LEGISLATIVE REPRESENTATION, BARGAINING POWER AND THE DISTRIBUTION OF FEDERAL FUNDS: EVIDENCE FROM THE US CONGRESS*

Brian Knight

This article investigates the relationship between representation in legislatures and the geographic distribution of federal funds. In a legislative bargaining model, we demonstrate that funds are concentrated in high representation areas, and two channels underlie this result. The proposal power channel reflects the role of representation in committee assignments, and the vote cost channel reflects the role of representation in coalition formation. In our empirical analysis, we find that small states, relative to large states, receive more funding in the US Senate, relative to the House. We also find empirical support for the two channels underlying this relationship.

Throughout the history of federations and international organisations, a key constitutional design question involves how jurisdictions of different sizes are represented in legislative bodies and how their votes will ultimately be weighted. Large jurisdictions typically argue for delegation sizes or voting weights that are roughly proportional to population whereas small jurisdictions typically argue for an equal number of delegates or votes for each jurisdiction. This conflict over representation is presumably driven, at least in part, by a belief that increased representation strengthens bargaining power in the battle over the geographic distribution of funds authorised by these legislative bodies.

In this article, I analyse the relationship between legislative representation, bargaining power and the geographic distribution of federal funds from both a theoretical and empirical perspective. In addition to investigating the existence of a relationship between representation and federal funds, I also explore the underlying mechanisms. That is, I ask which political institutions, such as agenda control provided to committees and the formation of coalitions under majority rule, shape the bargaining power of jurisdictions in central legislatures and thus facilitate the relationship between representation and the distribution of federal spending.

To provide a framework for analysing these issues, I begin by developing a simple theoretical model of a central legislature that determines the cross-jurisdiction distribution of local public goods, which are financed by a national tax base. Both delegation sizes and population are variable, allowing me to consider a variety of institutional arrangements for jurisdictional representation. Legislators represent local interests, and, given the common pool problem associated with concentrated benefits and dispersed tax costs, attempt to increase own-jurisdiction spending but to restrain spending elsewhere. The model follows a recent literature on legislative bargaining, in which a proposer must secure a majority of votes in order to pass legislation and, due to the

* Thanks for helpful comments to David Baron, Dhammika Dharmapala and Jim Snyder and to participants at the Public Choice Society Conference, the Constitutional Design Conference, the Warwick/IFS Public Economics Weekend, the Research Group on Political Institutions and Economic Policy, the University of Pennsylvania, Duke University and Columbia University.
common pool problem, has an incentive to form the cheapest possible coalition. Under reasonable conditions, I show that increases in representation lead to increases in jurisdiction spending. Moreover, this relationship is a result of two underlying channels. The proposal power channel reflects the fact that increased representation improves the likelihood of being represented as the proposer. To the extent that proposal power is valuable to legislators in securing disproportionate spending for their home jurisdictions, it is clear that jurisdictions will benefit from this increased control over the agenda. The second channel, the vote cost channel, reflects the fact that increases in representation make jurisdictions more attractive from the perspective of proposers attempting to form the cheapest possible coalition.

In addition to generating these testable predictions, the theoretical model provides insights regarding the identification of the effect of representation on the geographic distribution of public spending in empirical analyses. In particular, cross-jurisdiction methods, as typically employed in the existing literature, may not be able to identify the effect of representation separately from the effects of population and other state-level characteristics. To overcome this limitation associated with cross-jurisdiction analyses, I conduct a within-state, cross-chamber analysis using a new dataset that identifies projects as originating in the US House, where representation is roughly proportional to population, or in the US Senate, where all states have two delegates. Thus, I examine whether small states receive more funding in the Senate, relative to the House, and whether large states receive more funding in the House, relative to the Senate. I also investigate empirically the two theoretically-identified channels underlying the relationship between representation and federal spending.

1. Related Literature

Much of the related empirical work has focused on the effect of representation in the US Senate on the cross-state distribution of federal funds. Atlas et al. (1995) report a statistically significant relationship between per capita federal spending, as measured by $G_s/N_s$ where $G_s$ is federal spending located in state $s$ and $N_s$ is state population, and per capita representation in the US Senate, as measured by $2/N$. This correlation is interpreted as evidence of a small state bias associated with Senate representation. In a related analysis, Lee (1998) finds evidence of a small state advantage using data from the US Domestic Assistance Programs Database, as compiled by Bickers and Stein (1991). As noted above, these cross-jurisdiction analyses rely solely on variation in population across states and may not be able to identify the effects of representation separately from the independent effects of population. This idea will be developed more fully below in the context of the theoretical model.

Several papers have examined the effect of representation on the geographic distribution of public spending in other legislative settings. Rodden (2002) finds a connection between votes per capita and net transfers per capita in the EU. Ansolabehere et al. (2002) find that Baker v. Carr, which required that state legislative districts be re-apportioned every ten years, significantly altered the flow of state transfers to county

---

1 These data contain federal spending by Congressional districts, although the author aggregates up to the state level in order to focus on differences in per capita Senate representation.
governments. The authors calculate that, as a result of this ruling, approximately $7 billion were diverted annually from formerly over-represented counties, which were primarily rural, to formerly under-represented counties, which were primarily urban.

More broadly, this article is related to a growing empirical political economy literature on common pool problems in central legislatures. Del Rossi and Inman (1999) find that the size of federal water project requests fell significantly when cost sharing requirements for local governments increased. Knight (2004) finds that legislators are more likely to vote in favour of federal spending the higher the fraction of funds spent in their own district but the lower the tax liability associated with spending elsewhere. Baqir (2002) finds a positive relationship between the size of city councils and public spending, and Gilligan and Matsusaka (1995) and Bradbury and Crain (2001) find similar results for legislatures in US states. These empirical results relating to the size of the legislature are consistent with the theoretical model of Weingast et al. (1981), who argue that spending will increase in the number of legislative districts given that each legislator only internalises tax costs for their own constituents. Several studies have also examined how the cross-state conflict associated with the common pool problem is ultimately resolved. Knight (2005) finds that committee members use their proposal power to secure disproportionate transportation projects for their home districts. Levitt and Snyder (1997) find that Congressional districts with a large fraction of Democratic voters received significantly more federal funds between 1984 and 1990, a period in which Democrats controlled the US House. Levitt and Poterba (1999) find that states with powerful representatives experience more rapid economic growth than do other states; this effect cannot be explained, however, via higher federal spending.

This article also contributes to a growing literature, primarily theoretical, on legislative bargaining. Recent applications of this legislative bargaining model include comparative politics (Persson et al., 2000), federalism (Besley and Coate, 2003; Lockwood, 2002), intergovernmental transfers (Knight, 2002), legislative elections (Chari et al., 1997; Coate, 1997), legislative seniority (McKelvey and Riezman, 1992), pork barrel inefficiencies (Baron, 1991), social choice (Banks and Duggan, 2000), special interest politics (Persson, 1998; Helpman and Persson, 2001; Bennedsen and Feldmann, 2001), tax expenditures (Dharmapala, 1999) and public investment goods (Leblanc et al., 2000).

2. Theoretical Framework

Here I present the simple theoretical model underlying my empirical analysis. Given my empirical motivation, I keep things simple and use specific parametric forms and distribution functions where appropriate. The role of the assumptions will be examined at the end of this Section.

2.1. Economic Environment

Consider a collection of $S$ states, or jurisdictions, indexed by $s$, with a total national population equal to $N$. In a given state $s$, each of $N_s$ residents is assumed to have the following quasilinear preferences over consumption of a statewide public good ($G_s$) and consumption of a private good ($z_s$):
where the $h$ is a strictly increasing function and is normalised so that zero utility is received from zero spending [$h(0) = 0$]. The congestion parameter $\gamma \in [0, 1]$ captures the degree of rivalry in consumption; this specification nests the case of private goods ($\gamma = 1$) as well as the case of pure public goods ($\gamma = 0$). Finally, I assume that each resident in state $s$ is endowed with $m_s$ units of the private good, which can be converted into public goods at a dollar-for-dollar rate.

2.2. Political Environment

Following a literature on legislative bargaining, I assume that a central legislature determines the cross-state distribution of local public goods from a fixed budget, which I normalise to one. Thus, federal spending ($G_s$) can be interpreted as the fraction of the federal budget located in state $s$. These public goods are financed from a national, or common pool, tax base, and each resident pays an equal fraction 1/$N$. Finally, private consumption is determined residually and equals $z_s = m_s - 1/N$ in the event that federal spending is approved in the legislature.

The legislature is composed of delegates who represent the interests of individual states and I normalise the total number of delegates to one. Thus, the delegation size of state $s$, given by $V_s$, can be interpreted as a fraction of the legislature affiliated with delegation $s$. I allow for delegation sizes to be variable, and the model thus nests representation in the US Senate, in which each state has the same number of delegates ($V_s = 1/S$), as well as representation in the US House, in which representation is proportional to population ($V_s = N_s/N$).

The political process is modelled as a one-shot version of the legislative bargaining model due to Baron and Ferejohn (1987, 1989) and Persson and Tabellini (2002). In the first stage, a single delegate is chosen to be the agenda-setter, which I alternatively refer to as the proposer. Let $A_s \in \{0, 1\}$ indicate whether state $s$ is represented by the agenda-setter and let $\Pr(A_s = 1)$ denote this probability. While I allow this probability to be general for now, note that it is equal to the fraction of the legislature represented by state $s$ if each delegate is selected with equal probability [i.e. $\Pr(A_s = 1) = V_s$].

In the first stage of the game, the agenda-setter proposes a geographic distribution of the federal budget, subject to a balanced budget condition ($\Sigma_s G_s \leq 1$). In the second stage, each delegate votes over whether to accept or reject the proposed budget. If the proposal receives a majority of votes from delegates in support, it is implemented and taxes are levied; otherwise, a zero reversion budget is implemented. Projects are funded via a national tax base and each citizen is assumed to pays taxes equal to 1/$N$. Thus, while project benefits are concentrated within states, costs are dispersed; this common pool problem will be a key component of the model.

2.3. Subgame Perfect Equilibrium

Working backwards, each delegate will support proposals for which the total benefits accruing to a representative constituent exceed the tax costs associated with aggregate provision:

© The Author(s). Journal compilation © Royal Economic Society 2008
Given that, within a state, delegates use equivalent voting rules, the cost \((C_s)\) to the proposer of securing votes from all of the delegates of state \(s\), the spending level \(G_s\) that sets (2) to binding, can be expressed as follows:

\[ G_s = \theta N_s^\gamma \]  \hspace{1cm} (3)

where \(\theta\) is a constant that depends upon the function \(h\) as well as national population \((N)\). Thus, the cost of securing the votes of delegation \(s\) is increasing in state population so long as some congestion is present \((\gamma > 0)\). In the special case of pure public goods \((\gamma = 0)\), this cost is independent of population. Regardless of the degree of congestion, however, the cost of securing the votes from a given delegation is independent of its size \(V_s\).

In order to maximise the level of the public good provided to his home state, the agenda-setter has an incentive to form the cheapest possible coalition \((M)\). Let \(M_s \in \{0,1\}\) indicate inclusion in the coalition. In the context of this model, the agenda-setter, taking voting rules as given, selects delegates from those states with the highest vote yields, defined as the ratio of delegation sizes to the total cost of securing the votes from a given delegation:

\[ \frac{V_s}{C_s} = \frac{V_s}{\theta N_s^\gamma}. \]  \hspace{1cm} (4)

As shown, holding population constant, vote yields are proportional to delegation sizes. In order to smooth the probability of inclusion in the coalition, I assume that, in addition to selecting states with vote yields above some cutoff \(\eta\), the agenda-setter experiences a mean-zero cost equal to \(\xi_v\), which has a distribution function \(F\) from including state \(s\) in the coalition. This cost could represent, for example, a distaste for including members of the opposition party, an issue that I will examine empirically below. Including this additional cost, the payoff to the proposer from including state \(s\) equals \(V_s/C_s - \eta - \xi_v\), and the probability of inclusion in the coalition is thus given by:

\[ \Pr(M_s = 1) = F\left(\frac{V_s}{C_s} - \eta\right). \]  \hspace{1cm} (5)

To summarise, the agenda-setter forms a minimum-sized coalition of states, who tend to have high vote yields and who receive a payment equal to their cost \((C_s)\), which is independent of the delegation size. States excluded from the coalition receive no spending. Finally, the agenda-setter receives a residual equal to \(d\), which is independent of both state population and delegation size.

At the agenda-setter’s optimal proposal, expected spending \((G_s)\), unconditional on the identity of the agenda-setter, can be expressed as follows:

\[ E(G_s) = \Pr(A_s = 1)\delta + [1 - \Pr(A_s = 1)]\Pr(M_s = 1)C_s \]  \hspace{1cm} (6)

Taking the derivative of (6) with respect to representation, holding state population constant, we have that:

\[ \frac{\partial E(G_s)}{\partial V_i} = \frac{\partial \Pr(A_i = 1)}{\partial V_i} \left[ \delta - \Pr(M_s = 1)C_s \right] + \frac{\partial F(V_s/C_s - \eta)}{\partial V_i} \left[ 1 - \Pr(A_s = 1) \right]C_s \]  \hspace{1cm} (7)

Proposal-power channel \hspace{1cm} Vote-cost channel
Thus, so long as proposal power is valuable \( (\delta > \Pr(M_s = 1) C_s) \), there is a positive relationship between representation and expected federal spending. Moreover, this relationship can be decomposed into two underlying channels. The proposal power channel reflects the fact that increases in representation may lead to increased representation as the agenda-setter. Indeed, this partial derivative equals one if each delegate is selected with equal probability (i.e. \( \Pr(A_s = 1) = V_s \)). To the extent that proposal power is valuable, \( [\delta > \Pr(M_s = 1) C_s] \), it follows that federal spending is increasing in representation due to this proposal power channel.\(^2\) The vote-cost channel reflects the fact that, holding constant population, increases in representation improve the attractiveness as a potential coalition member from the perspective of the agenda-setter attempting to form an optimal coalition.\(^3\) This channel is increasing in the probability that the state is not represented as the agenda-setter \( [1-\Pr(A_s = 1)] \) and in the ultimate payment for the state conditional on inclusion in the coalition \( (C_s) \).

2.4. Role of the Assumptions

This Section discusses the role of key assumptions and the robustness of the results to alternative assumptions. The first of these assumptions, statewide public goods, is made for tractability. In particular, this assumption guarantees that delegations vote cohesively and thus the agenda-setter receives the votes of all delegates so long as spending is at least as high as the state-specific cost \( (G_s \geq C_s) \). An alternative model would allow for states to be sub-divided into legislative districts and for proposers to be able to channel funds to specific districts within a state. It is clear that the proposal power channel will continue to operate if \( \Pr(A_s = 1) \) is increasing in \( V_s \). I conjecture that the cost-channel will also continue to operate so long as some congestion is present. To see this, note that the cost of securing the vote of a single delegate from a state divided into \( D \) equally-sized legislative districts equals \( \theta(N_s/D) \) and the cost of securing a single vote is thus decreasing in the number of districts \( D \) if \( \gamma > 0 \). While it seems that my model is robust to local public goods, I choose to retain my assumption of statewide public goods for two reasons. First, my data has no geographic information below the level of the US state. Second, my analysis is motivated by the battle over representation and funds across, rather than within, states.

The second issue involves the result that the cutoff for inclusion in the coalition is independent of the identify of the agenda-setter; this result in turn guarantees that the

\(^2\) Knight (2005) finds that Congressional districts represented on the transportation authorisation committee received significantly more federal funds in earmarked transportation projects than did districts not represented on the committee.

\(^3\) Note that the vote-cost channel operates regardless of the degree of congestion. Under the special case of pure public goods \( (\gamma = 0) \), vote yields are proportional to delegation sizes \( (V_i/C_s = V_s/\theta) \) and are independent of population under Senate-type representation but are higher for large states under House-type representation. Thus, increases in representation lead to an increased probability of inclusion in the coalition. Under the special case of pure private goods \( (\gamma = 1) \), by contrast, vote yields are proportional to delegates-per capita \( (v_i = V_i/N_s) \) and are thus independent of population under House-type representation but are higher for small states under Senate-type representation. Again, increases in representation make jurisdictions more attractive from the perspective of the agenda-setter.
agenda-setter’s share $\delta$ is homogeneous across states. This constant cutoff may be violated by integer problems in my model and I have abstracted from this issue above.\footnote{For example, suppose that an agenda setter in the US Senate needs to secure the votes of 24 other states and finds it optimal to choose the smallest states. Then, an agenda setter from a small state will include the 25th smallest state, whereas an agenda setter from a large state will exclude this 25th smallest state.} This problem of a non-constant cutoff for inclusion in the coalition is related to the problem of non-homogeneity described by Snyder et al. (2005). Coalitions are said to be homogeneous if all minimum winning coalitions have the same number of votes. They address this problem of non-homogeneity by multiplying the number of players of each type, or state in our case, by some positive integer $r$, which they refer to as the number of replications; my baseline model above has $r = 1$. They then show that the problem of non-homogeneity grows small as $r$ increases.

The third issue involves the assumption of a default of zero public spending in the event that a federal budget is not passed. There are two ways to relax this assumption. The first would allow for continued bargaining in the event of a failed vote. Snyder et al. (2005) examine a model with multiple rounds of bargaining, variable voting weights, and recognition probabilities that are proportional to voting weights. Although their model includes neither variation in population nor a congestion parameter, it predicts that expected payoffs are proportional to voting weights. The intuition is that votes are perfect substitutes and thus a player with $k$ votes should command a price for those votes equal to the total price of $k$ players with one vote each. The second way to relax the assumption of a zero default involves introducing an exogenous and state-specific reversion equal to $R_s$ and assuming that citizens pay taxes equal to $1/N$ regardless of the voting outcome. In this case, the cost of securing the votes of the delegation from state $s$ is equal to this reversion level and vote yields are thus equal to $V_s/R_s$. Regarding the two underlying channels, it is clear that the proposal power channel will continue to operate under this alternative model and the vote-cost channel should as well, so long as reversion levels are independent of the number of delegates. While it seems that my model is robust to alternative assumptions regarding reversion levels, I choose to retain my assumption of a reversion of zero spending for two reasons. First, my data contain no information on reversion levels. Second, this assumption seems natural given my empirical analysis of spending on new projects, as opposed to incremental spending on existing programmes.

Finally, I assume that tax prices are equal to $1/N$ and are thus both independent of project spending and homogeneous across states. Regarding the first issue, it is straightforward to introduce cost-sharing by the state such that the federal government pays a fraction $f$ of project costs with the state paying a fraction $1 - f$. In this case, delegates support legislation if $h(G_s/N_s) \geq 1/N + (1 - f)G_s/N_s$. Again, the proposal power channel continues to operate, and I conjecture that the vote cost channel will continue to operate as well. Although one cannot solve explicitly for costs ($C_s$) with $f > 0$, it is clear that this cost continues to be independent of delegation sizes. Regarding the second issue, the homogeneity assumption, it is straightforward to introduce a state-specific per capita tax price $p_s$, which are assumed to sum to one nationally, and delegates now support the proposal if $h(G_s/N_s) > p_s$. Thus, the costs of securing the votes from a given delegation are increasing in $p$, as high-tax states must be compensated accordingly.

\cite{Snyder2005}
2.5. Implications for Empirical Analysis

In addition to formalising the theoretical relationship between representation and public spending, the baseline model also provides insights for existing cross-state analyses focusing on the US Senate, which attributes to Senate representation the finding that small states receive more funding on a per capita basis. In particular, the model demonstrates that such a cross-state strategy is hampered generally by the role of state-level characteristics and specifically by the role of state population. To see this, consider a special case of the above model in which the probability of selection as agenda-setter is equal to delegation sizes (Pr(A_s = 1) = V_s) and in which ξ, is uniformly distributed, in which case the probability of inclusion in the coalition is linear [F(V_s/C_s−η) = β + τV_s/C_s]. In this special case, expected per capita federal spending can then be written as follows:

\[ E(g_s) = v_sδ + (1 - V_s)(βc_s + τv_s) \]  

where \( c_s = θN_s^{γ−1} \) represents the per capita cost of securing the votes from the delegation representing state \( s \). As can be seen, state population exerts an independent effect upon per capita federal spending through \( c_s \) so long as congestion is incomplete (\( γ < 1 \)). Thus, it is difficult to disentangle the effects of Senate representation from the effects of population using only cross-state variation. More generally, cross-state analyses are limited by the role of unobserved state-level factors. For example, a finding that small states receive more federal spending on a per capita basis could be due to any state-level factor that is correlated with population.

To overcome these limitations associated with existing studies, I focus on cross-chamber variation. That is, I exploit the bicameral feature of the US Congress and compare federal spending originating in the US Senate, relative to federal spending originating in the US House. This allows me to difference out all state-level characteristics, including population, that do not vary at the state level between the House and Senate chambers. To the extent that representation matters, I then expect small states to receive relatively more Senate funds and for large states to receive relatively more House funds.

3. Empirical Analysis

In this Section, I provide an empirical analysis of the relationship between representation and the geographic distribution of federal spending. This analysis focuses on a cross-chamber analysis of House-sponsored projects, under which large states are predicted to receive relatively more spending, versus Senate-sponsored projects, under which small states are predicted to receive relatively more spending. I also investigate the two theoretically-identified channels underlying the relationship between representation and federal spending.

3.1. Data

The main source of data on federal spending is provided by the Citizens Against Government Waste (CAGW), a private, non-partisan, non-profit organisation that has...
catalogued funds for projects that were earmarked in annual appropriations bills. As shown in Table 1, Congressional earmarks are an important component of federal spending; the 9,362 projects listed in the fiscal year 2003 appropriations bills totalled $23 billion. Moreover, as shown, these appropriations earmarks have increased over time, rising from about $13 billion in fiscal year 1995.

For each of the 37,336 projects catalogued between 1995 and 2003, data include

(i) the state in which the project is to be located,
(ii) the dollar amount appropriated,
(iii) relevant appropriations bill (Congress appropriates funds through 13 separate bills each fiscal year), and
(iv) the sponsoring chamber (i.e. the chamber, House or Senate, responsible for the appropriation).

For the purposes of the econometric analysis, I focus on the subset of projects for which the relevant state is listed; the remaining analysis thus ignores those projects that were not targeted to a specific state. To provide a sense of the types of projects funded, Table 2 lists the largest projects, those with appropriations in excess of $60 million in 2003 dollars. As shown, the three largest projects were all included in the Defense appropriations bill, were sponsored by the Senate, and were located in Mississippi, a relatively small state and also the home state of Sen. Trent Lott, who served as Senate Majority Leader during the period in which the projects were funded. The remaining projects listed span a wide variety of states, appropriations bills and sponsoring chambers. To provide a broader sense of spending categories, Table 3 details the number of projects and total funding by appropriations bill over this time period. As shown, military construction and transportation appropriations bills provided the most funding, followed by VA/HUD, defence, and energy. Table 4 provides a breakdown of project spending by sponsoring chamber, one of the key

<table>
<thead>
<tr>
<th>Fiscal year</th>
<th>Number of projects</th>
<th>Appropriations (billions of $2003)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>1,439</td>
<td>$12.8</td>
</tr>
<tr>
<td>1996</td>
<td>958</td>
<td>$10.1</td>
</tr>
<tr>
<td>1997</td>
<td>1,596</td>
<td>$16.3</td>
</tr>
<tr>
<td>1998</td>
<td>2,143</td>
<td>$14.6</td>
</tr>
<tr>
<td>1999</td>
<td>2,838</td>
<td>$15.1</td>
</tr>
<tr>
<td>2000</td>
<td>4,926</td>
<td>$18.6</td>
</tr>
<tr>
<td>2001</td>
<td>6,333</td>
<td>$19.0</td>
</tr>
<tr>
<td>2002</td>
<td>8,341</td>
<td>$20.0</td>
</tr>
<tr>
<td>2003</td>
<td>9,962</td>
<td>$22.5</td>
</tr>
</tbody>
</table>

5 Many projects in the database fit into this category. For example, the fiscal year 2003 Senate Defense Appropriations bill earmarked $500,000 for M48 masks (Procurement - Defense Wide), and this project was thus not targeted to a specific state. While this earmarked project may clearly benefit some states more than others, it is not clear how the analyst would measure these cross-state differences for each project and I thus ignore these projects without a targeted state in the empirical analysis.

6 The District of Columbia, foreign operations, legislative and Treasury provided, perhaps not surprisingly, the least project funding given the difficulty of locating projects in a specific state in these cases.
variables in the econometric analysis as it allows us to identify the effects of representation separately from the effects of population. As shown, Senate projects provided the most funding, at $20 billion, followed by House projects, which funded just over $13 billion. Conference committees added a smaller amount and the CAGW did not classify an additional $9 billion in project spending (presumably due to missing or ambiguous data).

A sceptical reader might be justifiably concerned over possible biases in the database. While CAGW is non-partisan, they do appear to have a political agenda of reducing earmarks, and this agenda could arguably lead to biases in the database. While a complete audit of the CAGW database is beyond the scope of this analysis, I did conduct a case study of projects included in the 2003 Military Construction Appropriations Bill. Encouragingly, I found that all 191 projects listed in the database were

© The Author(s). Journal compilation © Royal Economic Society 2008
also detailed in the legislation. Regarding the converse, while almost all projects in the legislation were included in the database, I did find a couple of exceptions (i.e. projects listed in the legislation that were not included in the database). I suspect that these exceptions were due to differences in interpretation rather than a systematic bias in the construction of the database. Although limited to a single appropriations bill, this analysis is encouraging and increases our confidence in the quality of the database.

For the purposes of the econometric analysis, the federal spending measures are scaled in two ways. First, for comparability across chambers and years and for consistency with the theoretical model, which interpreted federal spending as the fraction of the budget located in each state, I scale the spending measure by total cross-state spending in each chamber and year. Second, some specifications use per capita measures and data on population by year from the Census Bureau are thus incorporated into the analysis. Although the Census is only conducted every 10 years, the Census Bureau provides annual estimates of population on July 1.

Finally, in order to gauge the importance of underlying channels, I also incorporate data on assignments to House and Senate appropriations subcommittees, each of which has agenda control over a single appropriations bill. Thus, while the theoretical model envisions a single proposer, agenda-setting powers in Congress are conferred...
upon sets of legislators, or committees. To make the committee measure consistent with the agenda-setter variable in the theoretical model, I compute the fraction of the committee represented by state $s$. That is, let $A$ denote an agenda-setting committee with $N_A$ members. Then, the fraction of the committee represented by state $s$ is given by $A_s = 1(s \in A) / N_A$, where $1(s \in A)$ indicates state representation on the committee. Intuitively, the power conferred upon each committee member is decreasing in the overall size of the committee.

3.2. Basic Results

Table 5 provides results of tests for the presence of relationship between representation and the geographic distribution of federal spending. In particular, we estimate the following simple linear difference equation relating federal spending to representation:

$$
\Delta G_{st} = \beta_0 + \beta_1 \Delta V_{st} + U_{st}
$$

(9)

where $s$ indexes states, $t$ indexes time. The key dependent variable is the Senate–House difference in spending ($\Delta G_{st} = G_{st}^{SENATE} - G_{st}^{HOUSE}$), while the key independent variable is the Senate–House difference in representation ($\Delta V_{st} = V_{st}^{SENATE} - V_{st}^{HOUSE}$). Finally, $U_{st}$ represents unobserved determinants of the geographic distribution of federal spending that differ within a state-year between the Senate and House.

Figure 1, which plots the Senate–House difference in spending by state (averaged over time) against the Senate–House difference in representation by state (averaged over time), provides a graphical sense of the identification strategy. As shown, California, the largest state, has roughly 12% of delegates in the House but only 2% in the Senate; the Senate–House difference in representation is thus roughly –10 percentage points. At the same time, California receives roughly 10 percentage points less of the fraction of the federal budget in the Senate, relative to the House. Small states, by contrast, have more representation in the Senate and, as shown, also tend to receive more project spending in Senate bills than in House bills.

Table 5

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senate–House spending difference</td>
<td>$1.0619^{**}$</td>
<td>$1.0440^{**}$</td>
<td>$1.2834^{**}$</td>
<td>$1.2704^{**}$</td>
</tr>
<tr>
<td>Senate–House per-capita spending difference</td>
<td>$1.0419^{**}$</td>
<td>$1.0440^{**}$</td>
<td>$1.2834^{**}$</td>
<td>$1.2704^{**}$</td>
</tr>
<tr>
<td>Observations</td>
<td>450</td>
<td>450</td>
<td>3,350</td>
<td>3,350</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.3044</td>
<td>0.2088</td>
<td>0.1189</td>
<td>0.0927</td>
</tr>
</tbody>
</table>

Notes. dependent variable is budget shares, std errors in parentheses, ** denotes 95% significance, * denotes 90%, constant not reported.

© The Author(s). Journal compilation © Royal Economic Society 2008
This relationship is also reflected in column 1 of Table 5. As shown, a one percentage point increase in representation, which could be achieved, for example, by giving a single state an additional Senator, is associated with an increase in the fraction of federal funds located in that state of roughly one percentage point. This effect is statistically significant and the 95% confidence interval is tight at [0.91,1.21]. For comparability with the existing empirical literature, which has examined the relationship between \textit{per capita} federal spending and \textit{per capita} Senate representation, Column 2 of Table 5 provides the results from a specification in \textit{per capita} terms:

\[
D_{gst} = b_0 + b_1 D_{vst} + u_{st}.
\] (10)

As shown, the results in column 2 are quite similar to those in column 1, and the baseline results are thus robust to scaling my key measures by state population.

Given that my underlying channels analysis to follow is at the level of the appropriations bill and appropriations subcommittee, I next conduct an analysis at that level of aggregation. In particular, I estimate the following equation:

\[
\Delta G_{sta} = \beta_0 + \beta_1 \Delta V_{st} + U_{sta}
\] (11)

where \(a\) indexes appropriations bills. As shown, my results are even stronger at this disaggregated level and the point estimates suggest that a one percentage point increase in representation is associated with an increase in the fraction of federal funds located in that state of 1.28 percentage points.\(^7\) As shown in column 4, the results from a corresponding \textit{per capita} analysis are similar in magnitude to those in column 3.

It is important to note that these results in Table 5 rest upon an implicit assumption of strategic independence across legislative chambers. Under this strategic independence

\(^7\) Given that there are 13 appropriations bills, this should yield a total sample size of 5,850. Many bills, such as the appropriations bill for the District of Columbia and Foreign Operations, however, have zero or negligible spending. After eliminating these bills, we are left with a matched sample of 3,350.
assumption, the House and the Senate independently develop an allocation of projects and these projects are then simply merged during the conference committee, which reconciles the House and Senate appropriations bills. Unfortunately, I cannot directly test this implicit assumption as projects that were removed during the conference committee are excluded from the CAGW data. As an alternative to this direct test, I conducted a case study of the fiscal year 2003 Military Construction appropriations bill. Based upon an analysis of projects included in the initial House bill, the initial Senate bill and the final conference bill, over 80% of projects included in the House and Senate bills were also included in the conference version. The large proportion of projects funded is consistent with this implicit assumption of strategic independence across chambers.\(^8\)

3.3. Underlying Channels

In order to provide further interpretation of these results, we next investigate the two underlying channels identified by the theoretical model. As shown in (7), the size of the proposal-power channel depends upon the relationship between the representation and selection as the agenda-setter \((\partial \Pr(A_s = 1)/\partial V_s)\) along with the premium associated with membership on the committee \((\delta - \Pr(M_s = 1)C_s)\). We investigate each of these two components of the proposal power channel in turn. To measure the effect of representation on committee assignments, we regress the Senate–House difference in representation on the Senate–House difference in committee assignments:

\[
\Delta A_{sta} = \beta_0 + \beta_1 \Delta V_{st} + U_{sta}. \tag{12}
\]

Figure 2 provides a plot of this relationship, where the horizontal axis is identical to that in Figure 1 and the vertical axis reflects the difference in the fraction of appropriations bills under which state delegations were assigned to the appropriations committee (averaged over time and appropriations bills). As shown, the relationship is strongly positive as large states are represented far more often on House committees and small states tend to be represented more often on Senate committees.

As shown in column 1 of Table 6, the coefficients from the corresponding regression demonstrate that the relationship is positive and statistically significant. Although the coefficient is statistically different from one and I can thus reject proportional assignments \((\Pr(A_s = 1) = V_s)\), it is large at 0.8 and clearly different from zero. Of course, increased likelihood of assignment to the committee will translate into higher federal spending only if committee members tend to receive more funds than do other states. To investigate this issue, I next regress the Senate–House difference in spending on the Senate–House difference in committee assignments:

\[
\Delta G_{sta} = \beta_0 + \beta_1 \Delta A_{sta} + U_{sta}. \tag{13}
\]

\(^8\) An alternative approach to addressing this issue would involve explicitly modelling the bicameral negotiations between the House and Senate. Snyder \textit{et al.} (2003) develop such a model and find that a small state bias emerges if the Senate has proposal power or if federal dollars fund statewide public goods. While their model is more involved than the model in this article and would thus be more difficult to analyse empirically, it is encouraging to note that a small state bias emerges in their model under the framework of this article, which assumes statewide public goods and focuses on the role of proposal power in the US Senate.
As shown in column 2, a delegation assigned to a committee of size one would receive roughly 43 percentage points more of the budget than do other states, and this effect is statistically significant. Taken together, results from the first two columns of Table 6 provide significant evidence in favour of the proposal power channel.

I next investigate the vote yield channel, which relates the size of the delegation to the probability of inclusion in the coalition \( \Pr(M_i = 1) / \Pr(V_i) \). To measure this relationship, I estimate the following specification relating inclusion in the coalition to delegation sizes.

As shown in column 2, a delegation assigned to a committee of size one would receive roughly 43 percentage points more of the budget than do other states, and this effect is statistically significant. Taken together, results from the first two columns of Table 6 provide significant evidence in favour of the proposal power channel.

I next investigate the vote yield channel, which relates the size of the delegation to the probability of inclusion in the coalition \( \Pr(M_i = 1) / \Pr(V_i) \). To measure this relationship, I estimate the following specification relating inclusion in the coalition to delegation sizes.
\[ \Delta M_{sta} = \beta_0 + \beta_1 \Delta V_{sta} + U_{sta} \]  
\[(14)\]

where \( \Delta M_{sta} = \Delta M^{\text{SENATE}}_{sta} - \Delta M^{\text{HOUSE}}_{sta} \) represents the Senate–House difference in the indicator for coalition membership, which I measure by the presence of positive project spending.\(^9\) This regression is run over the sample of appropriations bills for which the state is represented on neither the House nor the Senate appropriations subcommittee. As shown in column 3, there is a positive and statistically significant relationship, providing evidence in favour of this vote-cost channel.

### 3.4. Parties and Coalition Sizes

This Section investigates two additional issues raised by our theoretical model: the role of parties and the size of coalitions. Regarding the first issue, recall that our model includes a cost to the agenda-setter from including members in the coalition \((\xi)\). The most obvious interpretation of this cost is a distaste for including members of the opposition party in the coalition. Given that committee chairs, who have significant power within the committee, are members of the majority party, I hypothesise that committees may thus reflect the interests of the majority party. In this case, one would expect members of the minority party to have higher cost terms and thus have a lower probability of inclusion in the coalition with opposite effects for members of the majority party. To investigate this issue empirically, I develop measures of the Senate–House difference in the fraction of the delegation associated with the majority party. Consistent with the above hypothesis and as shown in column 4 of Table 6, there is a positive and statistically significant relationship as delegations with a higher fraction associated with the majority party are more likely to be included in the coalition. Importantly, however, after controlling for the partisan composition of delegations, the coefficient on the size of delegations is similar to that in column 3, suggesting that our evidence in favour of the vote cost channel is not driven by underlying partisan composition variables.

Another prediction of our model involves the size of coalitions. In particular, given the incentive of the agenda-setter to form minimum-sized coalitions, the empirical prediction is that just over 50% of delegate-weighted states should receive positive spending. As a test of this hypothesis, Figure 3 provides the empirical distribution of delegate-weighted states receiving positive spending across appropriations bills, fiscal years and chambers. As shown, there are roughly twice as many appropriations bills with positive spending for 50% to 60% of delegate-weighted states compared to

\(^9\) An alternative approach to measuring inclusion in the coalition would involve roll-call voting records. This would allow us to investigate the prediction that delegations behave cohesively in this model and thus all vote yes or all vote no. The problem with constructing such a voting measure is three-fold. First, some appropriations bills are passed via a voice vote and individual legislator votes are thus not recorded. Second, appropriations bills are, in some cases, passed in an omnibus bill, in which all thirteen appropriations bills are combined into a single bill. Third, appropriations bills contain other provisions in addition to projects earmarked for specific states. Evidence from Knight (2004), which demonstrates a positive correlation between transportation project spending in Congressional districts and support by representatives, as expressed in roll-call votes over an amendment to strip projects in all districts from a larger bill, suggests that project spending-based measures and roll call voting-based measures of inclusion in the coalition may yield similar results.

© The Author(s). Journal compilation © Royal Economic Society 2008
appropriations bills with positive spending for 40% to 50% of delegate-weighted states, suggesting an important discontinuity at 50%. On the other hand, many appropriations bills provide positive spending for over 60% of delegate-weighted states, suggesting inefficiently-sized coalitions. There are at least two explanations for the presence of over-sized coalitions. First, Congressional legislation is subject to Presidential veto and bills with a large number of earmarks are a common target for veto threats. Thus, in order to guarantee the two-thirds vote required to over-ride any vetoes, proposers may optimally choose to expand coalition sizes beyond 50%. Second, there could be logrolling across appropriations bills. For example, committees with jurisdiction over water projects may fund projects in states represented by committee members with jurisdiction over transportation projects in exchange for transportation projects.

4. Conclusion

This article provides a theoretical and empirical investigation of the relationship between representation in legislatures in the geographic distribution of federal funds. I begin by developing a simple legislative bargaining model in which both delegation sizes and state population are variable. In this model, increases in representation lead to increases in federal spending, and this relationship is the result of two underlying channels. First, increased representation leads to an increased likelihood of being represented by the agenda-setter. Second, increased representation increases the attractiveness of the district from the perspective of agenda setters attempting to form optimal coalitions.

While the existing empirical literature has focused on cross-state analyses, I show that such analyses may be hindered by the presence of unobserved state-level variables, and I instead focus on a within-state analysis. In particular, I show that small population
states, which are over-represented in the Senate but not the House, receive significantly more funding in Senate bills than in House bills, and the reverse is true for large population states. As evidence regarding the underlying channels, I find that these small states are also significantly more likely to be assigned to key Senate Appropriations committees than to corresponding House committees, and the reverse is true for large states. Committee members also receive significantly more funding than do non-committee members, and I thus find evidence in favour of the proposal power channel. I also find that small states are significantly more likely to be included in coalitions in Senate bills than in House bills. Again, the reverse is true for large states. This finding in support of the vote-cost channel is robust to the inclusion of measures of the affiliation of the delegation with the majority party. Taken together, the empirical analysis provides strong support for the hypothesised relationship between representation and federal spending along with the underlying channels identified in the theoretical model.

Brown University

Submitted: 15 November 2006
Accepted: 11 November 2007

References


