenno's (1992) study of Senator Mark Andrews, the halting of the Garrison Diversion water project significantly decreased appropriations to North Dakota in the late 1970s. It is possible that increased representation could have offset this setback.

[12] An LAD regression is just a special case of a quantile regression at the median. That is, it estimates the conditional median of the response variable rather than the conditional mean (as in OLS), thus making it robust to the influence of outlying cases.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

Table 3 Effect of changes in logged representation shares on first difference in changes in logged outlay shares (n = 53 switches)

Dependent variable (all in difference log first difference form) *Regression type*	*(1)* *Share outlays OLS*	*(2)* *Share outlays LAD*	*(3)* *Share outlays OLS*	*(4)* *Share outlays LAD*	*(5)* *Share outlays OLS*	*(6)* *Share outlays LAD*
$\ln(\text{rep}_{s,t}) - \ln(\text{rep}_{s,t-2})$	0.067 (0.039)	0.050 (0.059)	0.079 (0.042)	0.050 (0.063)	0.067 (0.072)	0.041 (0.098)
$\ln(\text{rep}_{s,t}) - \ln(\text{rep}_{s,t-2}) \times$ large state			−0.102 (0.087)	−0.087 (0.190)		
$\ln(\text{rep}_{s,t}) - \ln(\text{rep}_{s,t-2}) \times$ lose seat(s)					0.000 (0.100)	0.040 (0.130)

Note. Regressions also include year fixed effects. Standard errors clustered by state in parentheses.

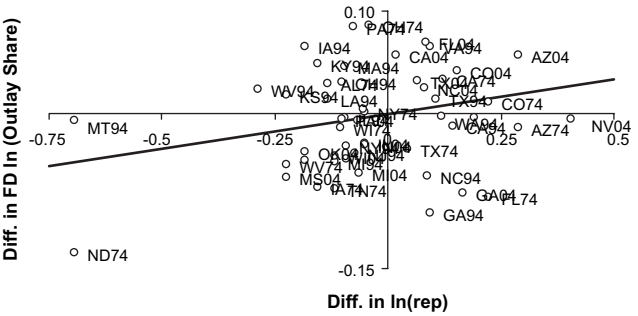


Fig. 7 Effect of changes in House representation on changes in federal outlays.

although the magnitude attenuates somewhat and the standard error increases.[13] Hence, when estimating a conditional median, the steepness of the slope is similar to OLS estimates of a conditional mean, but is less precisely estimated.

As discussed earlier, one advantage of our identification strategy is that the identifying source of variation occurs in large states as well as small states, which allows us to test for contextual effects. We find that the effects of representation on distribution are concentrated in the smaller states. We include an interaction term between change in representation and a dichotomous measure of whether the state is big (see column 3 of Table 3).[14] We find the effects to be solely concentrated in small states, where a 10% increase in representation is associated with a 0.8% increase in per capita outlays ($\beta = .079$, $p < .10$). Conversely, in large states, the effect of representation is nearly 0.[15] The importance of representation on Congressional committees is one plausible explanation for this effect since gaining an extra representative means a lot more to a small state like Utah than it does to a large state like California, which already has a representative on almost every committee. Finally, we find no evidence of an asymmetric effect for gains versus losses in representation (see column 5 of Table 3).

6.1 *Alternative Specification*

Following Ansolabehere, Gerber, and Snyder (2002), we construct an alternate population-normalized measure of outlays to use as the dependent variable. The relative outlays index (ROI) is constructed in an analogous manner to the RRI, where $\mathrm{ROI}_{s,t} = \frac{\mathrm{outlays}_{s,t}/\mathrm{pop}_{s,t}}{\mathrm{outlays}_{\bullet,t}/\mathrm{pop}_{\bullet,t}}$. To make this specification comparable to our main model, we include an indicator variable, Switch $_{s,t}$, that is equal to 1 if $\mathrm{rep}_{s,t} \neq \mathrm{rep}_{s,t-2}$ and 0 otherwise.[16] We then estimate the

[13] Another approach to eliminating the influence of outliers is to conduct a nonparametric sign test evaluating the pattern of results displayed in Fig. 7. We tested whether the observations were more likely to fall in quadrants I and III as opposed to the null of an equal distribution. We find that 31 of the 53 observations are in these two quadrants and reject the null of a homogenous distribution at $p = .136$. For the alternative specification (see Section 6.1), 33 of the 53 observations are in these two quadrants and we reject the null of a homogenous distribution at $p = .049$.

[14] We define the big states as those with at least 15 representatives in every Congress since the 1970 apportionment (e.g. California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas). The result is robust to other definitions of big.

[15] Again, similar patterns hold when estimating a LAD regression, although the coefficients are slightly smaller and the standard errors larger.

[16] Nonswitchers drop from the main specification of equation (5) by construction via differences.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

following regression model via ordinary least squares, clustering standard errors at the state level:

$$\begin{aligned}&\left[\ln\left(\mathrm{ROI}_{s,t}\right)-\ln\left(\mathrm{ROI}_{s,t-2}\right)\right]-\left[\ln\left(\mathrm{ROI}_{s,t-2}\right)-\ln\left(\mathrm{ROI}_{s,t-4}\right)\right]\\&=\delta\left(\left[\ln\left(\mathrm{RRI}_{s,t}\right)-\ln\left(\mathrm{RRI}_{s,t-2}\right)\right]-\left[\ln\left(RRI_{s,t-2}\right)-\ln\left(\mathrm{RRI}_{s,t-4}\right)\right]\right)\mathrm{Switch}_{s,t}+\epsilon_{s,t}.\end{aligned} \tag{6}$$

The alternative specification produces slightly more precise estimates than our main model (Table 4). The point estimate of δ in equation (6) is 0.068, which is nearly identical to the β estimated in equation (4) (see column 1). Substantively, a 10% increase in the number of representatives causes a 0.68% increase in a state's ROI. Again, the point estimate becomes somewhat smaller when estimating an LAD regression model (see column 2). The interaction effects also confirm that small states experiencing switches drive the result. As shown in the third column, a 10% increase in representation within small states is associated with a 0.81% increase in per capita outlays ($p < .05$). This effect is somewhat smaller but still approaches statistical significance when estimating the LAD model ($p < .12$). Finally, as shown in the last two columns, we do not observe any asymmetries between states that gained and lost seats.

7 Conclusion

Fairness in representation, embodied in the concept of one person, one vote, is one of the most important normative standards against which democratic institutions are measured. Analyzing the impact of fairness empirically, however, is made difficult by the challenge of isolating the distributional effects of representation from those caused by underlying demographic and economic conditions. In this paper, we have developed a method of identifying the causal effects of representation per se that can be exported across time and space. Moreover, contrary to previous analyses that have used apportionment formulas as instruments, we have developed a differences-in-first-differences estimator that alleviates concerns about small samples and local treatment effects. Scholars of comparative politics can apply the method in any democratic setting where census-based apportionment is carried out regularly and nonstrategically. For instance, decennial censuses in the American style are conducted in several nations including Austria, India, and Greece, among others.

One limitation of our difference-in-first difference estimator is that some effects of representation might not occur immediately following reapportionment. Our methodology could be extended to look at differences in the growth rates of spending in periods before and after switches. We do not do so in this paper because we only currently have data from a consistent source for multiple periods in the pre- and post-apportionment period for the 1990 reapportionment.

Our methodology could also be used to predict other dependent variables of substantive interest. For instance, we can explore whether committee representation or federal appointments are affected by having more members representing a state in Congress. Alternatively, we can explore if increased representation improves the reelection of members from a state's entire delegation or has positive spillover effects in the electoral prospects of governors and state legislators. Finally, we can use apportionment cycles as an exogenous source of variation in distribution to analyze the effects of money on a host of outcomes, including taxation and spending (i.e., ''the flypaper effect''), economic growth, and the electoral benefits of particularistic spending.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

Table 4 Effect of changes in logged RRI on first difference in changes in logged ROI ($n = 53$ switches)

Dependent variable (all in difference log first difference form)	*(1)* *ROI*	*(2)* *ROI*	*(3)* *ROI*	*(4)* *ROI*	*(5)* *ROI*	*(6)* *ROI*
Regression type	*OLS*	*LAD*	*OLS*	*LAD*	*OLS*	*LAD*
$\ln(RRI_{s,t}) - \ln(RRI_{s,t-2}) \times$ switch	0.069 (0.036)	0.042 (0.047)	0.081 (0.038)	0.063 (0.039)	0.073 (0.071)	0.037 (0.084)
$\ln(RRI_{s,t}) - \ln(RRI_{s,t-2}) \times$ switch $\times$ large state			−0.119 (0.089)	−0.054 (0.115)		
$\ln(RRI_{s,t}) - \ln(RRI_{s,t-2}) \times$ switch $\times$ lose seat(s)					−0.006 (0.096)	0.045 (0.110)

Note. Regressions also include year fixed effects. Standard errors clustered by state in parentheses.

Appendix: State Population and Seats Shares, 1935–2007

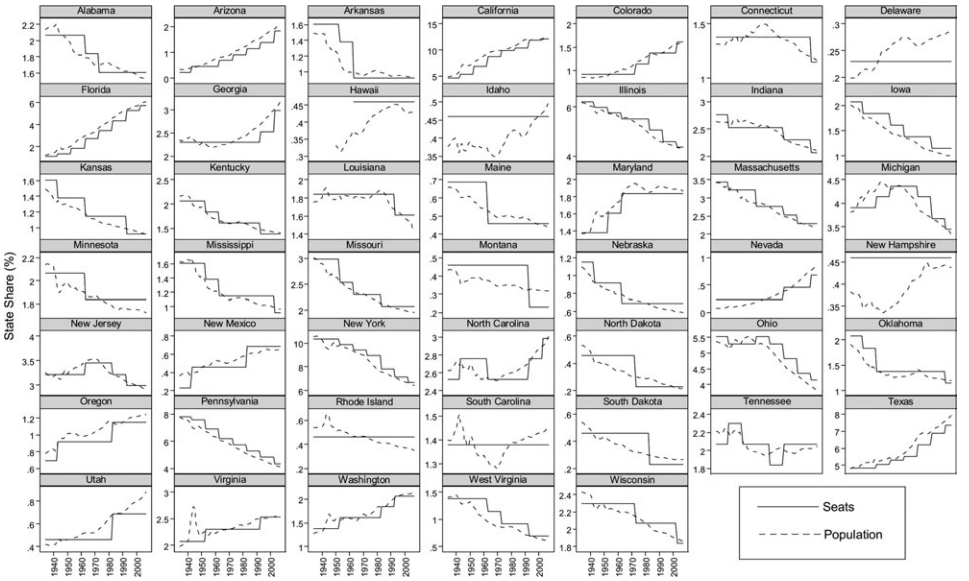


Fig. A1 Representation and Population U.S. House of Representatives, 1935–2007.

References

Alesina, A., and E. Spolaore. 1997. On the number and size of nations. *Quarterly Journal of Economics* 112:1027–56.

Angrist, J. D., and V. Lavy. 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics* 114:533–75.

Ansolabehere, S., A. Gerber, and J. M. Snyder. 2002. Equal votes, equal money: court-ordered redistricting and public expenditures in the American states. *American Political Science Review* 96:767–77.

Ansolabehere, S., J. M. Snyder, and M. M. Ting. 2003. Bargaining in bicameral legislatures: when and why does malapportionment matter? *American Political Science Review* 97:471–81.

Atlas, C. M., T. W. Gilligan, R. J. Hendershott, and M. A. Zupan. 1995. Slicing the federal government net spending pie: who wins, who loses, and why. *American Economic Review* 85:624–29.

Balinski, M. L., and H. P. Young. 2001. *Fair representation: meeting the ideal of one man, one vote*. Washington, DC: Brookings Institution Press.

Banks, J. S., and J. Duggan. 2000. A bargaining model of collective choice. *American Political Science Review* 94:73–88.

Baron, D., and J. Ferejohn. 1989. Bargaining in legislatures. *American Political Science Review* 83:1181–206.

Barone, M., and R. E. Cohen. 1972–1980. *The Almanac of American Politics*. Washington, DC: National Journal Group.

Chen, J., and N. Malhotra. 2007. The law of k/n: the effect of chamber size on government spending in bicameral legislatures. *American Political Science Review* 101:657–76.

Congressional Quarterly Weekly Report. 1970–1980.

David, P. T., and R. Eisenberg. 1961. *Devaluation of the urban and suburban vote*. Charlottesville, VA: University of Virginia Press.

Falk, J. 2006. *The effects of congressional district size and representative's tenure on the allocation of federal funds*. University of California, Berkeley, Manuscript.

Fenno, R. F. 1992. *When incumbency fails: the senate career of Mark Andrews*. Washington, DC: Congressional Quarterly Press.

Fiorina, M. P. 1989. *Congress: keystone of the Washington Establishment*. New Haven, CT: Yale University Press.

Hauk, W., and R. Wacziarg. 2007. Small states, big pork. *Quarterly Journal of Political Science* 2:95–106.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press

Hirano, S. 2006. Electoral institutions, hometowns, and favored minorities: evidence from Japan's electoral reforms. *World Politics* 59:51–82.
Hirano, S., and M. M. Ting. 2008. *Direct and indirect representation*. Columbia University, Manuscript.
Hoover, G., and P. Pecorino. 2005. The political determinants of federal expenditure at the state level. *Public Choice* 123:95–113.
Imbens, G. W., and J. D. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62:467–75.
Knight, B. G. 2008. Legislative representation, bargaining power, and the distribution of federal funds: evidence from the US senate. *Economic Journal* 118:1785–803.
Larcinese, V., L. Rizzo, and C. Testa. 2007. *Do small states get more federal monies? myth and reality about the US Senate malapporionment*. London School of Economics, Manuscript.
Lee, F. E. 1998. Representation and public policy: the consequences of senate apportionment for the geographic distribution of federal funds. *Journal of Politics* 60:34–62.
———. 2000. Senate representation and coalition building in distributive politics. *American Political Science Review* 94:59–72.
———. 2003. Geographic politics in the us house of representatives: coalition building and distribution of benefits. *American Journal of Political Science* 47:714–28.
Levitt, S. D., and J. M. Snyder. 1995. Political parties and the distribution of federal outlays. *American Journal of Political Science* 39:958–80.
Mansbridge, J. 2003. Rethinking representation. *American Political Science Review* 97:515–28.
Mayhew, D. 1974. *The Electoral Connection*. New Haven, CT: Yale University Press.
Office of Economic Opportunity. 1970–1980. *Geographic distribution of federal funds*. Washington, DC: U.S. Government Printing Office.
Snyder, J. M., M. M. Ting, and S. Ansolabehere. 2005. Legislative bargaining under weighted voting. *American Economics Review* 95:981–1004.
Ullah, A. 2004. *Finite sample econometrics*. Oxford: Oxford University Press.
U.S. Census Bureau. 1985–2007. *Consolidated federal funds report*. Washington, DC: U.S. Census Bureau.
———. 1995–1997. *Federal expenditures by state*. Washington, DC: U.S. Census Bureau.
———. 1970–2007. *Statistical abstract of the United States*. Washington, DC: U.S. Census Bureau.
Weingast, B. R., K. A. Shepsle, and C. Johnsen. 1981. The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics. *Journal of Political Economy* 89:642.
Woon, J. 2007. Direct democracy and the selection of representative institutions: voter support for apportionment initiatives, 1924–1962. *State Politics and Policy Quarterly* 7:167–86.

https://doi.org/10.1093/pan/mpp012 Published online by Cambridge University Press