

Communication in Federal Politics: Universalism, Policy Uniformity, and the Optimal Allocation of Fiscal Authority*

Anke S. Kessler[†]

Revised Version: April 2013

Abstract. The paper presents a positive model of communication in federal legislatures to study the incentives of members to engage in a meaningful exchange of information, and how this shapes policy outcomes. Depending on the type of policy under consideration, communication between delegates generally suffers from a bias that make truthful revelation difficult and sometimes impossible. This generates inefficient policy choices at the federal level that are often endogenously characterized by overspending, universalism and uniformity. Building on these findings, I develop a simple theory of fiscal (de-)centralization, which revisits Oates (1972)' decentralization theorem in a world of incomplete information and strategic communication. Empirical results from a cross section of U.S. municipalities are consistent with the predicted pattern of spending.

JEL Classification: D72, H77, D82, H50

Keywords: Legislative Behavior, Fiscal Federalism, Authority, Communication, Debate

*I am indebted to Philip Reny (the editor) for his detailed and thoughtful comments on an earlier draft, and two anonymous referees for valuable suggestions. I also wish to thank Daron Acemoglu, Philippe Aghion, Elhanan Helpman, Kevin Milligan, Krishna Pendakur, Torsten Persson, Guido Tabellini, as well as various participants of the 2012 IIPF congress, the 2008 Canadian Public Economic Group Meeting, and seminars at the Universities of Basel, Cologne, and Munich for helpful discussions, with special thanks belonging to Tim Besley and Ken Shepsle. I am grateful to the Canadian Institute for Advanced Research (CIFAR) and SSHRC for financial support, and to Ross Hickey for invaluable research assistance. The usual disclaimer applies

[†]Simon Fraser University, CIFAR and CEPR. Address of Correspondence: Simon Fraser University, Department of Economics, 8888 University Drive, Burnaby, B.C. V5A 1S6, Canada. E-mail: akessler@sfu.ca

1 Introduction

The majority of countries are federally organized; multiple tiers of government fulfill a variety of functions in revenue raising, taxation, and public expenditure. Among those, there has been a recent trend towards reallocating fiscal responsibilities from central governments to regional or local governments. Examples of countries that decentralized include the United States¹, the U.K. with its newly created regional legislatures in Scotland and Wales, a number of Latin American countries [Faguet (2004)], as well as Indonesia and Pakistan. Decentralization is also one of the World Bank stated policy objectives for developing countries.² At the same time, the debate among member countries of the European Union has evolved around “subsidiarity” versus centralization, i.e., which functions should remain with the regions, and which can sensibly be assumed by the European Parliament and the Council.

In popular discussions decentralization is often seen as preferable, fostering efficiency through inter-governmental competition and accountability through local say over service provision. The economic literature building on Oates’ famous Decentralization Theorem (Oates, 1972) promotes a more balanced view by emphasizing a trade-off between local policy decisions that are better tailored to the needs of the local population (“closer to the people”) and the obvious advantages of centralization if policies exhibit large economies of scale or spill-over effects across jurisdictions. While Oates’ perspective has shaped the theory of fiscal federalism for decades, it has recently come under criticism on the grounds that an essential aspect of policy formation is missing, namely, an understanding of how political actors on various levels of government are incentivised, and interact.

This paper addresses fiscal federalism from a political economy perspective. My purpose is twofold. First, I develop a model of legislative behavior which allows for communication as a key factor in legislative decision making on the federal (central) level.³ In a second step, I am then able to compare

¹In the U.S., the most prominent explicit decentralization initiative was to return responsibility over major welfare programs to the individual states through the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996, under which funding for state-run welfare programs switched from open-ended match grants to fixed block grants, and simultaneously increased the discretion of states to make decisions regarding their own welfare expenditures. Although the Obama administration has moved toward more centralized provision of some aspects of government like health insurance, federalism is still important when it comes to policies like the stimulus package and aid to states, which has totalled 2.4 trillion dollars over the past 3 years (source: Congressional Budget Office).

²In 2002, around 30 World Bank projects had decentralization components. A total of \$500 million USD was given in the form of various loans to countries like Mexico, Argentina, Brazil, India, Mexico and Pakistan. See <http://www1.worldbank.org/publicsector/decentralization/>.

³Communication in legislatures has received surprisingly little attention in the literature. Austen-Smith (1990) analyzes the informational role of debate in federal legislatures in a series of examples with majority voting over fixed or endogenous agendas. The author shows that communication allows legislators who would otherwise reveal their information through proposals to share this information prior to the agenda-setting stage. Gilligan and Krehbiel (1989),

federal decisions to local decisions in order to gain insights into the optimal degree of (de-)centralization and the determinants of fiscal authority.

The analytical framework I employ conceptually formalizes the process of federal policy formation through consultation with interested parties (local officials) or, alternatively, a policy debate among local delegates in a federal assembly. The aim is to capture the stylized elements of a country with multiple layers of government, and to study which factors determine the incentives of local representatives to truthfully reveal locally dispersed information on their constituents' preferences over public policy decisions. The economy is divided into a given number of (geographical) districts, which are defined by common tastes over a local public project (good) with spillover effects on other districts. Local preferences are local knowledge. Under federal fiscal authority, districts (through a representative) can communicate their preferences to the central government, which implements whichever policy it considers best given the transmitted information.⁴ Federally controlled public policies are funded through general taxation, which implies that costs are underestimated at the local level *ceteris paribus*. As I show, this implies that local representatives have an incentive to overstate the local benefit on average in order to seek federal ratification. Moreover, this tendency unambiguously increases in the number of districts, and is reinforced for projects whose benefits are locally concentrated ("distributive politics"). For policies that have public good character, however, this tendency is mitigated. Intuitively, delegates will have an incentive to understate the local benefit (respectively, overstate the local cost) for projects that are costly for their constituents but – due to their positive and large spillovers – are nevertheless likely to receive federal approval. Irrespective of the type of policy under consideration, the federal legislature recognizes the resulting communication bias, implying that meaningful transmission of information becomes difficult or even impossible. In the former case, the equilibrium is characterized by *universalism*: every interested district is assured a project. In the latter case, federal policies suffer from *uniformity of provision*: although it is commonly known that local preferences (generically) differ, those differences fail to be taken into account when the policy decision is made. The failure is rooted in the strategic considerations between the federal government and local officials (between local representatives in a federal assembly, respectively), which entirely

Epstein (1998) and Krishna and Morgan (2001) study legislatures in which informed committees can communicate information through policy proposals (submitting bills) to an uninformed floor. The objective of these papers is to understand how different rules for adding amendments affect equilibrium informational efficiency. In an abstract model outside the political economy context, Dessein (2002) studies communication versus delegation in a model which is similar to mine, but only has two actors, a principal and an agent. The latter has private information that is relevant to decision, which the principal can either take herself (after communicating with the agent), or delegate.

⁴In other words, the central government cannot commit to (message-contingent) policies that are different from what maximizes its own objective. Communication therefore takes the form of cheap talk [Crawford and Sobel (1982)].

prevents the credible transmission of any information. Thus, my results provide a theoretical foundation for two stylized facts about federal politics that have been frequently invoked as assumptions in the theoretical literature, namely, that the legislative process is characterized by universalism and that federal policies are uniform, i.e., do not take local circumstances into account.

Based on the above findings, I next compare the equilibrium policy under centralization with the policies that would have been chosen under decentralized decision making. The results give some understanding of what determines which level of government should have authority over a policy. The reasoning here, although conceptually different from Oates' original argument, is similar in its conclusions: centralization is better at internalizing externalities but worse at accommodating the needs of the local population. In this sense, the model provides a foundation for Oates' celebrated decentralization theorem that is derived from an explicit model of government behavior: policy uniformity is not assumed from the outset, but is a direct consequence of the impossibility to credibly communicate information about local tastes to higher levels of government in equilibrium. Because the problem is more severe the larger number of districts, however, I also find a congestion effect, which takes the form of dis-economies of scale in communication. As a result, centralization is better only if there are not too many districts, *ceteris paribus*.

The theoretical model has several implications regarding the relationship between the size of the legislature and both the level and the composition of public spending. In the final part of the paper, I show that the data are consistent with these predictions. In particular, I document that the number of legislatures has a strong positive effect on the extent of government spending, while it at the same time has a significant negative effect on the relative share of targetable expenditures in the budget. This second main prediction of the theory, namely that a larger legislative body should be associated with a shift in the composition of the government budget away from pork towards relatively more public goods, *ceteris paribus*, is unique to the model and, to the best of my knowledge, has never been empirically documented before. Using an instrumental variable strategy, these findings are shown to be robust to a possible endogeneity of legislative size.

Related Literature

This paper stands at the intersection of several strands of research in political economy. First, the paper contributes to the literature on legislative behavior and organization by providing a rationale for "universalism" in congress. This empirically relevant feature of legislative decision making generates

what is sometimes referred to as “the Law of $1/n$ ” [Weingast et al. (1981)], which postulates that (the inefficiency of) government spending is increasing in the size of the legislature.⁵ Weingast et al. (1981) and more recently, Primo and Snyder (2008), formally derive this law in models that are very similar to the present framework, but assume universalism from the outset and do not consider the possibility of delegates in the federal assembly communicating privately held information.⁶

Second, the empirical part of the analysis fits well into the burgeoning literature that examines the effects of political institutions on outcomes in politics and economic policy. In particular, Baqir (2002), Bradbury and Crain (2001), Gilligan and Matsusaka (2001), Perotti and Kontopoulos (2002), and more recently Egger and Koethenbueger (2010) empirically study the “law of $1/n$ ” and find support for the predicted positive relationship between government spending and legislature size. My results confirm this finding but go beyond these studies in that I am able to relate the size of the legislature to the *composition* of government spending as well. For this reason, the paper is also related to Persson et al. (2000); Persson and Tabellini (2004), Lizzeri and Persico (2001) and Milesi-Ferretti et al. (2002) who show how electoral rules and the types of democratic political institutions matter for the level and the composition of government spending.⁷ The empirical work in this literature mostly uses cross-country comparisons, however, which are prone to suffer from an omitted variables problem. Those are less likely to arise if one exploits within country variation as I do in the present paper.

Finally, the paper contributes to the literature on fiscal federalism. In particular, one key element in the classic theory of fiscal federalism [Oates (1972)] is that a central government will provide a single, uniform level of public output in all jurisdictions, a presumption that has come under criticism [see Seabright (1996), Lockwood (2002) and Besley and Coate (2003)].⁸ In the present paper, federal policies are *endogenously* uniform precisely in those circumstances where centralization is Pareto preferred to decentralization. Although previous contributions to the literature on fiscal federalism

⁵There have been several earlier papers that explain the norm of universalism. Those either rely on a simple expected value comparison between the rewards to legislators of the coalition of the whole as against the uncertainties associated with minimal winning coalitions. As noted by Niou and Ordeshook (1985), however, such a comparison fails to model directly the decision processes within a legislature and, thus, fails to explain why it is in no member’s self-interest to defect from such a norm.

⁶Although powerful, the logic in Weingast et al. (1981) and Primo and Snyder (2008) is unsatisfactory from a theoretical point of view because the legislature is assumed to adopt the “norm” of universalism, unanimously passing omnibus bills that ensure every legislator the projects he or she desires, even though all members would collectively be better off agreeing to a lower level of overall spending.

⁷See Besley and Case (2003) for an overview of results with emphasis on U.S. institutions.

⁸Empirically, it is not difficult to document cases where a central government provides public goods in a discriminatory manner across regions, although a tendency towards ‘equal treatment’ arguably remains. On the theoretical side, it is unclear why a central government does not differentiate between districts.

have looked at informational asymmetries between a central government and regional governments [see in particular, Klibanoff and Poitevin (2009) and Klibanoff and Morduch (1995)] I am not aware of any similar result. To my knowledge, the only other paper that explicitly addresses policy uniformity is by Harstad (2007) and quite different in focus. The author considers a model where two politically autonomous regions negotiate an agreement over local public goods with spillovers under asymmetric information. Although policy uniformity is not an equilibrium phenomenon as in the present framework, Harstad shows a mutual commitment to policy harmonization (uniform policies) may be advantageous in inter-regional negotiations because it reduces delay in bargaining.

The paper proceeds as follows. Section 2 introduces the basic framework, derives the equilibrium conditions, and shows how the information transmitted in equilibrium depends on various parameters of the model. The theoretical model is taken to the data in Section 3. Section 4 gives a brief discussion of the findings and concludes. All proofs are relegated to the Appendix.

2 The Model

2.1 The basic framework

Consider an economy divided into n constituencies or *districts* indexed by $i \in \{1, \dots, n\}$. Each district is composed of a continuum of homogenous households with exogenous income y and mass unity. For expositional purposes, it is convenient to think of these districts as being geographically distinct and I will often refer to them as regions. However, it is equally possible to interpret them as broadly defined constituencies, which are separated along observable demographic or economic (rather than geographic) lines and share a common objective with regard to the policy under consideration. There are $n + 1$ goods in the economy, one private consumption good x and n public projects $g_i \in \{0, 1\}$, one for each district. If the latter are geographic entities (regions, states, municipalities), the assumption is that policies are targeted to a particular locality and, as such, have the natural interpretation of public projects in infrastructure, recreation, urban renewal or the environment. Otherwise, the projects can be thought of as entitlement programs targeted to the respective constituency (families with children, say). Public spending on projects is financed by non-distortionary local income taxes $t_i \in \mathbb{R}$, which may differ across districts.

Within each district, a household gains a private benefit $\theta_i > 0$ from the locally realized public project. In addition, there are spill-overs from projects in other districts $j \neq i$, denoted by γ_{ij} .

The cost of realizing project g_i is $c_i \in (0, y)$. All costs and benefits are measured in terms of the private consumption commodity. The utility of a household in region i from the consumption vector (g_1, \dots, g_n, x_i) is thus equal to

$$u_i = \theta_i g_i + \sum_{j \neq i} \gamma_{ij} g_j + x_i.$$

and the Pareto optimal allocation of public projects is characterized by

$$g_i^*(\theta_i) = \begin{cases} 1 & \text{if } \theta_i + \sigma_i \geq c_i \\ 0 & \text{otherwise,} \end{cases}$$

where $\sigma_i \equiv \sum_j \gamma_{ji}$ measure the aggregate spillover of the public project associated with district i . The parameter measuring the private benefit θ_i from their own project g_i is private information of the individuals belonging to the respective constituency.⁹ People outside the constituency, including a central authority, only know that the θ_i 's are independently distributed according to some smooth distribution function $F_i(\theta_i)$ over the full support $\Theta_i = [0, \bar{\theta}_i]$. All other variables and parameters of the economy are common knowledge. To eliminate trivial cases, I will assume in the remainder that $\sigma_i \in [c_i - \bar{\theta}_i, c_i]$ for all projects g_i $i = 1, \dots, n$.

Since people within a district are identical, we can represent each district by a single agent $i = \{1, \dots, n\}$ who acts on behalf of all citizens. This agent's payoff from a policy vector $\{(g_i, t_i)\}_{i=1, \dots, n}$ is $u_i = \theta_i g_i + \sum_{j \neq i} \gamma_{ji} g_j + y - t_i$ which also captures the aggregate surplus generated in district i . Under *local authority*, the decision over project g_i lies with the local district and the agent should simply be thought of as an elected policy maker who directly determines local policy. In this case, each district is financially responsible for its own (and only its own) project, i.e., the (local) budget constraint reads $t_i = c_i g_i$. The agent chooses g_i to maximize

$$u_i^L = \theta_i g_i + \sum_{j \neq i} \gamma_{ji} g_j + y - c_i g_i, \tag{1}$$

taking the public good supply in all other districts $j \neq i$ as given. In equilibrium,

$$g_i^L(\theta_i) = \begin{cases} 1 & \text{if } \theta_i \geq c_i \\ 0 & \text{otherwise.} \end{cases} \quad i = 1, \dots, n.$$

where the subscript L stands for local authority. Because each district's government only takes the benefit received by its own citizens into account,¹⁰ the policy decisions are Pareto suboptimal: there

⁹I will assume that θ_i is 'soft' information, i.e., agents cannot certify or 'prove' the value of θ_i to others, even if they would like to.

¹⁰The presumption that politically autonomous jurisdictions do not co-ordinate their policies is standard in the literature and appears natural in many circumstances. If inter-jurisdictional contracts are enforceable, however, it is also conceivable that regions engage in Coasian bargaining. So far, only a few papers have studied this possibility of Pareto improving contractual arrangements under decentralization. See Kessler et al. (2009) for a model with efficient bargaining, and Harstad (2007) for an analysis of bargaining under asymmetric information.

is under-provision of public projects if the externality is positive $\sigma_i > 0$ and over-provision if the externality is negative, $\sigma_i < 0$. Moreover, the size of the distortion, measured by the expected difference between optimal and actual expected surplus,

$$\Delta_i^L = \left| \int_{c_i}^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) F_i(\theta_i) \right| > 0$$

is increasing in the degree of spillovers $|\sigma_i|$.

If the decision over project g_i lies with a central (federal) authority, the district representative may be thought of as a regional delegate to the federal assembly, an appointed local public official, or a lobbyist who advances the constituency's interests in the central government. Total federal expenditures $\sum_i c_i g_i$ are still funded by taxes on local residents t_i , $i = 1, \dots, n$ but since funds are often shared at the federal level, the link between a district's tax bill and the implementation of its project will generally not be perfect, i.e., local tax revenues could be raised independently from local project realization. To fix ideas, I will assume that there is some arbitrary but exogenous sharing rule $t_i = \sum_{j=1}^n s_{ij} c_j g_j$, where s_{ij} with $\sum_{i=1}^n s_{ij} = 1$ denotes district i 's share in the funding of project g_j . The central authority's objective is to choose policies (g_1, \dots, g_n) to maximize total surplus

$$u^C = \sum_i u_i = \sum_i \left[\theta_i g_i + \sigma_i g_i + y - \sum_{j=1}^n s_{ij} c_j g_j \right]. \quad (2)$$

There are two interpretations to this benevolent objective function. The first is that the central government is a policy maker whose constituency consists of the entire economy: while local policy makers only cares about his or her own (regional) district, the centre cares about the welfare of all citizens. There is an alternative and perhaps more compelling interpretation, however. Suppose all districts choose one agent as a delegate to send to the national legislature. Since the federal legislature is an enduring institution in which delegates interact repeatedly through communicating, bargaining, and voting over a fairly long period of time, one can expect regional representatives to overcome inefficiencies caused by majoritarian decision rules (minimum winning coalitions) and negotiate their way to the Pareto frontier. This is what Besley and Coate (2003) have called a 'co-operative' legislature. In this view, the above objective represents a utilitarian bargaining solution that could be mutually acceptable under the veil of ignorance, that is, legislators – prior to receiving private information – agree to ex post implement a policy vector that maximizes the joint surplus.¹¹ As we will see shortly, however, this (ex ante) efficient objective does not translate into ex post efficiency since

¹¹Some would argue that creating a forum for communication to debate policies and negotiate mutually beneficial agreements is one of the main purposes of a national assembly. Evidence backs this view: in the US house, for instance,

credible communication is not always feasible once delegates receive private information on their local project. Before analyzing this case, however, it is instructive to study the case without communication as a benchmark.

Thus, suppose the federal government does not know the realization of the vector of district-specific preference parameters $\theta = (\theta_1, \dots, \theta_n)$. Given prior beliefs $F_i(\theta_i)$, federal policies will satisfy

$$(g_1, \dots, g_n) \in \arg \max E \left(\sum_i \left[\theta_i g_i + \sigma_i g_i - \sum_j s_{ij} c_j g_j \right] \right)$$

or, $\forall \theta_i \in \Theta_i$ and $i = 1, \dots, n$

$$g_i^{EE}(\theta_i) = 1 \quad \Leftrightarrow \quad E[\theta_i] + \sigma_i \geq c_i, \quad (3)$$

where $E[\cdot]$ is the unconditional expectation operator and the superscript EE stands for the *ex ante efficient* decision, which naturally is the optimal policy for an uninformed federal authority that maximizes total surplus.

From (3), the lack of information for the central government yields uniform policies g_i , at least up to observable differences: consider any two districts i and j whose projects are ex ante indistinguishable, i.e., $(c_i, \sigma_i, \Theta_i) = (c_j, \sigma_j, \Theta_j)$. Even though local benefits will generally differ, $\theta_i \neq \theta_j$, a federal authority without further information must treat them equally, i.e., either $g_i^{EE}(\theta_i) = g_j^{EE}(\theta_j) \equiv 1$ or $g_i^{EE}(\theta_i) = g_j^{EE}(\theta_j) \equiv 0$. Moreover, the fact that federal policy is insensitive towards local preferences under incomplete information implies that decentralized policy making may be preferable. Indeed, because local authority is ‘closer to the people’, it is easy to see that it will often be strictly better if spillovers are very small, $|\sigma_i| \rightarrow 0$. For very large spillovers, in contrast, the inefficiency due to the lack of coordination under local authority dominates and central authority is the desirable mode of governance. Obviously, these conclusions mirror Oates (1972)’s classic arguments in favour or against decentralization within a framework of locally dispersed private information, taking the informational disadvantage of the central (federal) government as given [see also Oates (1999) and Oates (2005)].

While alluding to the lack of information on the part of the central government to explain policy uniformity is appealing at first glance, it remains unsatisfactory on second thought. If the federal government does not know regional preferences, and if this is what prevents it from adapting policies that are better suited for the local constituencies, why does it not ask local public officials? More

minimum winning coalitions are the exception rather than the rule. In the EU, the number of representatives is relatively small, which makes it likely that they will exploit the benefits of co-operation. Also, many decisions require unanimity, which may force legislators to co-operate. Also note that if the districts are sufficiently symmetric, implementing a joint surplus-maximizing objective under the veil of ignorance would not require side ‘payments’ (logrolling).

generally, what prevents regions from communicating their preferences to the federal government? Indeed, is communication not what a federal assembly *with regional delegates* is all about? We therefore next turn to the main part of our analysis, which endogenizes the lack of information at the federal level by allowing for communication between local representatives and the federal government (respectively, a debate among members in the legislature).

2.2 Legislative Communication

Now suppose authority over spending and taxation rests with the federal government but information can flow between the central government and the districts (local officials) in the sense that the latter can communicate their local benefit θ_i to the former. In doing so, however, they have to take into account the fact that the center will – upon having received the information communicated by the districts and possibly updated its prior on θ – implement its most preferred policy. There are two interpretations of this communication consistent with the two views of a central authority laid out above. First, if the central authority is a federal government that is distinct from the local representatives, one could imagine the centre consulting regional representatives and officials on the project before making a decision. Alternatively, the central authority may simply be a federal assembly, which itself is composed of regional delegates. In this interpretation, the information transmission stage can be seen as formalizing delegates communicating with each other – a policy debate. The constraint they operate under, however, is that, at the end of the day, the assembly will pass a bill that maximizes joint surplus.¹²

Formally, communication and policy selection under central authority is described by the following three stage game. In the first stage, nature chooses the vector of local benefits θ according to $F_i(\theta_i)$, and each district representative learns the $\theta_i \in \Theta_i$ for its local project g_i . The second stage is a communication stage where the central government consults local officials on, alternatively, the delegates in the federal assembly engage in a political debate. At a very general level, this communication can be formalized by the exchange of messages m_i . Upon observing m_i , the central authority forms new

¹²One might argue that allowing the federal government to design a mechanism to elicit the private information of the districts – as opposed to the communication outlined above – would be a natural next step. A mechanism (complete contract) requires full commitment, however, and by definition budget authority rests with the federal government in a centralized regime. Importantly, this includes the right to ‘renege’ on promises made, especially when the result would be a Pareto improvement. By assumption, the only way the federal government can commit not to implement its most preferred policy is to decentralize, i.e., formally place projects under local authority. By the same token, I also do not consider the possibility of regions ‘co-operating’ under decentralization. Consistent with the central government’s inability to commit, the analysis also implicitly assumes that a local policymaker cannot commit to any policy that does not maximize the respective objective function. See also the discussion following Theorem 1.

beliefs $\mu_i(m_i, \cdot)$ over $\theta_i \in \Theta_i$. Given $\mu = (\mu_1, \dots, \mu_n)$ the central authority then implements a policy vector (g_1, \dots, g_n) that maximizes expected social surplus. Note that because the federal government always chooses its most preferred policy, conditional on beliefs μ , the only thing communication may achieve is to change μ . Any communication is therefore ‘cheap talk’ and could in principle be quite complicated (the exchange of messages could be conditional and repeated). Under our assumptions that a) private information is not correlated across districts and b) as far as a single district representative is concerned, the decision on her own local project is *independent* of what happens in other districts, however, it is easy to see that the cheap talk game for each district can be analyzed separately. Moreover, there is no loss of generality restricting attention to a single message m_i that is transmitted from the local representative to the central authority, who then decides upon g_i given its updated beliefs $\mu_i(m_i)$.

2.2.1 Equilibrium Analysis

Cheap talk games with a single sender have been studied extensively in the seminal contribution of Crawford and Sobel (1982), who consider a generic version of the game where a better-informed sender can send arbitrary messages to a receiver who eventually makes an irreversible decision that affects the well-being of both. The authors show that the Bayesian Nash equilibria of the game will be characterized by a partition of the parameter space into connected sets (intervals). In each of these equilibria, the sender optimally and truthfully announces to which interval the realized state of nature belongs given the receiver’s decision rule, and the receiver maximizes her expected utility, given her updated beliefs based on the correct but coarse information about the state of nature she received from the sender.

Before analyzing equilibria with meaningful communication, however, it is important to note that there there always exists an equilibrium where no information is transmitted. Given the sender’s message is unrelated to his private information,¹³ the receiver rationally does not update her beliefs and picks the optimal action based on her prior. Conversely, since the receiver ‘ignores’ the message sent, any message is consistent with an equilibrium. Of course, there may be other equilibria with finer partitions in which informative communication is feasible. Since there is one decision taken for every message sent (which corresponds to an interval in the parameter space), the number of actions taken in equilibrium is limited by the number of elements of the partition. The converse is also true, however: if there are two messages (partition elements) that trigger identical actions, then we can combine

¹³For instance, the sender could truthfully announce that the the realized state belongs to the entire parameter space.

them into a single message (partition element) without changing the equilibrium outcome. Since there are at most two decisions in the present model, we can thus without loss of generality assume that the message space is $\{0, 1\}$; a strategy for district i is then a mapping $m_i : \Theta_i \rightarrow \{0, 1\}$ while a strategy for the federal government is $g_i^C : \{0, 1\} \rightarrow \{0, 1\}$. Finally, note that because preferences satisfy single crossing, any message sent by agents with private information on the partition containing higher values of θ_i ; otherwise, it would be optimal for local representatives with extreme values of θ_i to change their strategy (lie).

Lemma 1. *Communication is limited under centralization. In particular, for each district i , there are at most two types of Bayesian Nash equilibria (up to labelling differences):*

- (a) *Communication is completely uninformative. In this equilibrium, the local representative sends a message $m_i(\theta_i) = m_i(\theta'_i) \forall \theta_i, \theta'_i \in \Theta_i$ and the centralized policy consequently satisfies*

$$g_i^C(m_i(\theta_i)) = g_i^{EE}(\theta_i) \quad \forall \theta_i \in \Theta_i$$

- (b) *Communication is informative but incomplete. The corresponding equilibrium is characterized by a cut-off value $\tilde{\theta}_i \in (0, \bar{\theta}_i)$ such that*

$$m_i(\theta_i) = \begin{cases} 1 & \text{if } \theta_i \in [\tilde{\theta}_i, \bar{\theta}_i] \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad g_i^C(m_i(\theta_i)) = m_i(\theta_i) \quad \forall \theta_i \in \Theta_i$$

Proof. The proof follows directly from Crawford and Sobel (1982) and the above discussion □

The equilibrium in (b), which I refer to as the *informative communication* equilibrium in the remainder, is illustrated in Figure 1.

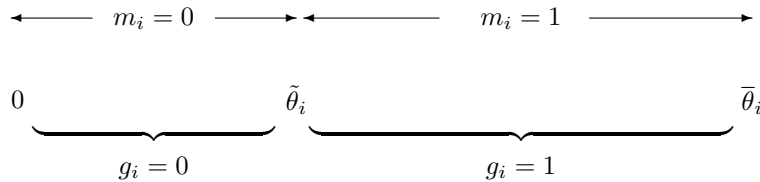


Figure 1: Informative Communication

Since $g_i(m_i(\theta_i)) = m_i(\theta_i)$, communication in this type of equilibrium has a simple structure. The local representative makes a ‘recommendation’ as to whether or not their policy project should be realized, and the federal government follows the recommendation. In equilibrium, all proposals are thus rubber-stamped:

Corollary. *Whenever communication between the federal government and a representative of district $i = 1, \dots, n$ is informative, the equilibrium policy is characterized by universalism, i.e., each representative asking for a project is assured it's approval (regardless of whether the project is socially desirable or not).*

Lemma 1 shows that whenever the informative communication equilibrium exists, it is not unique. So why should we focus on this equilibrium, i.e., is it a natural candidate for equilibrium selection? The answer is given in

Lemma 2. *The equilibrium with informative communication ex ante Pareto dominates the equilibrium where no information is transmitted.*

In what follows, I will assume that agents coordinate on the Pareto superior equilibrium, provided it exists. It remains to analyze when this is the case. To this end, consider an informative communication equilibrium and assume the representative of district i follows his or her equilibrium strategy as prescribed in Lemma 1, i.e., sends a message $m_i(\theta_i) = 1$ for values $\theta_i \geq \tilde{\theta}_i$ and $m_i(\theta_i) = 0$ for values $\theta_i < \tilde{\theta}_i$, respectively. For the federal authority to follow its own equilibrium strategy of rubber-stamping the district's proposal given beliefs μ , we therefore must have

$$E_{\theta_i \geq \tilde{\theta}_i} \{\theta_i + \sigma_i - c_i\} \geq 0 \quad \text{and} \quad E_{\theta_i < \tilde{\theta}_i} \{\theta_i + \sigma_i - c_i\} \leq 0,$$

where the first (second) inequality ensures that the government optimally chooses $g_i^C(m_i) = 1$ (respectively, $g_i^C(m_i) = 0$) after inferring $\theta_i \geq \tilde{\theta}_i$ (respectively, $\theta_i < \tilde{\theta}_i$) upon hearing the message $m_i = 1$ (respectively, $m_i = 0$). Next, consider the representative of the district. Since talk is 'cheap' (sending messages costless), for the communication strategy to be optimal given the government's prescribed equilibrium response of rubber-stamping, it must be the case that the representative prefers $g_i = 1$ to $g_i = 0$ whenever $\theta_i \geq \tilde{\theta}_i$ and $g_i = 0$ to $g_i = 1$ otherwise. Since θ_i is distributed with full support over the interval Θ_i , a representative with preference parameter $\tilde{\theta}_i$ must be indifferent between both outcomes. Denote by $s_i = s_{ii}$ a district's cost share in its own local project. From (1), replacing c_i with $s_i c_i$, whatever the (expected) values of g_j , $j \neq i$, the difference between a district's payoff between $g_i = 1$ and $g_i = 0$ is $\theta_i - s_i c_i$. Indifference at $\theta_i = \tilde{\theta}_i$ thus requires

$$\tilde{\theta}_i = s_i c_i \tag{4}$$

Hence,

Lemma 3. *In equilibrium, there is informative communication between the central government and district i if and only if*

$$E\{\theta_i + \sigma_i - c_i \mid \theta_i \geq s_i c_i\} \geq 0 \quad (5) \quad \text{and} \quad E\{\theta_i + \sigma_i - c_i \mid \theta_i < s_i c_i\} \leq 0. \quad (6)$$

In summary, we can characterize the information that is transmitted through a political debate and the subsequent course of action under centralization as follows. First, communication is imperfect in general, which implies that centralized decisions are never efficient, as would be the case under perfect information. Second, whenever informative communication is feasible, the central authority in effect rubber-stamps local proposals or, put differently, the central legislature operates with universalistic criteria: every district interested in carrying out a project is assured approval. However, the federal government still can – and generally will – discriminate among districts: in equilibrium, given the communicated information, the decision rule $g_i(m_i)$ depends on the realization of the local benefit θ_i . Third, whether or not meaningful communication is in fact feasible depends on the characteristics of the public project (σ_i, c_i) as well as on the fiscal cost sharing rule s_i .¹⁴ The following subsection studies these relations in greater detail.

2.2.2 How much Information flows? Local Bias Effect and Common Pool Effect

Lemma 3 has shown that the incentives to truthfully share information (respectively, the credibility of the shared information) and by extension the centralized equilibrium policy decision, is determined by two exogenous variables: the extent of the spill-overs and the rule governing cost-sharing. To develop some understanding of the mechanisms underlying (credible) information sharing, consider $s_i = 1$, i.e., the district ‘pays the piper’ even though the federal government ‘calls the tune’. Informative communication then requires $\tilde{\theta} = c_i$, i.e., the district proposes to implement its local project under federal authority whenever it would have implemented the project itself under local authority. Were the federal authority always to follow the local recommendation, the outcome under both governance structures would be identical. If spillovers are significant, however, the federal government will sometimes optimally ‘overrule’ the district, effectively undermining any informative communication.

To see this, suppose first spill-overs are negative. Then, the federal government may want to scrap the project even though the local representative is in favour. Because federal approval is gained *less often* than is desired by his constituency, the local representative’s incentives to truthfully commu-

¹⁴It is easy to construct examples to assess how pervasive informed communication (and, as a result, universalism) is in equilibrium. If θ_i is uniformly distributed on $[0, \bar{\theta}_i]$, for example, conditions (5) and (6) reduce to $\sigma_i \leq c_i(1 - \frac{1}{2}s_i) \leq \sigma_i + \frac{1}{2}\bar{\theta}_i$.

nicate θ_i are diminished. In particular, the desire to counteract the federal reluctance to realize the project by overstating local benefits θ_i may become sufficiently strong as to render any meaningful communication infeasible. Formally, this happens for (in absolute terms) large values of σ_i that do not satisfy condition (5). Similarly, if spillovers are positive the federal government will want to realize the project *more often* than is desired by the local population. Again, informative communication will become infeasible because at some point because the incentive of the local representative to counteract federal activism by understating local benefits θ_i is too pronounced: condition (6) is violated for sufficiently large (positive) values of σ_i . The federal government then rationally ignores the local representative's information and realizes the project against the expressed will of the local population.

Because the motive to overstate (respectively, understate) the local value of the project primarily depends on the discrepancy between the private value and the social value of the policy, we may call this effect the *local bias effect*. Also note that due to the simple structure of the model, the local bias effect operates only through a threshold value of σ_i above (respectively, below) which the informative equilibrium ceases to exist. It has no influence on the equilibrium policy as long as information is transmitted.

In practice, of course, a project that is decided upon at the federal level is almost certain to be funded at the federal level as well. In what follows, I will therefore focus on the empirically more relevant case where projects that are under federal jurisdiction are federally funded as well. Assuming a balanced budget, this *de facto* means that a jurisdiction will bear a share of the project cost equal to its share of the federal tax base. For example, irrespective of any (observable or unobservable) differences in the tax bases across districts, cost sharing will necessarily take place whenever the federal budget is at least partly financed by a uniform tax instrument such as a federal income tax or a federal consumption tax. In the simplest case of a central budget financed by a uniform tax on identical tax bases, for instance, we would have $s_i = 1/n$ or $t_i = \frac{1}{n}c_i$. More generally, I will assume¹⁵

¹⁵To see how the local tax burden t_i varies with the decision of the local project g_i , we can differentiate the federal budget to obtain $dt_i/dg_i = c_i/y \in (0, 1)$, so $s_i(n) \in (0, 1)$ follows directly from previous assumptions. Naturally, one can also think of situations where $s_i(n)$ is negative, or exceeds 1. For example, the federal government may promise a greater federal share in future local spending if the district agrees to a project that has high positive spillovers and low local returns. Similarly, there may be (implicit) penalties involved if districts push for projects with large negative spillovers. But this requires project (district) specific subsidies or penalties that dominate a district's share in the federal tax base that naturally results from uniform federal taxes such as income, payroll, or consumption taxes. In either case, allowing for $s_i < 0$ or $s_i > 1$ would not alter the results significantly.

Assumption 1. *The nature of federal taxation is such that costs are de facto shared on the federal level. Moreover, the district tax shares decline in the number of districts. For all districts $i = 1, \dots, n$, we have*

$$0 < s_i(n) < 1 \quad \text{and} \quad s_i(n) > s_i(n+1) \in (0, 1)$$

Public projects thus have “pork barrel” features: they largely benefit a single district at a cost to all others. These type of projects capture the important elements of many real-life policies, and are commonly assumed in the literature on legislative politics [e.g., Weingast et al. (1981), and Grossman and Helpman (2005).] Recall from the previous sections that how funds are raised and costs shared under federal authority had no impact on federal policies in the benchmark cases of perfect information and imperfect information without communication, respectively. Once we allow for communication between the federal government and local officials (among local delegates in a federal assembly) this is no longer true: in conjunction with the lack of information on the part of the federal government, cost sharing now creates a *common pool* problem. Project cost are not entirely borne by the local population, which leads local representative to underestimate the cost of the project ceteris paribus. This in turn creates an incentive to overstate the local benefits vis-a-vis the federal authority, again affecting the transmission of credible information.¹⁶

How the common pool effect and the local bias effect work together is summarized in

Theorem 1. *Consider any informative communication equilibrium. For parameter values $\sigma_i < (1 - s_i)c_i$, the federal policy g_i is characterized by over-provision and this tendency will grow with the number of districts. For values $\sigma_i > (1 - s_i)c_i$, the federal policy g_i is characterized by under-provision. This tendency is mitigated as the number of districts increases to some optimal n^* which minimizes the loss of information. If $\sigma_i = (1 - s_i)c_i$ communication is perfect and the federal policy decision g_i is socially optimal.*

The theorem states how informative the communication between central and a local government on the local policy depends on the extent of the policy’s spillovers and on how costs are shared at the federal level. To understand the result intuitively, suppose first $\sigma_i < (1 - s_i)c_i$. The informative communication equilibrium (assuming it exists) is then characterized by over-provision since the project is realized whenever $\theta_i \geq s_i c_i$ but should only be realized for values $\theta_i \geq c_i - \sigma_i > s_i c_i$. Adding additional districts imply $s'_i < s_i$ and clearly make matters worse: an already over-provided public

¹⁶In a more general formulation with variable project size, this effect holds whenever locally earmarked expenditures grow more rapidly with project scale than local taxes.

project is prone to be even more over-provided as the cost share of the constituency declines and the incentive to overstate its value increases further. Hence, the combination of imperfect communication of privately held information and a common pool problem *endogenously generates dis-economies of (organizational) scale*: the more districts there are, the more difficult it becomes to truthfully communicate a projects' true benefits in the political process, and the more distorted the resulting policy decision will be. The consequences are best seen in the limit case where $s_i \rightarrow 0$. From (4), if informative communication is feasible at all, we must have $g_i^{IC}(\theta_i) = 1$ almost always under federal authority. Intuitively, since the local district's share of the cost is almost nil, it has a strong incentive to overstate local benefits in order to persuade the federal government to realize the project. But we know that the latter always approves whenever it listens to the former, which is necessarily the case if communication is meaningful. For values $\sigma_i > (1 - s_i)c_i$, on the other hand, the local bias effect is sufficiently strong so that the informative communication equilibrium is characterized by under-provision: the project is realized for values $\theta_i \geq s_i c_i > c_i - \sigma_i$. The local-bias effect and the common-pool effect then work in opposite directions, and which effect dominates will depend on the extent of the local bias (σ_i) and on the degree of cost sharing (the number of districts n). In particular, additional districts now have the benefit of counteracting the under-provision problem: more information will flow and the decision will be more efficient as we increase the number of districts from n to $n + 1$ (ignoring the integer problem). Eventually, though, the common pool problem will dominate, assuming $s_i \rightarrow 0$ as n grows sufficiently large. In such a situation, there obviously is an optimal organizational size n^* that balances the common pool effect with the local bias effect and thus minimizes the loss of information under federal authority.

Theorem 1 also points to a case where federally chosen policies are efficient, namely if $\sigma_i = (1 - s_i)c_i$ or $s_i = 1 - \sigma_i/c_i$. Note that this value of the cost share parameter corresponds what is commonly known as the Glarke-Groves mechanism, which induces truth-telling in dominant strategies. As explained earlier¹⁷, the problem with setting the costs shares optimally is that this requires *full commitment* and complete contracting on the part of all parties involved. In contrast, we adopt the incomplete contracting approach by assuming that by its very nature, the central authority cannot commit to any policy that doesn't ex post maximize its objective (recall that the Clarke-Groves Mechanism is not balanced, and hence, not renegotiation proof). It is difficult to envision the federal government writing binding contracts on policies it has budget authority over absent an external enforcement mechanism. Also, a government may not be able to fully bind its successor. An entirely different, but

¹⁷See footnote 12.

perhaps equally compelling, argument is that optimally legislated tax shares will necessarily be both project specific and district specific. As local needs differ and potential projects change frequently (as reflected in annual budgets), the tax code must vary across districts and over time, which could give rise to considerable transaction cost: non-uniform taxation will result in inefficient re-allocation the tax base, and frequent changes create uncertainty hindering investment.

2.3 Assigning Fiscal Authority

I now briefly address how the local-bias effect and the common-pool effect together translate into the overall efficiency of policies chosen on a federal level, and then compare the outcome under federal authority with that under local authority. Note that the comparison is not trivial: as we saw earlier, for instance, communication under central authority improves as the local bias shrinks. Thus we would expect centralization to do well for small (absolute) values of σ_i . At the same time, though, the disadvantage of decentralized authority also disappears for small values of σ_i . Which effect dominates is thus not *a priori* clear. The analysis will shed light on which governance structure (local versus federal authority) the population of an economy would decide upon. In doing so, I will assume that this choice over fiscal authority is made so as to maximize social surplus, and thus coincides with the objective of the federal government.

Suppose first the project under consideration has negative spillovers. Projects in this category include investments with the potential to environmental damage such as water dams or power plants, or infrastructure that intensifies intra-jurisdictional competition (e.g. by attracting mobile capital). In this situation, the local-bias effect is positive, i.e., the local representative will want to overstate the project's benefits even if the district is financially fully responsible, $s_i = 1$. Irrespective of how much costs are shared on the federal level, the common pool and the local bias effect thus work in the same direction, namely, to generate a bias toward over-provision. We find

Theorem 2 (Negative Spillovers). *Suppose $\sigma_i < 0$. Then, local authority over policy g_i is socially preferred if $|\sigma_i| \leq E\{\theta_i - c_i \mid \theta_i \geq c_i\}$ and federal authority over g_i is socially preferred otherwise.*

Moreover, whenever federal authority is socially preferable, the policy decision will be taken under ignorance and is characterized by under-provision. Indeed, the project will not be undertaken.

Assuming that the surplus-maximizing governance structure is chosen and recalling the discussion at the end of Section 2.1 (page 8), it immediately follows that

Corollary. *Policy decisions with negative spillovers over which federal authority is socially preferred will not distinguish between ex ante indistinguishable districts: all such policies will be uniform.*

Next, consider policies with positive spillovers, e.g., projects that improve health care, foster education, or protect the environment. In this situation, the local-bias effect is negative, i.e., the local representative will want to understate the project's benefits if the district is financially fully responsible, $s_i = 1$. The common pool and local bias effects now work in opposite directions.

Theorem 3 (Positive Spillovers). *Suppose $\sigma_i > 0$. Then local authority over g_i is socially preferred if $\sigma_i \leq -E\{\theta_i - c_i \mid s_i c_i \leq \theta_i \leq c_i\}$ and federal authority is preferred otherwise.*

Moreover, the federal policy decision is taken under ignorance for all values $\sigma_i > -E\{\theta_i - c_i \mid 0 \leq \theta_i \leq s_i c_i\}$. In this case, the federal policy decisions are characterized by over-provision.

Again assuming that the socially preferred governance is chosen, we can conclude

Corollary. *Policy decisions with positive spillovers over which federal authority is socially preferred will not distinguish between ex ante indistinguishable districts if either the number of districts is small or spillovers are pronounced.*

To summarize, the model yields the following conclusions. First, communication between the federal government and its local counterparts necessarily remains incomplete. The resulting inefficiency of federal spending decisions is determined by the interplay between the local bias effect and the common pool effect, which in turn depend on the extent of local spillovers σ_i and the number of districts n . In particular, centralized political decision making results in overspending in policy domains with negative or small external effects, whereas too underspending persists in policy domains with large external effects. Second, whenever meaningful communication is feasible, federal spending decisions will be characterized by a form of *universalism*: each representative asking for a project gets the project approved. Third, if meaningful communication is not feasible, federal policy decisions are insensitive to local preferences: federal politics are *uniform*, that is, all districts are treated alike and either all projects are funded or none, possibly against the expressed will of the local population. In either case, the central government's information about the local consequences of its policies is incomplete.

Similar – but not identical – to Oates (1972)'s original argument, I thus identify a trade-off between a loss of coordination under local authority and a loss of information under federal authority, where the

latter become more severe the larger the number of districts. Theorems 2 and 3 relate to this trade-off: from society’s point of view, local policy decisions are preferable whenever the associated externalities are small, and whenever the number of districts (or preference heterogeneity, as measured by the variance of θ_i) is large. Moreover, optimal federal authority necessarily involves uniformity of public good provision for all policies with negative spillovers. Akin to Oates, therefore, federal authority is preferred when externalities are large in absolute terms. At the same time, however, the nature of the externality matters: while (small) positive externalities can result in ‘informed’ decisions at the federal level, and therefore Pareto superior outcomes compared to the ex ante efficient (uninformed) federal decision, the same is not true with negative externalities. Also, if authority were optimally assigned in practise, we should observe “universalism” only if spillovers are positive, and the degree of spillovers or the cost share parameter is not too large.

2.4 Empirical Implications

Some of these results are worth investigating empirically. To derive comparative statics properties of the model that can be linked to available data on expenditures of U.S. municipalities, I will take central authority for granted and introduce some heterogeneity in the model by supposing that the central government has a continuum of potential projects for each district, identified by their project cost $c \in [c^{min}, c^{max}]$, omitting the index i for ease of notation. It is helpful to distinguish between “public good” projects with $\sigma > 0$ and “pork” projects with $\sigma = 0$ (the case where $\sigma < 0$ is analogous to the pure pork case). Also, since the data do not allow for an adequate measure of the efficacy of government spending, the focus here will be on dollar expenditures.

First consider Figure 3 a) which illustrates the situation for public good projects ($\sigma > 0$). The figure is based on the case where an informative equilibrium exists for all values of c and n ; in particular, it assumes that public good projects are ex ante efficient [$E(\theta) + \sigma > c^{max}$], which ensures that for $s \rightarrow 0$ ($n \rightarrow \infty$), meaningful information transmission is still feasible even if costs are high. Because credible information always flows, projects are realized whenever $\theta \geq s(n)c$, which occurs with probability $1 - F(s(n)c)$. Aggregating over c , total expenditures on public goods E_{pg} are given by the area underneath $1 - F(s(n)c)$. An increase in the number of legislators increases (districts) from n to $n + 1$ and a corresponding drop in the local jurisdiction’s cost share from $s(n)$ to $s(n + 1)$ cause $1 - F(\cdot)$ to shift upwards by a factor proportional to c : the budget on public goods expands by an amount ΔE_{pg} , the area shaded in red.

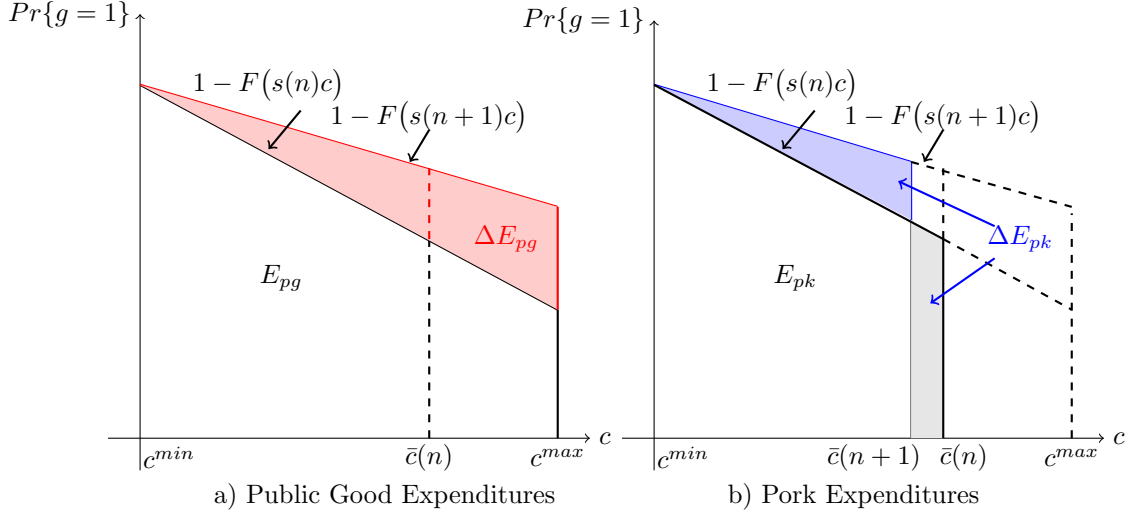


Figure 3

Next, consider pork projects ($\sigma = 0$). Again, if the informative equilibrium exists, spending on pork E_{pk} is equal to the area underneath $1 - F(s(n)c)$, and will increase in n ceteris paribus. Assuming high-cost pork projects are not ex ante efficient [$E(\theta) \leq c^{max}$], however, credible information will not flow for sufficiently high projects costs and large n , which will prompt the federal authority to *drop* projects for which $E[\theta|\theta \geq s(n)c] < c$. In this case, meaningful communication requires that c is below some threshold $\bar{c}(n)$ implicitly defined by $E[\theta|\theta \geq s(n)\bar{c}] = \bar{c}$, with $\bar{c}(n)$ decreasing in n .¹⁸ Figure 3 b) illustrates this case. Assuming local benefits for public goods and pork projects are drawn from the same distribution, and comparing E_{pg} with E_{pk} , we see that the total budget is strictly larger for public goods than for pork (this is also true in the data, as we will see below). Moreover, an increase in n now has two effects: it increases the likelihood of project realization of all projects that are still ‘on the agenda’, i.e., for which $c \leq \bar{c}$ holds. At the same time, it lowers \bar{c} , as the highest cost projects are no longer realized since meaningful communication became impossible. The overall effect ΔE_{pk} is ambiguous; it depends on whether the blue area is larger or smaller than the grey area. For smaller n , or if $E(\theta) > c^{max}$, in contrast, we will have $\bar{c} > c^{max}$, and the second effect (the grey area) disappears, yielding an unambiguous increase in the pork budget, similar to Figure 3 a).

To summarize, under the above assumptions the comparative statics of an increase in the number of legislators n is as follows: expenditures on public good projects will increase, while the effect on pork spending is positive for small n and ambiguous for large n since some high-cost pork projects will be

¹⁸Depending on $F(\theta)$, $c(n)$ may not be unique, in which case the argument is more cumbersome but still valid.

dropped from the agenda if $E(\theta) < c^{max}$. In either case, the increase in public good spending always weekly exceeds the increase in pork spending, and strictly so if $\bar{c}(n) < c^{max}$ (n large).¹⁹ Finally, total government expenditures increase if the rise in public good spending dominates the reduction (if any) in pork spending. This will generally depend on the distribution function and other parameters of the model. For the special case where $F(\cdot)$ is the uniform distribution, and assuming $c^{min} = 0$, $c^{max} = \bar{\theta}$ and $s = 1/n$, for example, it is straightforward to show that both pork spending and total spending unambiguously increase in n (the share of pork spending in the overall budget, however, still declines in n). Of course, how government spending and its composition varies with legislative size is ultimately an empirical question, which I address in the following section.

3 Empirical Analysis

3.1 Data Description

In this section, I will bring some of empirical implications of the theory to data from a cross section of city governments in the United States. These municipalities exhibit substantial variation in terms of institutional structures and policy outcomes, and have several advantages over cross-country data. First, they are more observations than in typical cross-country studies. Second, to estimate the effect of institutions on policies one needs to assume that the units of observation exhibit the same functional relationship between the institutional variable in question (in our case, the size of their legislature) and policy outcomes, which is more plausible in the case of sub-national units that share a common historical and institutional background. Countries are arguably much more heterogeneous. For this reason, omitted variables are also likely to be endemic to cross-country empirical work. Countries may be different for a variety of reasons and since it is impossible to control for these differences in practice, the observed correlations may be due to the fact that countries with different historical or economic backgrounds systematically differ in their institutions and their spending patterns. Exploiting variation within a country, in contrast, decreases the danger of unobserved heterogeneity biasing the results. Finally, of course, local jurisdictions are interesting in their own right because they affect economic outcomes and policies.

The data are drawn from three different sources, combining information about political institutions

¹⁹If some public good projects are ex-ante inefficient, $E(\theta) + \sigma < c^{max}$ the effect of n on the public good spending and relative size of public good to pork spending is no longer monotone, since some public good projects will also be discontinued as n increases. There is some \hat{n} , however, beyond which no further public good projects are discontinued, at which point the above argument applies.

Table 1: Summary Statistics

Variable	Obs.	Min	Max	Mean	Median	St. Dev	Units
Expenditures per capita	1431	0.14	9.74	1.39	1.12	0.96	\$ 1000 per person
Ratio Pork over Public Goods Exp.	1399	0	54.57	.423	.284	1.51	ratio
Pork Exp. per capita	1401	0	.883	.103	.082	.089	\$ 1000 per person
Public Good Exp. per capita	1401	0	2.05	.307	.283	.368	\$ 1000 per person
Population Size	1429	10,018	3,694,834	52,933	25,216	138,466	number of people
Population Density	1429	73.85	56,278.95	2977.94	2997.23	2833.24	people/square mile
Income per capita	1429	7,408	109,219	22,545	19,971	9666.78	\$ per person
Pop. with High School Degree	1429	.215	.988	.809	.824	.109	percentage
Pop. over 65 years of age	1429	.025	.574	.131	.129	.053	percentage
Owner-Occupied Houses	1429	.215	.986	.641	.633	.137	percentage
Ethnic Homogeneity	1429	.235	.971	.654	.660	.185	ratio
Median to Mean Income	1429	.459	.949	.7894	.797	.735	ratio
Council Size	1426	3	40	6.76	7	2.25	number of people
Council Size in 1976	568	3	25	6.66	6	2.40	number of people
Ward Elections	1393	0	1	.401	0	.490	indicator variable
Mayor-Council Form of Gov.	1429	0	1	.336	0	.473	indicator variable

^a **Note:** A city is classified as having ward elections if over half of the councilors are elected in districts, and as 'at large' otherwise

(form of government), population characteristics (municipal demographics) and public finance (city spending). The data on form of government are taken from a survey that is conducted every five years by the International City Management Association (ICMA); the analysis focuses mainly on 2001 but I will also use information from previous year to account for possible changes in government form over time.²⁰ The information on municipal finances comes from the 2002 Census of Governments conducted by the U.S. Bureau of Census, which also provides the data on city demographics from the 2000 Census of Population. The measure on government spending I use is general municipal expenditures per capita. In the data, this is broken down into a multitude of different categories ranging from expenditures on fire protection, police protection, roads and transit services, parks and recreation, utilities, and the like, to budgetary items such as debt service, airport maintenance, liquor store, insurance etc. A full description of all categories is available from the , Census of Governments classification manual provided by U.S. Census Bureau²¹ As I explain below, some of these items can be categorized fairly easily into measures of spending on “pork” versus “public goods”. The first type is comprised of goods and services for which legislators are likely to see the local benefits to their constituents outweigh the local cost (assuming those are shared with the population as whole). At the municipal level, examples are parks and recreational facilities such as swimming pools, as well as public libraries. The second type are those goods and services for which there is either an aggregate (city-wide) benefit that justifies their provision, or costs are locally concentrated, or both. These types of services include police stations, fire halls, as well as expenditures on projects that municipal planners classify as “locally unwanted land uses” (LULU’s) and that are commonly known as “not in my backyard” developments (NIMBY’s). Examples of the latter would be landfills and incinerators, power plants, halfway houses and other correctional facilities, hospitals, and public transit stations.

The empirical work to follow investigates the relationship between the number of city councilors (legislators) and the amount as well as composition of city spending, using a number of other plausible determinants of municipal expenditures as control variables.

²⁰Surveys are sent to roughly 7,000 municipalities with a response rate ranging from 50 to 70 percent.

²¹Available at the U.S. Census Bureau website for Federal, State, & Local Governments statistics, <http://www2.census.gov/govs/pubs/classification/>.

3.2 Results

3.2.1 Government Spending

The first question that emerges from the theory is whether or not the well established “law of 1/n” holds, which states that an increase in the number of legislators should be associated with higher per-capita spending. Recall from the previous section that this relationship is consistent with the model’s prediction, provided that public goods ‘dominate’ the budget. As one can see from Table 1, in the data, both median and mean ratio of pork to public good spending are below 1. Accordingly, I estimate a regression of the form

$$\ln(y_i) = \alpha + \beta \ln(Csize_i) + \gamma X_i + e_i \quad (7)$$

where y_i is per-capita government expenditures in municipality i , $Csize_i$ is the number of councilors in this municipality’s legislative body, and X_i is a vector of controls that mainly comprises demographic characteristics but also includes variables that are important for the institutional environment the municipal governments operate in.

The baseline specification in Table 5 below reports the resulting parameter estimates of ordinary least squares (OLS) regressions, heteroskedasticity-robust standard errors are in parenthesis. The explanatory variables are population size, density, per capita income, educational attainment (the percentage of the residents with a High School degree), population age structure (the percentage of the population over age 65), and the fraction of houses that are occupied by their owners as opposed to renters. All control variables except the fractions are in log form to facilitate the interpretation as elasticities.

Consistent with the theory, an increase in council size is associated with higher government spending per capita. The estimated effect is not only statistically significant but also qualitatively important. The point estimate of the coefficient is $\beta = .2$ percent, which for the median city with 7 councilors would roughly translate into 6 percent higher per capita spending if the size of the council increased from 7 to 9 members. Given average spending of \$ 1,400 in 2000 for the cities in the sample, this amounts to an increase of \$ 84 in annual expenditures per inhabitant.

The result confirms Baqir (2002)’s earlier finding, indicating that the relationship between spending and the size of the legislature is stable over time.²² The effects of the other co-variates are largely as

²²The point estimate is somewhat smaller than Baqir (2002)’s estimate of .302, which can be explained by the presence of two additional controls in the above regression, namely population density and the fraction of owner-occupied housing.

Table 2: OLS Regressions for Government Size

dependent variable	spending per capita			
independent variable	(1)	(2)	(3)	(4)
council size	.223*** (.052)	.215*** (.061)	.158** (.054)	.174** (.085)
population size	.079*** (.016)	.089*** (.018)	.079*** (.011)	.079*** (.015)
density	-.228*** (.019)	-.232*** (.021)	-.192*** (.043)	-.201*** (.031)
income per capita	.531*** (.067)	.588*** (.073)	.574*** (.109)	.513*** (.096)
% High School degree	-.568*** (.195)	-.783*** (.222)	-.839** (.368)	-.735*** (.255)
% of population over 65	1.58*** (.239)	1.51*** (.262)	1.32*** (.245)	1.43*** (.267)
% owner-occ. housing	-1.67*** (.124)	-1.82*** (.148)	-1.66*** (.208)	-1.53*** (.186)
median/mean income		.163 (.175)	-.009 (.262)	-.134 (.279)
ethnic homogeneity		.199** (.099)	.069 (.108)	-.012 (.125)
ward elections		.006 (.003)	-.007 (.039)	-.025 (.033)
strong mayor		-.036 (.033)	-.034 (.030)	-.057 (.038)
regional fixed effects	no	no	yes	no
state fixed effects	no	no	no	yes
number of observations	1427	1387	1387	1394
R^2	.22	.26	.26	.32

^a **Note:** Dependent variable is total municipal spending per capita. The standard errors reported in parentheses are heteroskedasticity-robust, and clustered at the state (respectively, regional) level for the regressions with fixed effects. Superscripts ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

one would expect and are highly significant. Government size increases with population size, average income, and the percentage of people over age 65, and decreases with population density and the fraction of owner-occupied housing. Perhaps surprisingly, a better educated population is associated with lower government spending if one controls for income (once the income variable is dropped from the regression, the coefficient becomes positive and significant).

The latter in particular is highly significant (it accounts for almost half of the explained variation in spending) and is negatively correlated with council size. Excluding both explanatory variables from the regression results in a coefficient of $\beta = .32$.

To check how sensitive the results are, I ran additional regressions allowing for further controls that are likely to affect government spending and may be correlated with council size. Income equality as measured by the ration of median to mean household income, an index of ethnic homogeneity,²³ the city’s electoral system, namely whether candidates are elected from the entire city (“at large”) or from districts (“ward”), and the city form of government which determines whether mayor has substantive authority over the budget (in the “mayor-council” form) or not (in the “council-manager” form). The first two variables capture demographic heterogeneity in the local electorate. This could be important if bigger councils are associated with a more ethnically diverse population structure, which in turn could cause government spending to be elevated. The second two variables consider further institutional factors. Since ward cities tend to have more councilors on average, failing to control for the electoral system may lead to spurious results if the conventional wisdom that cities with a ward-based electoral system spend more than cities where legislators are elected at large holds. Similarly, the mayor — who is generally chosen in city-wide elections — has a veto over the budget in major-council and one would therefore expect spending to be lower under this of government [see Coate and Knight (2009)] which could bias the estimates on the city council effect since there are systematically more council members in cities with strong mayors than in cities with weak mayors.

Column (2) of Table 5 reports the corresponding results.²⁴ We see that the estimated coefficient on the council size remains largely unaffected, indicating in particular that the positive council size effect is neither going through an effect of local heterogeneity in the composition of the electorate, nor through institutional characteristics that shape municipal policy making.

Notably, the reported regressions do not include intergovernmental revenue (transfers from higher level governments). Those are positively correlated with council size in the data, and thus increase its estimated effect on spending. Whether or not one would want to control for intergovernmental transfers is largely a question of what one seeks to measure. In particular, it is not obvious that this correlation necessarily reflects some connection between the size of the transfer and an unobserved municipal characteristic that itself is correlated with council size which would bias the coefficient

²³Formally, $\text{Ethn} = \sum_i s_i^2 \in \{0, 1\}$, where s_i denotes the share of the population of race i in the total population of the city, and $i \in \{\text{white, black, American Indian, Hispanic, Asian and Pacific Islander, other}\}$.

²⁴The finding that both income inequality and ethnic homogeneity are positively associated with government spending can largely be attributed to holding the fraction of owner-occupied houses constant: the latter is positively correlated with either measure of homogeneity in the population, and once it is dropped from the regression, the coefficients become negative (and significantly so for income equality) suggesting that more homogeneous population structures translate into fewer government expenses on average as one would expect. The effect of the ward variable is not significantly different from zero, and the data thus do not confirm the expectation that city councils whose members are predominantly elected in districts spend significantly more than cities where council members are predominantly elected at large.

estimate upward. Rather, an equally likely explanation is that a large council in line with the model desires large expenditures for, say, large public infrastructure projects (bridges, highways, public transit etc), and for that very purpose seeks auxiliary resources from higher level governments.²⁵

The regression in columns (3) and (4) add either regional or state fixed effects, respectively. The regions considered correspond to the U.S. census bureau’s nine “divisions” within in the continental United States.²⁶ Regional fixed effects are worth studying if unobserved regional spending preferences are correlated with preferences for smaller governments. In comparison to the state fixed effects model, this specification has the advantage that the minimum number of observations on the division level is considerably larger than at the state level, which improves the precision of the estimates. Of course, state-specific factors may also be important: state laws will generally affect both the extent of municipal spending (e.g., through the assigned number of municipal functions) as well as the local governance structure. In both regressions, we see that the estimated coefficient on council size drops in value, but remains positive and significant.

3.2.2 Composition of Government Spending

The theory predicts that an increase in the number of legislators (electoral districts, respectively) should be associated with a relatively higher increase in public good expenditures than in pork expenditures, *ceteris paribus*. To the best of my knowledge, this prediction is unique to the present model and has never been investigated empirically before. The regression model is:

$$\ln(r_i) = \alpha + \beta \ln(Csize_i) + \gamma X_i + e_i \quad (8)$$

where r_i is the log ratio of government outlays on pork relative to spending on public goods in municipality i , $Csize_i$ is the number of councilors in this municipality’s legislative body, and X_i is the vector of controls. The specific spending categories I used to account for spending on particularistic goods (pork) were general expenditures on public libraries, parks, and recreational/cultural-scientific facilities (the latter include everything from from golf courses, playgrounds, and swimming pools to stadiums, museums, zoos and celebrations, including public support of cultural activities). The

²⁵As an additional robustness check, I carried out all regression in Table 2 and Table 3 below accounting for intergovernmental transfers to the municipal government. The results are encouraging in that the main coefficient of interest, namely the effect of council size, remains significant, although the size of the effect is diminished. The regressions are obtainable from the author upon request.

²⁶The Census groups states into four broad “regions” and divides each of those into two or more “divisions”. The four census regions are Northeast, Midwest, South, and West. The results reported included fixed effects for the nine divisions, which are New England, Mid-Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific.

functional categories that comprise the spending on public goods are general expenditures on fire and police protection, public transit, as well as direct expenditures on construction in sewage and solid waste management.²⁷

The findings are reported in Table 6 with the same set of controls as in the previous section . The dependent variable in the first set of regressions is the (logged) ratio of spending on pork to spending on public goods. The results from the baseline specification in column (1) strongly support the theory: the relationship between council size and the ratio of spending on pork relative to public goods is negative and highly significant, i.e, cities with larger councils spend relatively more on public goods than on pork, *ceteris paribus*. We also see that the sign of the coefficients on the controls are largely as expected, which again suggests that the model is not grossly misspecified: larger cities spend more on pork relative to public goods while denser cities spend less (though the effect of density is only marginally significant). The fraction of the population with a High School degree is an important predictor for the pattern of spending, as is home-ownership: an electorate that is more educated and owns rather than rents demands, on average, more pork (parks, libraries etc.). Average income does not affect the composition of the budget in any significant way if one controls for education and home ownership; it is easily confirmed, though, that dropping the latter variables from the regression leads to large and significantly positive estimated effect of income.

The second column reports the coefficients from the regression using the full set of co-variates. We see that cities with ethnically more homogeneous populations tend to spend more on pork relative to public goods, while the opposite is true for cities with a strong mayor.²⁸ The latter effect could be explained by the fact that strong mayors are usually chosen in a separate city-wide election which gives them incentive to seek support of from broad coalitions of the population, implementing policies with public good character rather than policies specifically targeted at certain (geographically concentrated) groups. By the same token, though, cities with ward elections where a councilor's constituents live in smaller districts should spend more on pork relative to public goods. A number of recent papers have investigated the relationship between electoral rules and government spending [Persson and Tabellini (2000), Lizzeri and Persico (2001), and Milesi-Ferretti et al. (2002)] theoretically and predict that at-large systems tilt the composition away from pork spending towards spending on public goods

²⁷The categorization was driven primarily by data availability and common sense. I should emphasize, however that the results that follow are not sensitive to including more spending functions into any of the two broad groups of expenditures, or to excluding any one of the categories mentioned above.

²⁸One should note that the former effect is sensitive to the definition of the public goods variable. As an additional sensitivity analysis, I ran all regressions with police expenditures excluded from spending on public goods. The respective estimated coefficient on ethnic homogeneity became markedly smaller and eventually insignificant in the fixed effects regressions (none of the other coefficients were qualitatively affected in this exercise).

Table 3: OLS Regressions for Composition of Government Expenditures

dependent variable ^a independent variable	ratio of pork over public goods				pork	publ. goods
	(1)	(2)	(3)	(4)	(5)	(6)
council size	-.529*** (.909)	-.495*** (.111)	-.263** (.118)	-.225** (.112)	-.103 (.102)	.369*** (.065)
population size	.096*** (.027)	.111*** (.029)	.068** (.029)	.048 (.030)	.213*** (.030)	.107*** (.015)
density	-.078** (.035)	-.063** (.036)	-.029 (.038)	-.031 (.043)	-.142*** (.035)	-.086*** (.023)
income per capita	-.014 (.106)	-.001 (.121)	.204** (.137)	.125* (.151)	.572*** (.127)	.587*** (.772)
% High School degree	2.04*** (.351)	1.94*** (.351)	1.32*** (.383)	1.42*** (.426)	1.04*** (.398)	-.909*** (.212)
% of population over 65	-.377 (.474)	-.622 (.481)	-.432 (.501)	-.632 (.548)	.832** (.486)	1.66 (.262)
% owner-occ. housing	.533** (.211)	.532** (.211)	.296 (.237)	.423 (.302)	-.762*** (.229)	-.134*** (.158)
ethnic homogeneity		.374** (.160)	.409** (.182)	.364* (.197)	.192 (.169)	-.174** (.087)
ward electoral system		.012 (.052)	-.034 (.053)	-.016 (.055)	.008 (.054)	-.014 (.024)
strong mayor		-.123** (.053)	-.045 (.057)	-.023 (.065)	-.147** (.056)	-.013 (.029)
regional fixed effects	no	no	yes	no	no	no
state fixed effects	no	no	no	yes	no	no
number of observations	1380	1348	1340	1348	1349	1393
R^2	.11	.13	.17	.23	.14	.25

^a **Note:** Dependent variable in (1) to (4) is total municipal spending in areas categorized as pork relative to municipal spending in functions categorized as public goods, in (5) it is per capita spending on pork (parks, recreational facilities and libraries), and in (6) it is per capita spending on its per capita spending on public goods (police and fire protection, public transit, and expenditures on the construction of sewage and waste management facilities). All variables except percentages are in log form. Robust standard errors, clustered at the state (regional, respectively) level in the regressions with fixed effects. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

that benefit larger groups of the population. The result that the coefficient on the ward variable was insignificant is inconsistent with this view, and is similar to Baqir (2002) who also finds little support for this hypothesis.

Again, the effect of the number of councilors on the composition of municipal spending continue to hold qualitatively if I control for unobserved taste and institutional parameters by including either regional and state fixed effects, respectively. Results from those regressions are reported in columns (3) and (4). As before, the magnitude of the coefficient is smaller with regional fixed effects and drops further to about half of its former value in the regression with state fixed effects. The qualitative direction of

the effect remains unchanged, though: the size of the city council affects the ratio of pork relative to public good in a positive way, and – together with the effects of education and ethnic homogeneity – is the only explanatory variable that is significant in determining the observed expenditure patterns. To assess the estimate quantitatively, note that a point estimate of $\beta = -.5$ roughly implies a 15 percent drop in the share of pork relative to public goods in government spending if the size of council size increased from a median 7 members to the next higher level of 9 members, clearly a substantial reduction.

Finally, as an additional test columns (5) and (6) report separate regressions for per capita spending in either of the two categories, respectively. The average slope of the relationship between the number of councilors and the level of pork expenditures is negative but not significantly different from zero, which is certainly consistent with the theory (recall Figure 3 b)). In contrast, I find a strong and positive relation between council size and public good expenditures: larger-council cities spend more on public goods than smaller council cities, all else equal. Again, this result is implied by the theory and strongly significant.

3.2.3 IV Regressions

The assumption underlying the identification strategy in the previous sections is that the number of councilors is exogenous. While changing the number of legislators is not a straightforward matter²⁹, one concern with the estimates presented above is that cities whose spending patterns are different have adjusted the size of their legislative body in different ways. One could imagine, for instance, that municipalities were subject to shocks to the preferences of the electorate that drove the demand for institutional reform, those with stronger (weaker) preferences for a large government or more universalistic public spending increasing (decreasing) their council size over time. More generally, the concern is that both the size of the council and spending patterns are correlated with omitted right hand side variables that are not captured by regional or state fixed effects.³⁰

To deal with this issue, I employ an instrumental variable strategy, using both size of the municipal council in 1976 and the electoral system (ward or at large) as instruments for the number of legislators

²⁹The responses in the ICMA questionnaire from 2001 indicates that only about two percent of cities in the sample have attempted to change the size of their council in the five years leading up to the survey. Out of those, roughly two-thirds saw the attempted change approved (changes to the structure and form of municipal government frequently require voter approval, e.g., through a referendum).

³⁰Municipalities are regulated by State statutes. While some states impose upper limits on the number of councillors in a municipality (as a function of its population size, for example), cities often enjoy considerable freedom in choosing election procedures, number of councillors and the type of election (by district versus at large).

in 2001. Not surprisingly, council sizes in 1976 and 2001 are highly correlated (the covariance is .74) and given the length of the elapsed time period, it is not unreasonable to presume that any effect that the number of legislators 24 year ago has on today's spending is going through today's number of legislators [see also Baqir (2002)]. The electoral system is another potentially valid instrument: whether councilors are elected at large or in wards is fairly strongly correlated with councils size (the covariance is .39) and but had no effect effect on the amount or the composition of spending in the first set of IV regressions when only instrumented for council size by council size in 1976.

For brevity of exposition, Table 4 below shows only the estimated coefficient on the main explanatory variable of interest, namely the number of legislators in 2001 using council size in 1976 and electoral system (ward or at large) as instruments. The regressions are estimated two-stage least squares of the baseline model, and the extended model, where the latter is estimated both with and without fixed effects. The Council size in 1976 is only available for a smaller number of cities, resulting in 545 observations. To facilitate comparison, the first row therefore replicates the original OLS regression for the various specifications of the model with the smaller sample size. With the exception of the last specification that includes state fixed effects, all estimated coefficients using the smaller sample are very similar to the estimates from the full sample.³¹ Both instruments have considerable bite: the F-statistic for the first stage is 124, and t-values of 17 for the council size in 1976 and 5 for ward indicator variable make them the two primary predictors of council size in 2001. Since the model is over-identified, I can also whether the over-identification restrictions are satisfied. Using a Hansen-Sargan test, the null hypothesis that the instruments are jointly valid cannot be rejected for any of the specifications below.

The endogenous variable of interest in the first set of regressions is government spending per capita. When I instrument for the 2001 council size in those regressions, the estimated coefficients increase in magnitude for all specifications, which is an indication that the original estimates were too low due to omitted variables. Indeed, a variant of the Durbin-Wu-Hausman test (that is robust to violations of conditional homoskedasticity) shows that exogeneity of council size in 2001 is rejected for the baseline model at the 10 percent level, but cannot be rejected for any other specification.³²

³¹Including fixed effects for 48 states in this smaller sample soaks up a lot of variation in the data, and leaves some states with less than a handful of observations, which likely contributes to the larger standard errors. The estimated β in these regression is also highly susceptible to outliers. For instance, the value of the coefficient on council size increases considerably from .87 to .138 if a single observation is deleted. The observation in question is the city of Winfield, Kansas, which has a very small council of only 3 members in both 1976 and 2001, and spends per capita more than twice as much as the median city in Kansas. In the IV regression, the respective change is from .142 to .201.

³² Baqir (2002) finds qualitatively similar results using 1960 council size to instrument for 1990 council size. One (plausible) explanation he proposes for the observed increase in the coefficient is that, over time, cities with growing

Table 4: Estimated Coefficients on Council Size for Different 2SLS Regressions

dependent variable	model	baseline	extended	regional FE	state FE
per capita exp.	OLS	.243*** (.072)	.227*** (.078)	.296** (.125)	.087 (.103)
	2SLS	.361*** (.085)	.365*** (.095)	.381*** (.115)	.142 (.148)
pork/publ. good exp.	OLS	-.686*** (-.132)	-.697*** (.141)	-.395** (.124)	-.392** (.169)
	2SLS	-.719*** (.171)	-.719*** (.173)	-.404* (.219)	-.393 (.254)

^a **Note:** Instruments are the number of councilors in 1976 and the electoral system (ward or at large) . Number of observations is 546 for the first set and 538 for the second set. Robust standard errors in parenthesis, the errors are clustered at the regional and state level for the fixed effects regressions. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

The second set of regressions uses an IV strategy to study the effect of council size on the composition of spending. Here, the magnitude and pattern of the 2SLS point estimates is very similar to the OLS estimates, suggesting that endogeneity is not a problem for the composition of the budget. A Durban-Wu-Hausman test confirms this: the hypothesis that the OLS estimates are unbiased cannot be rejected. As before, the coefficient drops in magnitude once fixed effects are introduced, but remains negative (though insignificant for the state fixed effects).

To summarize, the results from the empirical section show that the size of the legislative body has a significant effect on the size and the composition of government spending. Cities with larger council spend more per capita on average. At the same time, they tend to spend less on pork barrel projects and more on public goods than cities with a smaller council which are otherwise comparable, leading to a strong and negative relationship between council size and the ratio of pork versus public good expenditures.

4 Discussion and Concluding Remarks

Federal authority results in political failure in the present model despite the fact that the legislature's decision making is efficient, i.e., the legislature acts 'cooperatively' [see Weingast et al. (1981)]. An alternative and non-cooperative approach to the behavior of legislatures is to assume that central

populations could have taken advantage of economies of scale in government spending (thus reducing their per-capita spending) while at the same time having to increase representation of the city council. This would lead to a downward bias in the estimated cross-sectional relationship between spending and council size in 2000, which could partly be corrected for by the instrumental variable approach.

decisions are taken by majority rule. This view is adopted in Lockwood (2002), who studies a model of legislative bargaining in the spirit of Baron and Ferejohn (1989) in which decisions require minimum winning coalition to form. Besley and Coate (2003) consider both cooperative and non-cooperative (majority rule) legislatures.³³ If legislators act non-cooperatively and decisions are taken by majority rule, centralized policies will be tilted towards those regions that are part of the ruling majority, and consistently disadvantage the minority regions. In the present framework, this effect would add another dimension of conflict, which likely renders a meaningful communication with jurisdictions outside the minimum winning coalition impossible. Otherwise, however, the flavor of the result would be preserved. In particular, as long as the cost of local public goods are still financed through a central budget, information transmission for those jurisdictions that belong to the minimum winning coalition should still be governed by the interplay of common pool effect and local bias effect.³⁴

The results in the present paper also shed light on the question whether centralization increase the size of government. In Besley and Coate (2003), regions seek to attract a larger share of central spending by delegating bargaining to representatives with high values for the local public good. In Persson and Tabellini (1994), local policy makers use contributions to persuade a central legislator to allocate public spending towards their regions, in the non-cooperative equilibrium, the central legislator supplies to many public goods. In all these studies, centralization for political decision making results in an over-provision of public goods. In the present paper, this is not true, but rather depends on the local bias effect. In particular, while centralization increases spending in policy domains where project have negative or positive but small spillovers, centralization may actually decrease public spending in policy domains with significant positive spill-overs. This finding is consistent with casual evidence in the European Union. The EU's common agricultural policy where spill-overs are absent or even negative, is arguably characterized by overspending. In other areas with significant spillovers such as the environment, in contrast, spending is remarkably low.

³³By incorporating political economy considerations and focusing on the (distributional) conflict of among regions, these papers thus look at federalism from an angle that is quite different from Oates (1972)'s original insight. In contrast, the model in the present paper pays tribute to the need to explicitly account for the incentive of political actors (although it admittedly does so in a fairly rudimentary way) but is much closer in spirit to Oate's.

³⁴Legislators will continue to be subject to the common pool problem since more districts still imply a lower cost share of one's own jurisdiction on any given project. Similarly, the motive to manipulate the perceived value of one's own project will still depend on the discrepancy between that value and whatever value the decisive legislator (agenda setter) places on the project, giving rise to a local bias effect. While these two main effects are thus likely to survive in a framework with legislative bargaining, it is difficult to assess the full characteristics of any communication equilibrium without an explicit model in mind. Moreover, any information provided may shift the the composition of the minimum winning coalition which will further complicated matters. Also note that direct majority vote is not a suitable model for legislative decision making in this framework of distributive politics. Due to the multidimensionality of the policy vector, a majority vote equilibrium (Condorcet winner) does not exist.

Finally, in order to keep the analysis concise and tractable, I have made specific assumptions concerning the nature of the public projects and the privately held information. In particular, the informational asymmetry only related to a projects idiosyncratic benefits, not to its cost or its spillovers. Moreover, project costs and benefits were additive separable across districts, implying that the efficient local project choice could be made independently of what happened in other districts. Some of these assumptions could easily be altered. For instance, models where districts are better informed about their cost of raising public funds, or about the spillover effect of other projects (rather than about their own project benefit) could be analyzed in an analogous fashion, with similar conclusions. Similarly, one could easily allow the federal government to collect additional information at the local level, provided some residual uncertainty remains.³⁵ It would also be fairly straightforward to study a situation where the realizations of local benefits are correlated, say, because of some common shock to the economy. How much information a local representative can credibly transmit in equilibrium would then hinge on what others have already said. My conjecture is that this effect could make sequential communication desirable; depending on the nature of the correlation, the order of speech would matter and there may well be an optimal sequence of speakers in the assembly that maximizes informational efficiency.

Going one step further, one could consider a scenario where project decisions interrelated, e.g. because the federal budget is fixed. The communicated information of others then directly affects the social desirability of each project. Again, the incentives of local delegates to truthfully convey their information will depend on the communication strategies of other delegates in equilibrium. Because of the interdependency in communication strategies, this case is much more difficult to analyze.³⁶ A formal analysis of these cases is beyond the scope of the present paper but constitutes a fruitful avenue of future research.

³⁵Technically, what is most important is that the variable capturing the privately held information continues to be one-dimensional. Obviously, this would still be the case if districts new more than the federal authority about the project cost or about the potential spillovers (as opposed to project benefits). It is easy to see, then that communication would still be influenced by the common pool and the local bias effect in ways analogous to the current model, yielding qualitatively comparable results. Also, it is straightforward to allow a federal government to gather information on its own, provided local representatives know which information the government possesses (i.e., there no two-sided informational asymmetry). In this case, one can just think of the uncertainty regarding θ as a *residual* uncertainty that is present once the federal government has updated its beliefs from its own information (for instance, the support of Θ could have narrowed as a result of federal information). As long as some residual uncertainty is left, the results will carry through.

³⁶In a series of examples with three delegates who have received private signals about the desirability of a common project, Austen-Smith (1990) shows that there are two opposing effect: on the one hand, delegates have an incentive to share information insofar as their preferences are correlated. On the other hand, information may be withheld due to diversity in ideal points.

References

- Austen-Smith, D. (1990). Information Transmission in Debate. *American Journal of Political Science*, 34(1), pp. 124–152.
- Baqir, R. (2002). Districting and Government Overspending. *Journal of Political Economy*, 110(6), pp. 1318–1354.
- Baron, D. P. and J. Ferejohn (1989). Bargaining in Legislatures. *American Political Science Review*, 83(4), pp. 1181–1206.
- Besley, T. and A. Case (2003). Political Institutions and Policy Choices: Evidence from the United States. *Journal of Economic Literature*, 41(1), pp. 7–73.
- Besley, T. and S. Coate (2003). Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach. *Journal of Public Economics*, 87(12), pp. 2611–2637.
- Bradbury, J. C. and W. M. Crain (2001). Legislative Organization and Government Spending: Cross-Country Evidence. *Journal of Public Economics*, 82(3), pp. 309–325.
- Coate, S. and B. Knight (2009). Government Form and Public Spending: Theory and Evidence from U.S. Municipalities. *NBER Working Papers*.
- Crawford, V. P. and J. Sobel (1982). Strategic Information Transmission. *Econometrica*, 50(6), pp. 1431–1451.
- Dessein, W. (2002). Authority and Communication in Organizations. *Review of Economic Studies*, 69(4), pp. 811–838.
- Egger, P. and M. Koethenbueger (2010). Government spending and legislative organization: Quasi-experimental evidence from Germany. *American Economic Journal: Applied Economics*, 2(4), pp. 200–212.
- Epstein, D. (1998). Partisan and Bipartisan Signaling in Congress. *Journal of Law, Economics, and Organization*, 14(2), pp. 183–204.
- Faguet, J.-P. (2004). Does Decentralization Increase Government Responsiveness to Local Needs? Evidence from Bolivia. *Journal of Public Economics*, 88(3-4), pp. 867–893.

- Gilligan, T. W. and K. Krehbiel (1989). Collective Choice without Procedural Commitment. In P. C. Ordeshook (editor), *Models of Strategic Choice in Politics*, pp. 295–314. Ann Arbor: Michigan University Press.
- Gilligan, T. W. and J. G. Matsusaka (2001). Fiscal Policy, Legislature Size, and Political Parties: Evidence from State and Local governments in the First Half of the 20th Century. *National Tax Journal*, 54(1), pp. 57–82.
- Grossman, G. M. and E. Helpman (2005). Party Discipline and Pork Barrel Politics. *NBER Working Paper # 11396*. National Bureau of Economic Research.
- Harstad, B. (2007). Harmonization and side payments in political cooperation. *American Economic Review*, 97(3), pp. 871–889.
- Kessler, A. S., C. Lülfesmann, and G. Myers (2009). Federations, Constitutions, and Political Bargaining. *CEPR Discussion Paper # 4244*. CEPR, London.
- Klibanoff, P. and J. Morduch (1995). Decentralization, Externalities, and Efficiency. *Review of Economic Studies*, 62(2), pp. 223–247.
- Klibanoff, P. and M. Poitevin (2009). A Theory of (De-)Centralization. Mimeo.
- Krishna, V. and J. Morgan (2001). Asymmetric Information and Legislative Rules: Some Amendments. *American Political Science Review*, 95(2), pp. 435–452.
- Langbein, L., P. Crewson, and C. Brasher (1996). Rethinking ward and at-large elections in cities: Total spending, the number of locations of selected city services, and policy types. *Public Choice*, 88(3), pp. 275–293.
- Lizzeri, A. and N. Persico (2001). The Provision of Public Goods under Alternative Electoral Incentives. *American Economic Review*, 91(1), pp. 225–239.
- Lockwood, B. (2002). Distributive Politics and the Costs of Centralization. *Review of Economic Studies*, 69(2), pp. 313–337.
- Milesi-Ferretti, G. M., R. Perotti, and M. Rostagno (2002). Electoral Systems and Public Spending. *The Quarterly Journal of Economics*, 117(2), pp. 609–657.
- Niou, E. M. and P. C. Ordeshook (1985). Universalism in Congress. *American Journal of Political Science*, 29(2), pp. 246–258.

- Oates, W. E. (1972). *Fiscal Federalism*. New York, NY: Harcourt.
- Oates, W. E. (1999). An Essay on Fiscal Federalism. *Journal of Economic Literature*, 37(3), pp. 1120–1149.
- Oates, W. E. (2005). Toward a Second-Generation Theory of Fiscal Federalism. *International Tax and Public Finance*, 12(4), pp. 349–373.
- Perotti, R. and Y. Kontopoulos (2002). Fragmented fiscal policy. *Journal of Public Economics*, 86(2), pp. 191–222.
- Persson, T., G. Roland, and G. Tabellini (2000). Comparative politics and public finance. *Journal of Political Economy*, 108(6).
- Persson, T. and G. Tabellini (1994). Does Centralization Increase the Size of Government? *European Economic Review*, 38(3-4), pp. 765–773.
- Persson, T. and G. Tabellini (2000). *Political economics: Explaining economic policy*. London and Cambridge: The MIT Press.
- Persson, T. and G. Tabellini (2004). Constitutional Rules and Fiscal Policy Outcomes. *American Economic Review*, 94(1), pp. 25–45.
- Primo, D. and J. Snyder (2008). Distributive Politics and the Law of 1/n. *The Journal of Politics*, 70(02), pp. 477–486.
- Seabright, P. (1996). Accountability and Decentralisation in Government: An Incomplete Contracts Model. *European Economic Review*, 40(1), pp. 61–89.
- Weingast, B. R., K. A. Shepsle, and C. Johnson (1981). The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics. *Journal of Political Economy*, 89(4), pp. 642–664.

Appendix A

Proof of Lemma 2. Note that the equilibrium without information transmission, either $g_i \equiv 1$ or $g_i \equiv 0$, irrespective of θ_i . Now consider an informative equilibrium and a local delegate with information $\theta_i \in [0, \tilde{\theta}_i)$ who sends a message $m_i = 0$ that triggers $g_i = 0$ in equilibrium. Since this agent can alternatively send a message $m_i = 1$ which would prompt the central authority to set $g_i = 1$, he or she

must weakly prefer the former, and strictly so whenever $\theta_i < \tilde{\theta}_i$. The converse is true for delegates with information $\theta_i \in [\tilde{\theta}_i, \bar{\theta}_i]$. The local districts are therefore better off. By a similar revealed preference argument, the central authority is also strictly better off. Since information is being transmitted, it changes the decision $g_i(\theta_i)$ for some states of the world, which must be strictly better than sticking with the (uninformed) decision that prevails without the information. \square

Proof of Theorem 2. For $\sigma_i < 0$, the project is over-provided under local authority. Theorem 1 tells us it is also over-provided under centralization if the informative communication equilibrium exists. But because the project is undertaken even more often in the latter, local authority is strictly preferred. Formally,

$$\Delta_i^C = \int_{s_i c_i}^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) > \int_{c_i}^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) = \Delta_i^L.$$

Next, note that local authority is also socially preferred whenever $|\sigma|$ is sufficiently small such that the federal authority would always realize the project without additional information, i.e., if the informative communication equilibrium does not exist because (6) is violated. Hence, if centralization is ever optimal, it must be the case that a) the federal authority acts under ignorance (no information flows) and b) (5) is violated, which implies $g_i \equiv 0$. The expected surplus lost under centralization is then

$$\Delta_i^C = \int_{c_i - \sigma_i}^{\bar{\theta}_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) > 0. \quad (9)$$

The expected loss in surplus under local authority is

$$\Delta_i^L = - \int_{c_i}^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) > 0. \quad (10)$$

Comparing (9) with (10), we see that local authority is preferred if

$$\begin{aligned} & \int_{c_i - \sigma_i}^{\bar{\theta}_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) \geq - \int_{c_i}^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) \\ \Leftrightarrow & \int_{c_i}^{\bar{\theta}_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) \geq 0 \\ \Leftrightarrow & \int_{c_i}^{\bar{\theta}_i} (\theta_i - c_i) dF_i(\theta_i) \geq -\sigma_i [1 - F(c_i)] \\ \Leftrightarrow & -E\{\theta_i - c_i \mid \theta_i \geq c_i\} \leq \sigma_i. \end{aligned}$$

Finally, note that for $\sigma_i < -E\{\theta_i - c_i \mid \theta_i \geq c_i\}$, condition (5) is indeed violated, i.e., the informative communication equilibrium under centralization does not exist. For those values of σ_i we must have $g_i^C(\theta_i) \equiv 0$ for all $\theta_i \in \Theta_i$, which proves the second part of the theorem. \square

Proof of Theorem 3. Consider first parameter values $0 < \sigma_i \leq (1 - s_i)c_i$. For those values, (6) holds trivially. Hence, if the informative communication equilibrium does not exist in this range, (5) must be violated and we must have $g_i(\theta_i) \equiv 0$ i.e., the project would never be undertaken under federal authority. But then local authority is always preferred because it at least ensures that $g_i = 1$ in some states of the world. If the informative communication equilibrium exists in this range, we know from Theorem 1 that the project is undertaken too often. The expected surplus lost is

$$\Delta_i^C = - \int_{s_i c_i}^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) > 0. \quad (11)$$

Conversely, the project is undertaken not often enough under local authority. The corresponding expected loss in surplus is

$$\Delta_i^L = \int_{c_i - \sigma_i}^{c_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) > 0. \quad (12)$$

Comparing (11) with (12), local authority is preferred if

$$\begin{aligned} & - \int_{s_i c_i}^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) \geq \int_{c_i - \sigma_i}^{c_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) \\ \Leftrightarrow & - \int_{s_i c_i}^{c_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) \leq 0 \\ \Leftrightarrow & - \int_{s_i c_i}^{c_i} (\theta_i - c_i) dF_i(\theta_i) \geq \sigma_i [F(c_i) - F(s_i c_i)] \\ \Leftrightarrow & -E\{\theta_i - c_i \mid s_i c_i \leq \theta_i \leq c_i\} \geq \sigma_i. \end{aligned}$$

Finally, note that for $\sigma_i > -E\{\theta_i - c_i \mid s_i c_i \leq \theta_i \leq c_i\}$, both conditions (5) and (6) hold, i.e., the informative communication equilibrium under centralization indeed exists (and the outcome is preferred to that under local authority).

Next, consider parameter values $(1 - s_i)c_i < \sigma_i$. If the informative communication equilibrium exists in this range, Theorem 1 tells us it will be characterized by under-provision. But because the project

is undertaken even less often under local authority, federal authority is strictly optimal. Formally,

$$\Delta_i^C = \int_{c_i - \sigma_i}^{s_i c_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) < \int_{c_i - \sigma_i}^{c_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) = \Delta_i^L.$$

If the informative communication equilibrium does not exist for $\sigma_i > (1 - s_i)c_i$, we must have (6) violated,

$$\int_0^{s_i c_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) > 0 \quad \Leftrightarrow \quad \sigma_i F(s_i c_i) > - \int_0^{s_i c_i} (\theta_i - c_i) dF_i(\theta_i),$$

or $\sigma_i > -E\{\theta_i - c_i \mid 0 \leq \theta_i \leq s_i c_i\}$. Federal authority is still better since

$$\Delta_i^C = - \int_0^{c_i - \sigma_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) < \int_{c_i - \sigma_i}^{c_i} (\theta_i + \sigma_i - c_i) dF_i(\theta_i) = \Delta_i^L.$$

for those values of σ_i . Federal policy decisions will then be taken under ignorance and will be characterized by excessive provision. Finally, note that $\sigma_i > -E\{\theta_i - c_i \mid 0 \leq \theta_i \leq s_i c_i\}$ will hold if either σ is large or the cost share parameter s_i is large (the number of districts n small). \square

Appendix B

This appendix presents the results of the empirical analysis in the main text when in inter-governmental transfers are controlled for. Table 5 replicates Table 2 with the logarithm of inter-governmental transfers (transfers from federal or state government to the municipal government) as an additional co-variate:

Table 5: Council Effect on Spending Controlling for Intergovern. Transfers

dependent variable	spending per capita			
independent variable	(1)	(3)	(3)	(4)
council size	.124** (.051)	.135** (.061)	.137** (.062)	.153** (.065)
population size	.059*** (.015)	.059*** (.017)	.037** (.017)	.028 (.017)
density	-.246*** (.019)	-.242*** (.021)	-.177*** (.023)	-.173*** (.025)
income per capita	.508*** (.067)	.522*** (.073)	.585*** (.074)	.567*** (.079)
% High School degree	-.594*** (.195)	-.672*** (.220)	-.730*** (.228)	-.663*** (.229)
% of population over 65	1.04*** (.245)	.954*** (.266)	.883*** (.283)	.896*** (.279)
% owner-occ. housing	-1.51*** (.120)	-1.54*** (.145)	-1.55*** (.152)	-1.34*** (.160)
median/mean income		.003 (.169)	.132 (.175)	-.063 (.181)
ethnic homogeneity		.089 (.095)	.115 (.102)	-.016 (.108)
ward elections		.021 (.031)	-.008 (.031)	-.042 (.032)
strong mayor		-.056 (.031)	-.022 (.031)	-.048 (.033)
inter-gov transfers per capita	.140*** (.015)	.142*** (.015)	.171*** (.031)	.212*** (.028)
regional fixed effects	no	no	yes	no
state fixed effects	no	no	no	yes
number of observations	1418	1385	1387	1385
R^2	.28	.29	.32	.35

^a **Note:** Dependent variable is total municipal spending per capita. The standard errors reported in parentheses are heteroskedasticity-robust. Superscripts ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Table 6 replicates the composition of government regressions in Table 3. The dependent variable

Table 6: OLS Regressions for Composition of Government Expenditures

dependent variable ^a independent variable	ratio of pork over public goods				pork	publ. goods
	(1)	(2)	(3)	(4)	(5)	(6)
council size	-.644*** (.935)	-.581*** (.116)	-.196* (.117)	-.212* (.123)	-.103 (.102)	.339*** (.067)
population size	.079*** (.029)	.094*** (.032)	.034 (.032)	.010 (.032)	.213*** (.030)	.107*** (.015)
density	-.122*** (.037)	-.104*** (.039)	-.062 (.041)	-.059 (.046)	-.142*** (.035)	-.086*** (.023)
income per capita	-.052 (.116)	-.077 (.136)	.245** (.148)	.238* (.160)	.572*** (.127)	.587*** (.772)
% High School degree	2.61*** (.389)	2.51*** (.436)	1.65*** (.448)	1.42*** (.426)	1.04*** (.398)	-.909*** (.212)
% of population over 65	-.797 (.505)	-1.17** (.543)	-.756 (.567)	-.632 (.548)	.832** (.486)	1.66 (.262)
% owner-occ. housing	.528** (.231)	.554* (.284)	.121 (.305)	.423 (.302)	-.762*** (.229)	-.134*** (.158)
ethnic homogeneity		.388** (.188)	.566** (.209)	.364* (.197)	.192 (.169)	-.174** (.087)
ward electoral system		.005 (.057)	-.045 (.056)	-.016 (.055)	.008 (.054)	-.014 (.024)
strong mayor		-.155** (.061)	-.028 (.086)	-.023 (.065)	-.147** (.056)	-.013 (.029)
regional fixed effects	no	no	yes	no	no	no
state fixed effects	no	no	no	yes	no	no
number of observations	1373	1340	1333	1348	1349	1393
R^2	.14	.15	.23	.23	.14	.25

^a **Note:** Dependent variable in (1) to (4) is total municipal spending in areas categorized as pork relative to municipal spending in functions categorized as public goods, in (5) it is per capita spending on pork (parks, recreational facilities and libraries), and in (6) it is per capita spending on its per capita spending on public goods (police and fire protection, public transit, and expenditures on the construction of sewage and waste management facilities). All variables except percentages are in log form. Robust standard errors, clustered at the state (regional, respectively) level in the regressions with fixed effects. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.