

# Efficient Centralization: Political Accountability and Regional Disparities\*

Federico Boffa

Amedeo Piolatto

IEB and Università di Macerata

IEB and Universitat de Barcelona

Giacomo A. M. Ponzetto

CREI, Universitat Pompeu Fabra, and Barcelona GSE

First draft, 15 October 2011 – This draft, 15 March 2012

## Abstract

Government accountability varies across regions in a single country. Centralization may increase it. This paper develops a model of political agency that explains both facts through regional differences in voter information. In our model, rent-seeking politicians are constrained by career concerns. As a result, voter information reduces equilibrium rent extraction, but its returns are decreasing. Then centralization is efficient: it reduces rents in less informed constituencies more than it increases them in better informed ones. Meanwhile, centralization transfers political power from the less to the more informed. We show that centralization can be Pareto efficient, if the right balance is struck between distributive policies that privilege the informed and uniform public goods that benefit the uninformed.

---

\*We are grateful to seminar participants at CREI and Erasmus Univeristy Rotterdam for helpful comments. We are indebted to Mark Strazicich for sharing data with us. Piolatto acknowledges financial support from IVIE, the Spanish Ministry of Science and Innovation (grant ECO2009-12680), the Barcelona GSE Research Network and the Generalitat de Catalunya (grant 2009 SGR 102). Ponzetto acknowledges financial support from the Spanish Ministry of Science and Innovation (grants ECO2008-01666, CONSOLIDER-Ingenio 2010 CSD2006-00016, and Juan de la Cierva JCI-2010-08414), the Barcelona GSE Research Network and the Generalitat de Catalunya (2009 SGR 1157). E-mail: federico.boffa@unimc.it, piolatto@ub.edu, gponzetto@crei.cat.

# 1 Introduction

The U.S. federal government has grown steadily since its founding, both in absolute terms and relative to state and local governments. The New Deal represented the most dramatic push towards centralization (Wallis 2000a), but the trend well predates the twentieth century and has continued throughout its second half and into the present, with few exceptions such as Reagan’s decentralization efforts (Inman and Rubinfeld 1997). The past fifty years have witnessed, for instance, a significant expansion of federal environmental policy, with the creation of the Environmental Protection Agency and ensuing national standards. The federal government has also greatly increased its involvement in health care, particularly by the creation and expansion of Medicare, and most recently through the 2010 federal health care reform. Europe has witnessed a parallel sharp increase in the scope and extent of the powers of the European Union since the 1970s (Alesina, Angeloni, and Schuknecht 2005).

This pattern is difficult to understand through the lens of traditional models of fiscal federalism. Oates’s (1972, 1999) classic theory posits two rationales for centralization: economies of scale and policy coordination. The subsequent literature in political economy has mostly emphasized the benefits of decentralization (Lockwood 2006). In particular, decentralized government is held to be more responsive to regional preferences, since legislative bargaining at the national level reduces preference-matching (Lockwood 2002; Besley and Coate 2003). Moreover, decentralization may enhance accountability by enabling yardstick competition between local jurisdictions (Besley and Case 1995; Besley and Smart 2007).<sup>1</sup> This paper, on the other hand, explains the pull towards centralized government, uncovering reasons for its efficiency beyond economies of scale and policy coordination. We show that centralization increases efficiency when voters’ information, and therefore political accountability, is heterogenous across regions.

The quality of public institutions and public officials varies greatly within as well as across countries. Political corruption is unevenly distributed across the United States (Glaeser and Saks 2006). In Oregon, less than one public official per million inhabitants was convicted of corruption-related crimes in an average year from 1976 to 2002. In the same period, the corruption rate was above 5 in Alaska, Mississippi, and Louisiana.<sup>2</sup> Similarly, the European Union encompasses large disparities in the quality of government both across member states, and across regions in a single country (Charron, Dijkstra, and Lapuente 2011).<sup>3</sup>

---

<sup>1</sup>On the other hand, the central government could be less susceptible to capture by special interest groups. However, formal analysis of this possibility has reached ambiguous conclusions (Bardhan and Mookherjee 2000, 2006; Blanchard and Shleifer 2001).

<sup>2</sup>The average is 2.8 per million and the standard deviation 1.3.

<sup>3</sup>Italy is the most striking and best known European example (Putnam 1993; Del Monte and Papagni 2001, 2007; Golden and Picci 2005). Regional variation in political corruption is also widely documented in

In this paper we present a model of political agency in which accountability is determined by voter information, in keeping with the finding that government performance improves with media scrutiny (Besley and Burgess 2002; Ferraz and Finan 2008; Snyder and Strömberg 2010; Ponzetto 2011), as well as with voters' human capital (Glaeser and Saks 2006) and social capital (Nannicini et al. 2010), which facilitate respectively the acquisition and the sharing of information. We show that there are decreasing returns to monitoring politicians. As a consequence, when voter information varies across regions, fiscal centralization improves aggregate policy outcomes. The central politician is held accountable by voters in all regions, so his incentives depend on the average level of information. This represents a large improvement for regions whose residents are less informed than average, and a smaller loss for the better informed regions, yielding aggregate gains for the citizens.

Historical evidence bears out our theoretical result. Centralization increased rulers' accountability and the provision of public goods in precolonial Africa (Gennaioli and Rainer 2007). In the United States, federal oversight of relief spending, introduced by the New Deal, eliminated the corruption that had been hitherto endemic to transfer programs managed at the state and local level (Wallis 2000b; Wallis, Fishback, and Kantor 2006). The power of the states relative to the federal government and political corruption both declined from their respective nineteenth-century peaks (Glaeser and Goldin 2006).<sup>4</sup>

U.S. environmental policy since 1970 provides a more recent example of centralization. As an example of its effectiveness, White (1982) points to federal standards on car emissions. The introduction in 1970 of national standards for pollutant emissions marks a structural break in the time series of state-level emissions per capita. Sulfur dioxide emissions have since been converging across a large majority of states (72%). Instead, they were diverging across most states (52%) under the pre-existing local standards (Bulte, List, and Strazicich 2007). Consistent with our theory, federal intervention improved average environmental quality while reducing variation across states. Moreover, we find that states with below-average voter information benefited from centralization, and those with above-average information did not. Our measure of information is daily newspaper circulation per capita in the state in 1970.<sup>5</sup> We split the sample at the average level of circulation, and compare average annual sulfur dioxide emissions per capita recorded in the decades before (1961–70) and after the reform (1971–80).<sup>6</sup> Emissions declined by 13.16% on average in less informed states, while

---

developing countries, including Brazil (Ferraz and Finan 2008, 2011), China (Dong and Torgler 2010), India (Charron 2010), Madagascar (Francken, Minten, and Swinnen 2009), Russia (Diniño and Orttung 2005), and Uganda (Reinikka and Svensson 2004).

<sup>4</sup>Contemporary evidence on decentralization and corruption yields conflicting findings, but countries with more layers of government seem to have more corruption (Treisman 2007; Fan, Lin, and Treisman 2009).

<sup>5</sup>Newspaper circulation data are from the *Statistical Abstract of the United States*.

<sup>6</sup>EPA pollutant data from Bulte, List, and Strazicich (2007).

they increased by 2.62% in better informed ones. For the whole 1961–80 period, we estimate the regression

$$Y_{s,t} = \beta \text{Uninformed}_s \times \text{Federal}_t + d_s + d_t + \varepsilon_{s,t},$$

where  $s$  indexes states and  $t$  years, with respective dummies  $d_s$  and  $d_t$ .  $Y_{s,t}$  denotes per-capita sulfur dioxide emissions.  $\text{Uninformed}_s$  is a dummy for states with newspaper circulation below the national average in 1970.  $\text{Federal}_t$  is a dummy for periods after the reform, i.e., from 1971 onwards. The differential impact of centralization on less versus more informed states is measured by the coefficient  $\beta$ . The estimated value is  $\hat{\beta} = -.0314159$  (.0092304), which is significantly negative at the 1% confidence level. As predicted by our model, centralization provides differential benefits that favor less informed states.<sup>7</sup>

Similar evidence exists for Europe. Fredriksson and Gaston (2000) consider an EU directive introducing uniform standards for packaging waste, and conclude that it “was less stringent than the existing German, Danish and Dutch laws, but was significantly stricter than the Greek, Irish and Portuguese requirements.” Durante, Labartino, and Perotti (2011) study the consequences of decentralization in the Italian public university system. A 1998 reform transferred responsibility for faculty hiring from the national level to the individual universities. As a result, the quality of academic recruitment fell in provinces with lower newspaper readership. Those with higher readership experienced no decline but at most a marginal improvement, implying an aggregate efficiency loss from decentralization.<sup>8</sup>

Section 2 sets up our model of public good provision in an imperfectly representative democracy. The ruling politician allocates an exogenous government budget among multiple public goods. He can also extract rents that provide him with private benefits at the taxpayers’ expense. Although politicians’ only goal is rent extraction, they are disciplined by career concerns. Their skill at providing public goods is stochastic, and voters rationally infer that the incumbent is more capable when they observe greater public-good provision. Hence, the incumbent moderates his current rent extraction to increase his probability of being re-elected and being able to continue extracting rents. Within this framework of political agency, we introduce variations in political accountability by considering imperfectly informed voters. Public-good provision increases support for the incumbent only among voters who can observe it. Thus a rational rent-seeking politician extracts smaller rents when his constituents are better informed of his performance. In addition to improved incentives,

---

<sup>7</sup>The sign and significance of  $\beta$  are unaltered if we use as our dependent variable  $Y_{s,t} - Y_{s,t-1}$ .

<sup>8</sup>Quality is inversely measured by the prevalence of the same surname among faculty in the same department. Homonymy signals the practice of hiring and promoting close relatives, which in turn is linked to a decline in research performance. The introduction of the reform is a structural break that marks a discontinuous and permanent divergence in the extent of familism in provinces with newspaper readership above and below the median.

greater voter information also leads to better selection of politicians, since higher-ability incumbents are more likely to be recognized and reelected.

Section 3 derives the novel efficiency benefits of centralization implied by our model. Centralized provision of public goods reduces aggregate rent extraction through three complementary channels. First, when voter information is heterogeneous across regions. Second, when there are information spillovers across regions. Then voters may learn about public goods in other regions even if they fail to observe them in their own. This additional information has no social value under decentralization. Instead, under centralization it enables voters to infer the competence of the national politician. As a consequence, both his incentives and his selection are better than those of local politicians. This result is exactly symmetric to yardstick competition, reflecting opposite assumptions about voters' information. In our model, voters know the ideal provision of public goods, but they are imperfectly informed about their actual provision. Yardstick competition emerges instead when voters are uncertain about optimal policy, but observe perfectly enacted policy. In that case, decentralization leverages information spillovers. In ours, centralization does. Our result is consistent with Gentzkow's (2006) empirical findings. As television crowded out newspapers, media coverage shifted from local to national news. Voters' knowledge of local politics declined, and so did their turnout in local relative to presidential elections. Changes in media focus over the twentieth century thus suggests a further reason for the observed increase in centralization.

Interregional externalities in public-good provision are the third and last determinant of the efficiency of centralization. This is the only force that is already considered in the traditional theory of fiscal federalism. The classic analysis is summarized by Oates's (1972) Decentralization Theorem, which focuses on the trade off between the benefits of policy coordination and the costs of policy uniformity.<sup>9</sup> The theorem continues to hold in our framework only if information is homogeneous and without spillovers. Otherwise, centralization is superior whenever voters' preferences are homogeneous, regardless of externalities. Conversely, decentralization is no longer unambiguously optimal even when preferences vary across regions and the central government must provide public goods uniformly.

Section 4 considers policy uniformity from a distributive perspective. If voters in different regions are differently informed, centralization reduces rent extraction through an efficient transfer of accountability from the better to the worse informed regions. Then public-good provision also falls for the former and rises for the latter if the central government operates under a uniformity constraint. The direction of the transfer is reversed if different regions can

---

<sup>9</sup>The same focus has been applied to the study of international unions (Alesina and Barro 2002; Alesina, Angeloni, and Etro 2005).

receive different local public goods under centralization. Then the informed regions to whom the central politician is more accountable receive the lion’s share of government spending, as shown by Strömberg (2004). These opposite forces provide a reason for imposing a uniformity constraint on some public goods, but not all. The resulting balance can make centralization Pareto efficient, allowing all regions to share the gains of decreased rent extraction.

Real-world polities seem capable of striking such a balance. For instance, in the case of U.S. environmental policy, introducing uniform national standards benefited less informed states. But the federal government conversely benefited urban northern states—the most informed ones—at the expense of the rural South through the policy of “prevention of significant deteriorations” (Pashigian 1985).<sup>10</sup> In such an equilibrium our model implies that the benefits from centralization are a convex function of voter information. We show that American National Election Studies survey data are consistent with this prediction. Trust in the federal government is a convex function of media exposure to election campaigns.

## 2 Political Agency with Imperfectly Informed Voters

### 2.1 Setup

The economy is populated by infinitely lived agents, whose preferences are separable over time and quasilinear across a set of public goods  $p = 1, \dots, P$ . Individual  $i$  in period  $t$  derives utility

$$u_t^i = \tilde{u}_t^i + \sum_{p=1}^P \alpha_p^i \log g_{p,t}, \quad (1)$$

where  $\tilde{u}_t^i$  is utility from private consumption, and  $g_{p,t}$  the provision of public good  $p$ . The relative importance of each good for individual  $i$  is described by the shares  $\alpha_p^i \geq 0$  such that  $\sum_{p=1}^P \alpha_p^i = 1$ . We focus on public-good provision, treating  $\tilde{u}_t^i$  as an exogenous shock.

Public goods are provided by a politician who allocates a given government budget  $b$ , invariant over time and subject to a balanced-budget constraint every period. Expenditure on each public good  $p$  is  $x_{p,t}$ , and residually the politician can extract an unproductive private rent  $r_t \geq 0$ . Thus

$$r_t + \sum_{p=1}^P x_{p,t} = b. \quad (2)$$

The politician’s objective is to maximize rent extraction over his term in office. His discount factor is  $\delta \in (0, 1]$ .

---

<sup>10</sup>The policy prohibits areas with air quality superior to the minimum national standards from permitting a significant deterioration of air quality. The prohibition places limits on economic development.

For a given level of spending, public goods are produced with technology

$$g_{p,t} = e^{\eta_{p,t}} x_{p,t}. \quad (3)$$

Productivity  $\eta_{p,t}$  represents the stochastic competence of the politician in providing good  $p$ . It is independent across public goods, and follows a first-order moving average process

$$\eta_{p,t} = \varepsilon_{p,t} + \varepsilon_{p,t-1}. \quad (4)$$

The shocks  $\varepsilon_{p,t}$  are independent and identically distributed across policies, over time, and across politicians. They have support  $[\check{\varepsilon}, \hat{\varepsilon}]$ , mean zero and variance  $\sigma^2$ .

Within each period  $t$ , events unfold according to the following timeline.

1. The incumbent politician's past competence shock  $\varepsilon_{t-1}$  becomes common knowledge.
2. The incumbent chooses expenditures  $\mathbf{x}_t$ , and residually rent  $r_t$ , without knowing the realization of his period- $t$  competence shock  $\varepsilon_t$ .
3.  $\varepsilon_t$  is realized and the provision of public goods  $\mathbf{g}_t$  is determined.
4. Each voter  $i$  observes the provision  $g_{p,t}$  of public good  $p$  with probability  $\theta_i$ ; with probability  $1 - \theta_i$  he remains completely uninformed about  $g_{p,t}$ .<sup>11</sup> The arrival of information is independent across voters. No voter can have any direct observation of  $\varepsilon_t$ ,  $r_t$ , or  $x_{p,t}$  for any  $p$ .
5. An election is held, pitting the incumbent against a single challenger, randomly drawn from the same pool of potential office-holders.

The electorate consists of a continuum of atomistic voters. It can be partitioned into  $J$  internally homogeneous groups. Group  $j$  comprises a fraction  $\lambda_j$  of voters, who have identical preferences  $\alpha_p^j$ , and identical probabilities  $\theta_j$  of information acquisition.

Each voter's political preferences consist of two independent elements, following the probabilistic-voting approach (Lindbeck and Weibull 1987). First, agents have preferences  $\mathbb{E}u_{t+1}^i$  over the provision of public goods they expect from either politician in the following period. Given information  $\Omega_t^i$ , individual  $i$  has policy preferences

$$\Delta_i(\Omega_t^i) \equiv \mathbb{E} \left[ \sum_{p=1}^P \alpha_p^i (\log g_{p,t+1}^I - \log g_{p,t+1}^C) \mid \Omega_t^i \right], \quad (5)$$

---

<sup>11</sup>This is not inconsistent with knowledge of one's own utility  $u_t^i$ . The exogenous component  $\tilde{u}_t^i$  may include a stochastic shock, and uninformed voters are unable to distinguish between the effects of the shock and those of  $g_{p,t}$ .

where  $g_{p,t+1}^I$  denotes public-good provision if the incumbent is re-elected, and  $g_{p,t+1}^C$  if the challenger defeats him. In addition, voters have preferences for candidates' non-policy characteristics, such as their likability or the ideology of their party. Thus voter  $i$  votes for the incumbent if and only if

$$\Delta_i(\Omega_t^i) \geq \Psi_t + \psi_t^i, \quad (6)$$

where  $\Psi_t$  and  $\psi_t^i$  are independent draws from common-knowledge probability distributions. The common shock  $\Psi_t$  to the incumbent's popularity accounts for the aggregate uncertainty in the electoral outcome. The idiosyncratic shock  $\psi_t^i$  to each voter's tastes provides the intensive margin of political support, and is i.i.d. across agents. Both variables have uniform distributions with supports respectively  $[-1/(2\phi), 1/(2\phi)]$  and  $[-\bar{\psi}, \bar{\psi}]$ , sufficiently wide that neither any voter's ballot nor the outcomes of the election are perfectly predictable on the basis of policy considerations alone.

**Assumption 1** *The support of the electoral shocks  $\Psi_t$  and  $\psi_t^i$  is sufficiently wide, and that of the competence shocks  $\varepsilon_{p,t}$  sufficiently narrow, that*

$$\frac{1}{2\phi} - \bar{\psi} \leq \check{\varepsilon} < \hat{\varepsilon} \leq \bar{\psi} - \frac{1}{2\phi} \quad \text{and} \quad -\frac{1}{2\phi} \leq \check{\varepsilon}\bar{\theta} < \hat{\varepsilon}\bar{\theta} \leq \frac{1}{2\phi},$$

where

$$\bar{\theta} = \sum_{j=1}^J \lambda_j \theta_j.$$

## 2.2 Equilibrium

The incumbent's and the challenger's competence shocks are known to be independent draws from a common distribution. Moreover, voters have rational expectations that any politician in every period will choose the same allocation  $\bar{\mathbf{x}}$ , because the environment is stationary and performance is separable in effort and ability. Voter's  $i$  information is described by the set  $\Omega_t^i$  of public goods  $p$  whose provision  $g_{p,t}$  he has observed. These observations allow him to infer the incumbent's competence  $\eta_{p,t}$ , whereas he has no information about the challenger. His policy preferences are

$$\Delta_i(\Omega_t^i) = \sum_{p \in \Omega_t^i} \alpha_p^i \mathbb{E}(\varepsilon_{p,t} | g_{p,t}) = \sum_{p \in \Omega_t^i} \alpha_p^i (\log g_{p,t} - \log \bar{x}_p - \varepsilon_{p,t-1}).$$

Each group  $j$  comprises a continuum of agents and the arrival of information is independent across agents, so a share  $\theta_j$  of its member have observed public-good provision  $g_{p,t}$ , while the remainder  $1 - \theta_j$  have not. Given the independent realizations of the uniform



idiosyncratic shock  $\psi^i$ , the share of members of group  $j$  who vote for the incumbent is

$$v_t^j = \frac{1}{2} + \frac{1}{2\psi} \left[ \theta_j \sum_{p=1}^P \alpha_p^j (\log g_{p,t} - \log \bar{x}_p - \varepsilon_{p,t-1}) - \Psi_t \right], \quad (7)$$

conditional on the realizations of  $\mathbf{g}_t$  and  $\Psi_t$ . Taking into account the uniform aggregate shock  $\Psi_t$ , the incumbent's probability of re-election is

$$\pi(\mathbf{x}_t) = \frac{1}{2} + \phi \sum_{j=1}^J \lambda_j \theta_j \sum_{p=1}^P \alpha_p^j (\log x_{p,t} - \log \bar{x}_p) \quad (8)$$

as a function of his policy choices  $\mathbf{x}_t$  (and residually  $r_t$ ).

The politician understands that if he is re-elected he will have further occasions to extract rents. Denote by  $R$  their expected present value. The trade-off between current and future rent extraction leads to policy choices

$$\mathbf{x} = \arg \max_{\mathbf{x}_t} \left\{ \tau y - \sum_{p=1}^P x_{p,t} + R \pi(\mathbf{x}_t) \right\}, \quad (9)$$

namely

$$x_p = \phi R \sum_{j=1}^J \lambda_j \theta_j \alpha_p^j \text{ for all } p = 1, \dots, P, \quad (10)$$

and thus current rent extraction

$$r = \tau y - \phi \bar{\theta} R. \quad (11)$$

In equilibrium, the probability of re-election is  $1/2$ , because the politician does not have private information at the time of the policy choice, voters have rational expectations, and their non-policy tastes do not have a permanent bias against incumbency nor in its favor. As a consequence, the present value of re-election is

$$R = \delta \sum_{t=0}^{\infty} \left( \frac{\delta}{2} \right)^t r = \frac{2\delta}{2-\delta} r. \quad (12)$$

Substituting this relationship into equations 10 and 11, solving the latter, and plugging it into the former, we obtain the equilibrium allocation of the government budget. To economize on notation, we define

$$\tilde{\delta} \equiv \frac{\delta}{2-\delta} \in (0, 1], \quad (13)$$

a convenient rescaling of the politician's discount factor.

Finally, rational expectations imply that the incumbent is re-elected if and only if

$$\Psi_t \leq \sum_{j=1}^J \lambda_j \theta_j \sum_{p=1}^P \alpha_p^j \varepsilon_{p,t}. \quad (14)$$

Let  $\chi_t$  be an indicator variable for this condition. The competence of ruling politicians evolves according to

$$\hat{\eta}_t = \chi_{t-1} (\varepsilon_{t-1}^I + \varepsilon_t^I) + (1 - \chi_{t-1}) (\varepsilon_{t-1}^C + \varepsilon_t^C), \quad (15)$$

where the superscripts  $I$  and  $C$  refer to the incumbent and challenger in the election at the end of period  $t - 1$ . The unconditional expectation of ability is then

$$\mathbb{E} \hat{\eta}_{p,t} = \mathbb{E} (\chi_{t-1} \varepsilon_{p,t-1}). \quad (16)$$

This completes the characterization of the stationary equilibrium of our model of political career concerns.

**Proposition 1** *In equilibrium, a politician with budget  $b$  extracts rent  $r = \rho b$ , with rent extraction*

$$\rho \equiv \left(1 + 2\tilde{\delta}\phi\bar{\theta}\right)^{-1}$$

*He sets expenditure on the provision of public good  $p$  to  $x_p = \beta_p (1 - \rho) b$ , with relative shares*

$$\beta_p \equiv \sum_{j=1}^J \lambda_j \frac{\theta_j}{\bar{\theta}} \alpha_p^j.$$

*His expected ability at providing public good  $p$  is*

$$\mathbb{E} \hat{\eta}_p = \phi \sigma^2 \sum_{j=1}^J \lambda_j \theta_j \alpha_p^j.$$

Equilibrium rents are proportional to the government budget  $b$ . Rent extraction declines as voters are more informed ( $\partial \rho / \partial \bar{\theta} < 0$ ), since information allows them to monitor the politician more tightly. Rents are also reduced when voters are more likely to be swayed by policy than non-policy considerations ( $\partial \rho / \partial \phi < 0$ ). More conventionally, rent extraction declines when politicians are more patient ( $\partial \rho / \partial \delta > 0$ ), because they are more willing to sacrifice current benefits for a higher probability of remaining in office in the future.

The same factors that make elections a better incentive device also make them a better screening mechanism. The average productivity of government spending is proportional to

the variance of the underlying distribution of ability ( $\sigma^2$ ), which measures the gains available from screening. The effectiveness of screening rises when voters are more informed about public-good provision ( $\partial \mathbb{E} \hat{\eta}_p / \partial \theta_j > 0$ ), and thus have the ability to cast their ballots on the basis of a signal of the incumbent's skill. Moreover, screening is more effective when citizens are more willing to vote on the basis of observed performance, rather than out of idiosyncratic non-policy tastes ( $\partial \mathbb{E} \hat{\eta}_p / \partial \phi > 0$ ). All these effects indeed raise not only the expectation of ability, but its entire distribution, in the sense of first-order stochastic dominance.

The allocation of spending across public goods ( $\beta$ ) reflects both differences in preferences and differences in information. A utilitarian social welfare planner would set budget shares

$$\beta_p = \bar{\alpha}_p \equiv \sum_{j=1}^J \lambda_j \alpha_p^j, \quad (17)$$

reflecting average preferences for each public good. However, the politician is only imperfectly accountable, and specifically he is only held accountable by informed voters. Thus his choices deviate from social welfare optimization even beyond the extraction of a rent. The allocation of expenditure across public goods obey a weighted social welfare function in which each group's preferences are weighted by its level of information  $\theta^j$ .

The equilibrium welfare of each member of group  $j$  is

$$\mathbb{E} u_j = \sum_{p=1}^P \alpha_p^j \mathbb{E} \log g_{p,t} = \log b + \log(1 - \rho) + \sum_{p=1}^P \alpha_p^j (\mathbb{E} \hat{\eta}_p + \log \beta_p) \quad (18)$$

### 3 Fiscal Federalism with Heterogeneous Regions

Each locality  $l = 1, \dots, L$  is inhabited by a measure-one population of homogeneous residents. There are then  $LP$  public goods: their indexing is expanded so that  $g_{l,p,t}$  is the provision of public good  $p$  in locality  $l$  at time  $t$ . Politicians are drawn independently across regions, from identical local pools.

Externalities in public-good provision are measured by  $\xi_p^l \in [0, 1]$ . A resident of  $l$  derives utility

$$\alpha_{l,p}^l = \left(1 - \frac{L-1}{L} \xi_p^l\right) \alpha_p^l > 0 \quad (19)$$

from public goods provided to his own region, but he may also derive additional utility

$$\alpha_{m,p}^l = \frac{1}{L} \xi_p^l \alpha_p^l \text{ for } l \neq m \quad (20)$$

from public goods provided in any other region. Thus we can write individual utility

$$u_t^i = \tilde{u}_t^i + \sum_{p=1}^P \left[ (1 - \xi_p^i) \alpha_p^i \log g_{l,p,t} + \frac{1}{L} \xi_p^i \sum_{m=1}^L \alpha_p^i \log g_{m,p,t} \right]. \quad (21)$$

Regions can also interact through the inter-regional spread of information, measured by  $\iota_l \in [0, 1]$ . A resident of  $l$  is informed of the provision  $g_{l,p,t}$  of a public good in his own region with probability  $\theta_l \in [0, 1]$ . He may also observe its provision in any other region with probability  $\iota_l \theta_l$ . The arrival of information is independent across localities.

Decentralization means that in each region a local politician with ability  $\eta_{l,p,t}$  independently allocates the regional budget  $b_l^D$  to expenditures on local public goods  $x_{l,p,t}^D$ , and residually extracts a rent in proportion  $\rho_l^D$ .

Centralization means that a single politician with ability  $\eta_{p,t}$  allocates the central budget  $b^C$  to expenditures on public goods  $x_{p,t}^C$ , and residually extracts a rent in proportion  $\rho_l^C$ . Each region then receives public goods

$$g_{l,p,t}^C = \frac{1}{L} e^{\eta_{p,t}} x_{p,t}^C. \quad (22)$$

This formulation implies that there are no economies of scale in public-good provision, which removes the most obvious rationale for centralization. It also means that centralization imposes a constraint of policy uniformity:  $g_{l,p,t}^C = g_{p,t}^C$  for all regions  $l$ . This constraint has been highlighted, since Oates (1972), as the simplest rationale for decentralization (e.g., Alesina and Spolaore 1997; Alesina, Angeloni, and Etro 2005). Its presence allows a direct comparison between our results and Oates's classic Decentralization Theorem.

The exogenous government budget is identical under centralization and decentralization, and homogeneous across regions. Every region  $l$  has a local government budget  $b^D$  under decentralization, or under centralization it contributes  $b^D$  to the central government budget  $b^C = Lb^D$ . Differences in tax revenues resulting from heterogeneous income across regions would provide an immediate welfare benefit of centralization, since a utilitarian welfare function favors the egalitarian distribution of resources across regions that is implied by the uniformity constraint.

Our main focus is on the effect of centralization or decentralization on political accountability, which is summarized by the following proposition.

**Proposition 2** *Aggregate rent extraction is lower under centralization ( $\rho^C \leq \sum_{l=1}^L \rho_l^D / L$ ). It is strictly lower if information is heterogeneous across regions ( $\theta_l \neq \theta_m$  for some regions  $l \neq m$ ), if there are information flows across regions ( $\iota_l > 0$  for some  $l$ ), or if there are*

externalities in public-good provision ( $\xi_p^l > 0$  for some  $l$  and  $p$ ).

Average efficiency in providing each public good is higher under centralization ( $\mathbb{E}\hat{\eta}_p^C \geq \sum_{l=1}^L \mathbb{E}\hat{\eta}_{l,p}^D/L$ ). It is strictly higher if there are information spillovers across regions ( $\iota_l > 0$  for some  $l$ ) or externalities in public-good provision ( $\xi_p^l > 0$  for some  $l$  and  $p$ ).

The first and key result in the proposition is that if voters are heterogeneously informed ( $\theta_l \neq \theta_m$  for  $l \neq m$ ), and thus politicians are heterogeneously accountable, centralization has beneficial aggregate effects on accountability. The decline in rent extraction is an intuitive consequence of decreasing returns to monitoring. By joining heterogeneous regions into a single polity, centralization leads to an overall level of political information equal to the average  $\bar{\theta}$  of information across regions. For regions with low accountability under decentralization, this represents a large improvement, because the increase in accountability is powerful given the low starting point. For regions with high accountability under decentralization, the deterioration is not equally stark, because the marginal value of information is low when it is plentiful to begin with. The aggregate effect of centralization is thus an unambiguous decrease in rent extraction.

Centralization can increase political accountability through a second channel. The probability that a voter observes public good provision  $g_{p,t}$  increases if information spreads independently across regions. Intuitively, political accountability is imperfect in our model because voters need not observe public goods that are being provided to them. One way of receiving this information is via direct usage of the public services: for instance, a household would directly observe school quality if it includes school-age children, but not otherwise. This kind of direct experience is possible only for goods provided to one's own region. However, a voter can also learn of public goods indirectly, through reports from the media or personal acquaintances. This indirect channel of information acquisition can surely operate across regions. Then  $\iota_l > 0$  if information about different regions is obtained independently (or at least with less than perfect correlation).

For instance, local public good provision may be covered in a local newspaper. Whether or not a voter reads it, he may get information about public goods in another region from national television. In this scenario, under decentralization only newspaper readers have the information to cast an informed ballot in the local election. However, television watchers can correctly infer the ability of the central politician and cast an informed ballot in the national election. Although the information they received is not about public goods provided in their own region, the inference is equally valid. Thus centralization becomes more efficient as the media coverage shifts from local to national news. This has been the case in the United States with the rise of television and the decline of newspapers. Gentzkow (2006) documents that, as a consequence of the changing media environment, turnout declined for local elections

relative presidential elections.

The positive impact of information spillovers on the efficiency of centralization represents the mirror image of yardstick competition (Besley and Case 1995; Besley and Smart 2007). If voters are uncertain about the optimal provision of public goods, but perfectly informed about its actual provision, then decentralization provides positive information externalities. Voters have a chance to learn the optimum from the independent choices of neighboring regions: they can thus realize if their local politician is rent-seeking rather than benevolent, and deny him re-election accordingly. The opposite is true if instead voters are certain about the optimal provision of public goods, but imperfectly informed about its actual provision. Then centralization provides positive information externalities. Voters have a chance to learn from independent observations of neighboring region how capable the central politician is, and to support his reelection accordingly.

The resulting increase in voter information leads not only to improved incentives and lower rent extraction, but also to improved screening. If  $\iota_l > 0$  the voters are able to cast more informed ballots and thus select politicians whose average ability ( $\mathbb{E}\hat{\eta}_p$ ) is greater. Not only wasteful rents decline, but the efficiency of productive public spending rises.

A similar effect both on incentives and on selection plays out if there are spillovers in public goods across regions ( $\xi_p^l > 0$ ). Advantages of centralization in the presence of inter-regional externalities are present in all theories of federalism since Oates (1972). Yet the classical benefits result from a different mechanism and do not affect political accountability. Proposition 2 does not concern the benefits of coordination. We find that spillovers improve incentives even if there is no need to coordinate budget allocations because preferences are identical across regions.

Political accountability improves due to the intensive margin of electoral support. In a setting of probabilistic voting, each citizen is more likely to support the incumbent if he has proved to be more capable than average. The intensity of popular support, however, depends not only on the extent of ability, but also on its importance. A voter who is informed of the incumbent's skills may nonetheless vote for him because of his personal likability or ideological affinity. He is, however, less likely to do so as the stakes in the election grow higher. These are higher in a national election than in a local election if there are public-good spillovers. For  $\xi > 0$ , every voter is keener on electing a proficient politician at the central rather than at the regional level. Thus centralization reduces the influence of non-policy preferences on electoral outcomes, improving the monitoring value of elections.

On the other hand, in our model, as in Oates (1972), decentralization can be beneficial to avoid the cost of policy uniformity when regions have heterogeneous preferences. However, Proposition 2 establishes several forces that tend to make centralization more efficient than

decentralization. Thus, the Decentralization Theorem holds in our setting if voters are homogeneously informed and there are no information spillovers ( $\theta_l = \theta$  and  $\iota_l = 0$  for all  $l$ ). With heterogeneous information or information spillovers, centralization is systematically welfare increasing.

**Corollary 1** *Suppose that information is homogeneous and does not spill over regional boundaries ( $\theta_l = \theta$  and  $\iota_l = 0$  for all  $l$ ). Then:*

1. *If there are no externalities and preferences are homogeneous ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = 0$  for all  $l$  and  $p$ ), then centralization and decentralization yield identical outcomes.*
2. *If there are externalities and preferences are homogeneous ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = \xi_p$  for all  $l$ , with  $\xi_p^l > 0$  for some  $p$ ), then centralization yields higher welfare than decentralization.*
3. *If there are no externalities and preferences are heterogeneous ( $\xi_p^l = 0$  for all  $l$  and  $p$  while  $\alpha_p^l \neq \alpha_p^m$  for some  $l \neq m$  and  $p$ ), then decentralization yields higher welfare than centralization.*

*Suppose that information is heterogeneous ( $\theta_l \neq \theta_m$  for some  $l \neq m$ ) or there are information spillovers across regions ( $\iota > 0$ ). Then if preferences are homogeneous ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = \xi_p$  for all  $l$ ), centralization yields higher welfare than decentralization.*

The first three points coincide with Oates's Decentralization Theorem. When there are neither heterogeneity nor spillovers in information, these classic results obtain in spite of the distortions arising from imperfect political agency.

Without externalities, there are no benefits from policy coordination. With homogeneous preferences, there are no costs of policy uniformity. Furthermore, in this case imperfect agency causes the same distortions under centralization or decentralization. Constituency size affects political agency through two opposing forces (Seabright 1996; Persson and Tabellini 2000). Centralization reduces the probability that voters in any one region are pivotal in the election. Hence a central politician is less responsive to each voter's preferences than a local politician is to those of his fewer constituents. Conversely, centralization increases the scale of political rents. When the politician allocates the larger central budget instead of a smaller regional budget, re-election is more valuable. A greater value of re-election sharpens the incentives for the central politician to perform well. Corollary 1 shows that these forces are perfectly balanced. Centralization expands the budget by a factor  $L$ , while reducing the electoral clout of each region by a factor  $1/L$ . The politician's

incentives are thus invariant with respect to the scale of his constituency. Rent extraction is proportional to the government budget, as established in Proposition 1.

The second point of the Decentralization Theorem deals with the benefits of policy coordination. Oates (1972) assumed that local politicians maximize local welfare but cannot cooperate. In our setting, even if local politicians could cooperate across regions, they would have no incentives to do so. Local politicians are uninterested in changing each other's behavior. Their only goal is to signal his ability to their own constituents, which they do most effectively by ignoring all externalities. Thus, Corollary 1 reflects an endogenous inability to internalize externalities under decentralization. In addition, Proposition 2 showed that the incentives and screening both improve under centralization when externalities exist.

Finally, the third point highlights the standard cost of a binding uniformity constraint. Centralization is costly when regions have heterogeneous preferences, because it involves a suboptimal allocation of expenditure across regions. Furthermore, with imperfect political agency a binding uniformity constraint also worsens electoral screening. Politicians' skill sets are more congruent with their constituents' preferences when they are elected locally rather than in a single national election.

Corollary 1 concludes by showing that voter information generically modifies the findings of the Decentralization Theorem in favor of centralization. With homogeneous preferences, decentralization is strictly dominated not only when there are externalities, but also when information is homogeneous across regions or spreads from one region to another. Then it is no longer unambiguously true that decentralization is beneficial when preferences are heterogeneous: the costs of policy uniformity can be more than offset by the accountability benefits described by Proposition 2.

## 4 Centralization and Redistribution

The previous section focused exclusively on the aggregate efficiency properties of centralization and decentralization. Proposition 2 highlighted two informational channels through which centralization increases aggregate efficiency. Accordingly, Corollary 1 established that centralization is strictly welfare-increasing whenever preferences are homogeneous. Yet fiscal federalism also has distributional consequences when regions are heterogeneous. In our framework, redistribution occurs even if income is homogeneous, so long as voter information is not. The aggregate improvement in political accountability established by 2 is also a transfer from the more to the less informed. This transfer is most clearly displayed when heterogeneous information is the only determinant of the efficiency of centralization.



**Corollary 2** *Suppose that preferences are homogeneous, information is heterogeneous, and there are no externalities nor information spillovers across regions ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = \iota_l = 0$  for all  $l$  and  $p$ ;  $\theta_l \neq \theta_m$  for some  $l \neq m$ ). Centralization yields strictly higher aggregate welfare than decentralization. Region  $l$  is better off under centralization if and only if its voters are less informed than average ( $\theta_l \leq \bar{\theta}$ ).*

When information is heterogeneous, centralization reduces rent extraction by transferring political monitoring from the regions whose voters are more informed than average to those with less than average information. The more informed regions are better at incentivizing and selecting local politicians. The quality of their electoral process is diluted when the constituency is expanded to include regions with less informed voters, who worsen the overall monitoring ability of the electorate. Conversely, the less informed regions are plagued with rent-extracting and incompetent local governments. Centralization effectively enables them to outsource their governance to better informed voters in other regions.

The inter-regional transfer of accountability has a twofold redistributive effect. Changes in electoral screening are welfare-neutral in the aggregate, as shown in Proposition 2. The expected ability of central politicians is the same as the average of local politicians', which increases public-good provision in less informed regions as much as it decreases in better informed ones. In addition, a welfare-increasing transfer of government funds results from centralization, although every region contributes an equal amount ( $b$ ) to of the government budget and receives an identical provision of public goods ( $g$ ). In fact, the transfer is precisely a consequence of this two-sided uniformity. Under decentralization, instead, the contributions ( $b$ ) are identical, but productive government expenditures are slightly higher in more informed regions, while less informed regions suffer much higher rent extraction.

The last consideration highlights that centralization redistributes from more to less informed regions largely because of the uniformity constraint imposed on centralized public-good provision. This assumption, while standard, is not necessarily always realistic. Lockwood (2002) and Besley and Coate (2003) have modelled fiscal federalism under the alternative hypothesis that the central government can arbitrarily vary the levels of different public goods in different regions. In reality, discretionary federal spending is not required to be homogeneously spent across states, virtually by definition of discretionality. Strömberg (2004) has documented empirically that the regional allocation of discretionary government spending responds to voters' information. During the New Deal, state governors directed more public funds to counties with a greater share of radio listeners. Such a pattern of regional spending emerges in our model when the uniformity constraint is relaxed. The more informed regions provide more of the central politician's incentives, and thus they receive more government expenditure.

To analyze this phenomenon, we partition public goods into two sets. The set  $\mathcal{U}$  includes public goods whose centralized provision is subject to the uniformity constraint. The complementary set  $\mathcal{D}$  comprises public goods that the central government can provide in arbitrarily different amounts in different regions. The budget constraint of the central government then becomes

$$b^C = r_t^C + \sum_{p \in \mathcal{U}} x_{l,p,t}^C + \sum_{l=1}^L \sum_{p \in \mathcal{D}} x_{l,p,t}^C, \quad (23)$$

and each region then receives public goods

$$g_{l,p,t}^C = \frac{1}{L} e^{\eta_{p,t}} x_{p,t}^C \text{ for } p \in \mathcal{U} \text{ and } g_{l,p,t}^C = e^{\eta_{p,t}} x_{l,p,t}^C \text{ for } p \in \mathcal{D}. \quad (24)$$

As in Corollary 2, we keep our focus on heterogeneous information by removing all other forces that make centralization more efficient: information spillovers ( $\iota_l = 0$  for all  $l$ ) as well as externalities ( $\xi_p^l = 0$  for all  $l$  and  $p$ ). Moreover, we retain the assumption that preferences are homogeneous ( $\alpha_p^l = \alpha_p$  for all  $l$  and  $p$ ). Then the uniformity constraint is not binding for a benevolent central social planner, and its prevalence does not impact on the aggregate welfare ranking of centralization and decentralization. Let

$$\alpha_{\mathcal{U}} \equiv \sum_{p \in \mathcal{U}} \alpha_p \in [0, 1] \quad (25)$$

denote the welfare weight of public goods that are subject to the uniformity constraint under centralization.

**Proposition 3** *Suppose that preferences are homogeneous, information is heterogeneous, and there are no externalities nor information spillovers across regions ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = \iota_l = 0$  for all  $l$  and  $p$ ;  $\theta_l \neq \theta_m$  for some  $l \neq m$ ).*

*There exists a threshold  $\bar{\alpha}_{\mathcal{U}} \in (0, 1 - \rho^C)$  such that centralization yields higher aggregate welfare than decentralization if and only if  $\alpha_{\mathcal{U}} \geq \bar{\alpha}_{\mathcal{U}}$ .*

*There exists a threshold  $\bar{\sigma}^2 > 0$  such that centralization with an optimal uniformity constraint Pareto dominates decentralization if  $\alpha_{\mathcal{U}} = 1 - \rho^C$  and  $\sigma^2 \leq \bar{\sigma}^2$ . The gains from centralization for each region  $l$  are a convex function of its resident's information ( $\partial^2 (\mathbb{E}u_l^C - \mathbb{E}u_l^D) / \partial \theta_l^2 > 0$ ).*

The first result highlights that a uniformity constraint, which creates costs of centralization in the classic theory of fiscal federalism (Oates 1972), is instead strictly necessary for the efficiency of centralization with heterogeneously informed voters. Centralization increases overall accountability and reduces aggregate rent extraction. When public goods

are uniformly provided by the central government, the resulting increase in productive public spending benefits the regions who need it most, because their local politicians extract higher rents under decentralization. Indeed, uniformity induces an egalitarian inter-regional allocation, which is precisely what aggregate welfare maximization requires.

Instead, all public goods not subject to the uniformity constraint are provided preferentially to the most informed regions. Their provision is exactly proportional to the level of information:  $\beta_{l,p}^C/\beta_{m,p}^C = \theta_l/\theta_m$  for all  $p \in \mathcal{D}$ . The resulting geographic misallocation of government expenditures can be more inefficient than rent extraction. The share of public goods whose centralized provision is not subject to the uniformity constraint ( $1 - \alpha_U$ ) measures the share of the central budget that better informed regions can appropriate. In the limit as  $\alpha_U \rightarrow 0$ , uniformed voters certainly suffer more under centralization, when their taxes are channeled to public spending in better informed regions, than under decentralization, when they are defrauded by rent-extracting local politicians. The welfare losses of redistribution across heterogeneously informed voters loom larger than those of imperfect political agency. As evidence of such welfare costs, Ponzetto (2011) has shown that knowledge asymmetries across voters can account for the inefficient protectionist bias of trade policy.

On the other hand, Proposition 3 establishes that the uniformity constraint can mediate between efficiency and redistribution. If it can be set optimally at the constitutional table, centralization may be made Pareto dominant. Better incentives for ruling politicians create an aggregate surplus that can be shared across regions. For  $1 - \alpha_U \geq \rho^C$ , any region with more than average information ( $\theta_l > \bar{\theta}$ ) prefers the budget allocation induced by centralization. Gaining control of centrally provided goods not subject to the uniformity constraint ( $1 - \alpha_U$ ) is worth more than a decrease in rent extraction to  $\rho_l^D < \rho^C$ . For  $\rho^C \geq 1 - \alpha_U$ , any region with less than average information ( $\theta_l < \bar{\theta}$ ) prefers the budget allocation induced by centralization. Reducing rent extraction from  $\rho_l^D > \rho^C$  is worth more than the loss of control over distributive goods ( $1 - \alpha_U$ ). Hence, if  $1 - \alpha_U = \rho^C$  centralization induces a Pareto dominant allocation of expenditures.

Beyond the budget allocation, centralization inevitably improves the selection of politicians in less informed regions, and worsens it in more informed ones. Centralization can be Pareto efficient so long as the resulting gains and losses are small, because politicians' ability is not too variable (low  $\sigma^2$ ). Then the main problem in political agency is rent extraction (moral hazard) rather than the screening of more skilled politician (adverse selection). In this case, centralization can always provide a Pareto efficient improvement in accountability.

Proposition 3 concludes by noting that Pareto efficient centralization yields higher benefits to regions with more extreme levels of voter information. This distribution is intuitive: a region with average information ( $\theta_l = \bar{\theta}$ ) is simply indifferent between centralization and

decentralization. As information rises further above the average, there are greater benefits from taking effective control of the central government budget. As information falls further below the average, there are greater benefits from escaping rent-extracting local politicians.

This theoretical prediction is consistent with empirical evidence on Americans' attitudes toward the federal government. The model suggests that satisfaction with the federal government should be greater in states whose residents are less informed about politics. The relationship can be observed in survey data from the American National Elections Studies, 1964 to 2004.<sup>12</sup> The ANES asks respondents how much of the time they think they can trust the government in Washington to do what is right.<sup>13</sup> The answer to this question provides our dependent variable. We measure voter information by media exposure to electoral campaigns, constructed by the ANES as the count of the different types of media through which the respondent has followed a campaign.<sup>14</sup> Since the predictions of our theoretical model concern information and policy preferences across states rather than individual, we average both dependent and independent variables by state-year.

Table 1 displays the strong convex relationship that emerges from the data, providing suggestive evidence in support of the empirical predictions of our theory. The coefficients of a quadratic regression have the predicted signs. They are highly significant and hardly affected by the inclusion of state fixed effects, year fixed effects, or both state and year fixed effects. In the last case, the fitted quadratic has its minimum at a level of exposure to 1.86 media, reasonably close to the sample mean of 2.26. The second half of the table adds as controls the logarithms of state population and personal income per capita. Our analysis in Proposition 3 abstracted from heterogeneity in either, but more broadly they would be related to government revenues and expenditures, and could then systematically influence preferences for centralization. However, neither has a robust effect on confidence in the federal government, and their inclusion does not modify the basic finding of a convex effect of voter information.

## 5 Conclusion

Political accountability and the quality of government vary across regions within a single country, and across member states of international organizations like the European Union. In this paper, we have shown that such differences make centralization efficient because a

---

<sup>12</sup>Data are available for 1964, 1968, 1972, 1976, 1978, 1980, 1982, 1984, 1992, 1996, and 2004.

<sup>13</sup>Question VCF0604: How Much Does R Trust the Federal Govt To Do What is Right. Responses are coded on a four-point scale from "None of the time" to "Just about always."

<sup>14</sup>Question VCF0728: Media Exposure Count. Coded on a five-point scale from no media to all four media: television, radio, magazines, and newspapers.

central politician is more accountable than a collection of local ones. This result helps explain the steady growth of the federal government over the history of the United States, and the decline in corruption that accompanied its greatest expansion during the New Deal.

Our model emphasizes the role of differences in voters' information. Rent-seeking politicians have better incentives when their constituents are more informed about the provision of public goods. Crucially, we show that returns to monitoring are decreasing. Therefore a single central politician subject to average monitoring extracts lower rents than many local politicians, some monitored tightly and some loosely. Hence centralization is beneficial whenever voter information is heterogeneous across regions. Centralization is also efficient when there are information spillovers across regions. Then voters, by observing public goods across the entire union, gain more information about the central government than they have about the local politicians in their respective regions. As a consequence, decentralization worsens both politicians' incentives and their selection.

Moreover, we have shown that heterogeneous information has distributive effects within a union. Centralization slightly lowers accountability in regions with high information, and greatly raises it in those with low information. These opposite changes imply a transfer from the informed to the uninformed when public goods are uniformly provided by the central government. On the other hand, if the central politician can differentiate local public-good provision across regions, he caters to the preferences of the most informed voters, whom he is disproportionately accountable to. In this case, there is a transfer of power from the uninformed to the informed.

Thus we have identified a new role for the policy uniformity constraint. By striking a balance between uniform public goods that favor the uninformed and freely distributive goods that benefit the informed, it can make centralization Pareto efficient. Our model then implies that the benefits of centralization should be a convex function of voter information. We have shown suggestive evidence supporting this prediction. Americans' self-reported trust in the federal government is a convex function of their exposure to media coverage of politics.

# A Derivations and Proofs

## A.1. Centralization and Decentralization

Under centralization, the independent arrival of information about public goods in different regions implies that a resident of region  $l$  is informed of  $g_{m,p,t}$  for at least one region  $m$  with probability  $\theta_l + (1 - \theta_l) \Xi_l$ , where

$$\Xi_l \equiv 1 - (1 - \theta_l)^{L-1}. \quad (\text{A1})$$

Proposition 1 implies directly that rent extraction is

$$\rho^C = \left\{ 1 + 2\tilde{\delta}\phi \frac{1}{L} \sum_{l=1}^L [\theta_l + (1 - \theta_l) \Xi_l] \right\}^{-1}, \quad (\text{A2})$$

and the expected ability of a central politician is

$$\mathbb{E}\hat{\eta}_p^C = \frac{\phi\sigma^2}{L} \sum_{l=1}^L \alpha_p^l [\theta_l + (1 - \theta_l) \Xi_l]. \quad (\text{A3})$$

The relative shares of each local public good are

$$\beta_p^C = \frac{\sum_{l=1}^L \alpha_p^l [\theta_l + (1 - \theta_l) \Xi_l]}{\sum_{l=1}^L [\theta_l + (1 - \theta_l) \Xi_l]}. \quad (\text{A4})$$

Welfare in region  $l$  is

$$\mathbb{E}u_l^C = \log \frac{b^C}{L} + \log (1 - \rho^C) + \sum_{p=1}^P \alpha_p^l (\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C), \quad (\text{A5})$$

and aggregate welfare is

$$W^C = \log \frac{b^C}{L} + \log (1 - \rho^C) + \sum_{p=1}^P \bar{\alpha}_p (\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C). \quad (\text{A6})$$

Under decentralization, Proposition 1 implies directly that rent extraction is

$$\rho_l^D = \left[ 1 + 2\tilde{\delta}\phi \left( 1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p^l \alpha_p^l \right) \theta_l \right]^{-1}, \quad (\text{A7})$$

and the expected ability of a local politician is

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2 \left(1 - \frac{L-1}{L}\xi_p^l\right) \alpha_p^l \theta_l. \quad (\text{A8})$$

The relative shares of each local public good are

$$\beta_{l,p}^D = \frac{\left(1 - \frac{L-1}{L}\xi_p^l\right) \alpha_p^l}{1 - \frac{L-1}{L} \sum_{q=1}^P \xi_q^l \alpha_q^l}. \quad (\text{A9})$$

Welfare in region  $l$  is

$$\mathbb{E}u_l^D = \sum_{p=1}^P \alpha_p^l \left\{ \begin{aligned} & \left(1 - \xi_p^l\right) \left[\log b_l^D + \log(1 - \rho_l^D) + \mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D\right] \\ & + \frac{\xi_p^l}{L} \sum_{m=1}^L \left[\log b_m^D + \log(1 - \rho_m^D) + \mathbb{E}\hat{\eta}_{m,p}^D + \log \beta_{m,p}^D\right] \end{aligned} \right\}, \quad (\text{A10})$$

and aggregate welfare is

$$W^D = \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \left\{ \left[ \left(1 - \xi_p^l\right) \alpha_p^l + \frac{1}{L} \sum_{m=1}^L \xi_p^m \alpha_p^m \right] \left[\log b_l^D + \log(1 - \rho_l^D) + \mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D\right] \right\}. \quad (\text{A11})$$

## A.2. Proof of Proposition 2

### A.2.1 Rent Extraction

Aggregate rent extraction is lower under centralization if and only if

$$\rho^C \leq \frac{1}{L} \sum_{l=1}^L \rho_l^D, \quad (\text{A12})$$

which can be written

$$f \left( \frac{1}{L} \sum_{l=1}^L [\theta_l + (1 - \theta_l) \Xi_l] \right) \leq \frac{1}{L} \sum_{l=1}^L f \left( \left(1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p^l \alpha_p^l\right) \theta_l \right) \quad (\text{A13})$$

for

$$f(x) \equiv \frac{1}{1 + 2\tilde{\delta}\phi x} \quad (\text{A14})$$

a strictly decreasing and strictly convex function of  $x > 0$ :

$$f'(x) = -\frac{2\tilde{\delta}\phi}{(1+2\tilde{\delta}\phi x)^2} < 0 \text{ and } f''(x) = \frac{(2\tilde{\delta}\phi)^2}{(1+2\tilde{\delta}\phi x)^3} > 0. \quad (\text{A15})$$

Thus

$$\begin{aligned} f\left(\frac{1}{L}\sum_{l=1}^L[\theta_l + (1-\theta_l)\Xi_l]\right) &\leq f\left(\frac{1}{L}\sum_{l=1}^L\theta_l\right) \leq \frac{1}{L}\sum_{l=1}^L f(\theta_l) \\ &\leq \frac{1}{L}\sum_{l=1}^L f\left(\left(1 - \frac{L-1}{L}\sum_{p=1}^P\xi_p^l\alpha_p^l\right)\theta_l\right). \end{aligned} \quad (\text{A16})$$

The first inequality follows from monotonicity and holds strictly if  $\iota_l > 0 \Rightarrow \Xi_l > 0$  for some  $l$ . The second inequality follows from convexity by Jensen's inequality and holds strictly if  $\theta_l$  is heterogeneous across regions. The third inequality follows from monotonicity and holds strictly if  $\xi_p^l > 0$  for some  $l$  and  $p$ .

### A.2.2 Ability

Average efficiency in providing public good  $p$  is higher under centralization if and only if

$$\mathbb{E}\hat{\eta}_p^C \geq \frac{1}{L}\sum_{l=1}^L\mathbb{E}\hat{\eta}_{l,p}^D, \quad (\text{A17})$$

which can be written

$$\sum_{l=1}^L\alpha_p^l[\theta_l + (1-\theta_l)\Xi_l] \geq \sum_{l=1}^L\left(1 - \frac{L-1}{L}\xi_p^l\right)\alpha_p^l\theta_l. \quad (\text{A18})$$

The left-hand side is weakly higher than  $\sum_{l=1}^L\alpha_p^l\theta_l$ , and strictly higher if  $\iota_l > 0 \Rightarrow \Xi_l > 0$  for some  $l$ . The right-hand side is weakly lower than  $\sum_{l=1}^L\alpha_p^l\theta_l$ , and strictly lower if  $\xi_p^l > 0$  for some  $l$  and  $p$ .



### A.3. Proof of Corollary 1

#### A.3.1 The Decentralization Theorem

Suppose that information is homogeneous and does not spill over regional boundaries ( $\theta_l = \theta$  and  $\iota_l = 0$  for all  $l$ ). Then under centralization

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\theta\frac{1}{L}\sum_{l=1}^L\alpha_p^l, \text{ and } \beta_p^C = \frac{1}{L}\sum_{l=1}^L\alpha_p^l. \quad (\text{A19})$$

while under decentralization

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\theta\left(1 - \frac{L-1}{L}\sum_{p=1}^P\xi_p^l\alpha_p^l\right)\right]^{-1}, \quad (\text{A20})$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\left(1 - \frac{L-1}{L}\xi_p^l\right)\alpha_p^l, \quad (\text{A21})$$

and

$$\beta_{l,p}^D = \left(1 - \frac{L-1}{L}\xi_p^l\right)\alpha_p^l. \quad (\text{A22})$$

1. If there are no externalities and preferences are homogeneous ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = 0$  for all  $l$  and  $p$ ), then

$$\rho^C = \rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \mathbb{E}\hat{\eta}_p^C = \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\alpha_p \text{ and } \beta_p^C = \beta_{l,p}^D = \alpha_p. \quad (\text{A23})$$

2. If there are externalities and preferences are homogeneous ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = \xi_p$  for all  $l$ , with  $\xi_p > 0$  for some  $p$ ), then under centralization

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\theta\alpha_p, \text{ and } \beta_p^C = \alpha_p, \quad (\text{A24})$$

and

$$W^C = \log b + \log(1 - \rho^C) + \sum_{p=1}^P\alpha_p(\mathbb{E}\hat{\eta}_p^C + \log\beta_p^C). \quad (\text{A25})$$

Under decentralization

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\theta\left(1 - \frac{L-1}{L}\sum_{p=1}^P\xi_p\alpha_p\right)\right]^{-1} > \rho^C, \quad (\text{A26})$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta \left(1 - \frac{L-1}{L}\xi_p\right) \alpha_p < \mathbb{E}\hat{\eta}_p^C, \quad (\text{A27})$$

$$\beta_{l,p}^D = \frac{\left(1 - \frac{L-1}{L}\xi_p\right) \alpha_p}{1 - \frac{L-1}{L} \sum_{q=1}^P \xi_q \alpha_q}, \quad (\text{A28})$$

and

$$W^D = \log b + \log(1 - \rho_l^D) + \sum_{p=1}^P \alpha_p (\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D). \quad (\text{A29})$$

Centralization achieves the optimal allocation of productive expenditures, while decentralization does not unless  $\xi_p$  is homogeneous across goods.

Thus welfare is lower under decentralization due to increased rent extraction, lower government efficiency, and also misallocation of expenditures across public goods unless  $\xi_p$  is homogeneous across goods.

3. If there are no externalities and preferences are heterogeneous ( $\xi_p^l = 0$  for all  $l$  and  $p$  while  $\alpha_p^l \neq \alpha_p^m$  for some  $l \neq m$  and  $p$ ), then under centralization

$$W^C = \log b + \log(1 - \rho^C) + \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l (\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C), \quad (\text{A30})$$

while under decentralization

$$\rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1} = \rho^C, \quad \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\alpha_p^l, \quad \beta_{l,p}^D = \alpha_p^l \quad (\text{A31})$$

and

$$W^D = \log b + \log(1 - \rho_l^D) + \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l (\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D). \quad (\text{A32})$$

Decentralization achieves the optimal allocation of productive expenditures, while centralization does not. Moreover, decentralization achieves a better screening of politicians

$$\frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l \mathbb{E}\hat{\eta}_{l,p}^D > \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l \mathbb{E}\hat{\eta}_p^C \quad (\text{A33})$$

because for each public good  $p$

$$\frac{1}{L} \sum_{l=1}^L (\alpha_p^l)^2 > \left(\frac{1}{L} \sum_{l=1}^L \alpha_p^l\right)^2 \quad (\text{A34})$$

unless  $\alpha_p^l = \alpha_p$  for all  $l$ .

### A.3..2 Efficient Centralization

Suppose that preferences are homogeneous ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = \xi_p$  for all  $l$ ). Then under centralization

$$\rho^C = \left\{ 1 + 2\tilde{\delta}\phi \frac{1}{L} \sum_{l=1}^L [\theta_l + (1 - \theta_l) \Xi_l] \right\}^{-1}, \quad (\text{A35})$$

$$\mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\alpha_p \frac{1}{L} \sum_{l=1}^L [\theta_l + (1 - \theta_l) \Xi_l], \quad (\text{A36})$$

$$\beta_p^C = \alpha_p, \quad (\text{A37})$$

and

$$W^C = \log b + \log(1 - \rho^C) + \sum_{p=1}^P \alpha_p (\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C). \quad (\text{A38})$$

Under decentralization

$$\rho_l^D = \left[ 1 + 2\tilde{\delta}\phi \left( 1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p \alpha_p \right) \theta_l \right]^{-1}, \quad (\text{A39})$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2 \left( 1 - \frac{L-1}{L} \xi_p \right) \alpha_p \theta_l, \quad (\text{A40})$$

$$\beta_{l,p}^D = \frac{\left( 1 - \frac{L-1}{L} \xi_p \right) \alpha_p}{1 - \frac{L-1}{L} \sum_{q=1}^P \xi_q \alpha_q}, \quad (\text{A41})$$

and

$$W^D = \frac{1}{L} \sum_{l=1}^L \left[ \log b + \log(1 - \rho_l^D) + \sum_{p=1}^P \alpha_p (\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D) \right]. \quad (\text{A42})$$

1. The welfare cost of rent extraction falls with centralization:

$$\log(1 - \rho^C) > \frac{1}{L} \sum_{l=1}^L \log(1 - \rho_l^D) \quad (\text{A43})$$

which can be written

$$f\left(\frac{1}{L} \sum_{l=1}^L [\theta_l + (1 - \theta_l) \Xi_l]\right) > \frac{1}{L} \sum_{l=1}^L f\left(\left(1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p \alpha_p\right) \theta_l\right) \quad (\text{A44})$$

for

$$f(x) \equiv \log x - \log\left(1 + 2\tilde{\delta}\phi x\right) \quad (\text{A45})$$

a strictly increasing and strictly concave function of  $x > 0$ :

$$f'(x) = \frac{1}{x(1 + 2\tilde{\delta}\phi x)} > 0 \text{ and } f''(x) = -\frac{1 + 4\tilde{\delta}\phi x}{\left[x(1 + 2\tilde{\delta}\phi x)\right]^2} < 0. \quad (\text{A46})$$

Thus

$$\begin{aligned} f\left(\frac{1}{L}\sum_{l=1}^L[\theta_l + (1 - \theta_l)\Xi_l]\right) &\geq f\left(\frac{1}{L}\sum_{l=1}^L\theta_l\right) \geq \frac{1}{L}\sum_{l=1}^L f(\theta_l) \\ &\geq \frac{1}{L}\sum_{l=1}^L f\left(\left(1 - \frac{L-1}{L}\sum_{p=1}^P\xi_p\alpha_p\right)\theta_l\right). \end{aligned} \quad (\text{A47})$$

The first inequality follows from monotonicity and holds strictly if  $\iota_l > 0 \Rightarrow \Xi_l > 0$  for some  $l$ . The second inequality follows from convexity by Jensen's inequality and holds strictly if  $\theta_l$  is heterogeneous across regions. The third inequality follows from monotonicity and holds strictly if  $\xi_p > 0$  for some  $p$ .

2. Average ability increases under centralization for all public goods, as proved in Proposition 2.
3. Centralization achieves the optimal allocation of productive expenditures, while decentralization does not unless  $\xi_p$  is homogeneous across public goods.

Thus centralization increases welfare whenever information is heterogeneous ( $\theta_l \neq \theta_m$  for some  $l \neq m$ ), there are information spillovers ( $\iota_l > 0$  for some  $l$ ), or there are externalities in the provision of public goods ( $\xi_p > 0$  for some  $p$ ).

#### A.4. Proof of Corollary 2

Suppose that preferences are homogeneous, information is heterogeneous, and there are no externalities nor information spillovers across regions ( $\alpha_p^l = \alpha_p$  and  $\xi_p^l = \iota_l = 0$  for all  $l$  and  $p$ ). Then residents of region  $l$  have welfare

$$\mathbb{E}u_l^D = \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi\sigma^2\theta_l \sum_{p=1}^P \alpha_p^2 + \log \frac{2\tilde{\delta}\phi\theta_l}{1 + 2\tilde{\delta}\phi\theta_l} \quad (\text{A48})$$

under decentralization, and

$$\mathbb{E}u_l^C = \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi \sigma^2 \bar{\theta} \sum_{p=1}^P \alpha_p^2 + \log \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} \quad (\text{A49})$$

under centralization. Thus region  $l$  gains from centralization if and only if  $\theta_l \leq \bar{\theta}$ .

### A.5. Proof of Proposition 3

The allocation under decentralization is unaffected by  $a_U < 1$ . Aggregate welfare is

$$W^D = \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi \sigma^2 \bar{\theta} \sum_{p=1}^P \alpha_p^2 + \frac{1}{L} \sum_{l=1}^L \log \frac{2\tilde{\delta}\phi\theta_l}{1 + 2\tilde{\delta}\phi\theta_l} \quad (\text{A50})$$

From Proposition 1, rent extraction under centralization is

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\bar{\theta}\right)^{-1} \text{ for } \bar{\theta} = \frac{1}{L} \sum_{l=1}^L \theta_l \quad (\text{A51})$$

and the expected ability of a central politician is

$$\mathbb{E}\hat{\eta}_p^C = \phi \sigma^2 \bar{\theta} \alpha_p. \quad (\text{A52})$$

The relative shares of each local public good are

$$\beta_p^C = \alpha_p \text{ for } p \in \mathcal{U} \text{ and } \beta_{l,p}^C = \frac{1}{L} \frac{\theta_l}{\bar{\theta}} \alpha_p \text{ for } p \in \mathcal{D}. \quad (\text{A53})$$

Welfare in region  $l$  is

$$\mathbb{E}u_l^C = \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi \sigma^2 \bar{\theta} \sum_{p=1}^P \alpha_p^2 + \log \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} + (1 - \alpha_U) (\log \theta_l - \log \bar{\theta}), \quad (\text{A54})$$

and aggregate welfare is

$$W^C = \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi \sigma^2 \bar{\theta} \sum_{p=1}^P \alpha_p^2 + \log \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} + (1 - \alpha_U) \left( \frac{1}{L} \sum_{l=1}^L \log \theta_l - \log \bar{\theta} \right). \quad (\text{A55})$$

Then aggregate welfare is greater under centralization if

$$\alpha_U \geq \frac{\log \left( 1 + 2\tilde{\delta}\phi\bar{\theta} \right) - \frac{1}{L} \sum_{l=1}^L \log \left( 1 + 2\tilde{\delta}\phi\theta_l \right)}{\log \bar{\theta} - \frac{1}{L} \sum_{l=1}^L \log \theta_l} \equiv \bar{\alpha}_U \in (0, 1). \quad (\text{A56})$$

Region  $l$  gains from centralization if

$$\log \frac{1 + 2\tilde{\delta}\phi\theta_l}{1 + 2\tilde{\delta}\phi\bar{\theta}} - \alpha_U \log \frac{\theta_l}{\bar{\theta}} \geq \phi \sigma^2 \sum_{p=1}^P \alpha_p^2 (\theta_l - \bar{\theta}). \quad (\text{A57})$$

The function

$$f(x) \equiv \log \left( 1 + 2\tilde{\delta}\phi x \right) - \alpha_U \log x \quad (\text{A58})$$

has a unique minimum

$$f'(x) = \frac{2\tilde{\delta}\phi}{1 + 2\tilde{\delta}\phi x} - \frac{\alpha_U}{x} \geq 0 \Leftrightarrow x \geq \frac{\alpha_U}{2\tilde{\delta}\phi(1 - \alpha_U)}, \quad (\text{A59})$$

and convexity

$$f''(x) = - \left( \frac{2\tilde{\delta}\phi}{1 + 2\tilde{\delta}\phi x} \right)^2 + \frac{\alpha_U}{x^2} > 0 \Leftrightarrow \alpha_U > \left( \frac{2\tilde{\delta}\phi x}{1 + 2\tilde{\delta}\phi x} \right)^2 \quad (\text{A60})$$

Hence if

$$\alpha_U = \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} = 1 - \rho^C \quad (\text{A61})$$

the left-hand side of condition A57 is a convex function of  $\theta_l \in [0, 1]$  with minimum at  $\theta_l = \bar{\theta}$ .

This also proves that

$$\bar{\alpha}_U < 1 - \rho^C. \quad (\text{A62})$$

Moreover, let

$$m = \arg \min_{l=1, \dots, L} \{ \theta_l : \theta_l > \bar{\theta} \}. \quad (\text{A63})$$

Then for

$$\sigma^2 \leq \frac{1}{\phi \sum_{p=1}^P \alpha_p^2 (\theta_m - \bar{\theta})} \left\{ \log \frac{1 + 2\tilde{\delta}\phi\theta_m}{1 + 2\tilde{\delta}\phi\bar{\theta}} - \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} \log \frac{\theta_m}{\bar{\theta}} \right\} \equiv \bar{\sigma}^2 > 0, \quad (\text{A64})$$

centralization Pareto dominates decentralization for  $\alpha_U = \rho^C$ . Any region with  $\theta_l = \bar{\theta}$  is always indifferent between the two. Region  $m$  is also indifferent if  $\sigma^2 = \bar{\sigma}^2$ , and strictly prefers centralization otherwise. Any regions with  $\theta_l < \bar{\theta}$  or  $\theta_l > \theta_m$  strictly prefer centralization. Since the left-hand side of condition A57 is convex and the right-hand side linear, the benefits from centralization are a convex function of  $\theta_l$ .

## References

- [1] Alesina, Alberto, and Robert J. Barro. 2002. “Currency Unions.” *Quarterly Journal of Economics* 117(2):409–36.
- [2] Alesina, Alberto, Ignazio Angeloni, and Federico Etro. 2005. “International Unions.” *American Economic Review* 95(3):602–15.
- [3] Alesina, Alberto, Ignazio Angeloni, and Ludger Schuknecht 2005. “What Does the European Union Do?” *Public Choice* 123:275–319.
- [4] Alesina, Alberto, and Enrico Spolaore. 1997. “On the Number and Size of Nations.” *Quarterly Journal of Economics* 112(4):1027–56.
- [5] Bardhan, Pranab, and Dilip Mookherjee. 2000. “Capture and Governance at Local and National Levels.” *American Economic Review* 90(2):135–9.
- [6] ——. 2006. “Decentralization, Corruption, and Government Accountability.” In *International Handbook on the Economics of Corruption*, edited by Susan Rose-Ackerman, 161–88. Cheltenham: Edward Elgar.
- [7] Besley, Timothy, and Robin Burgess. 2002. “The Political Economy of Government Responsiveness: Theory and Evidence from India.” *Quarterly Journal of Economics* 117(4):1415–51.
- [8] Besley, Timothy, and Anne Case. 1995. “Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition.” *American Economic Review* 85(1):25–45.
- [9] Besley, Timothy, and Stephen Coate. 2003. “Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach.” *Journal of Public Economics* 87:2611–37.
- [10] Besley, Timothy, and Michael Smart. 2007. “Fiscal Restraints and Voter Welfare.” *Journal of Public Economics* 91:755–73.
- [11] Blanchard, Olivier, and Andrei Shleifer. 2001. “Federalism With and Without Political Centralization: China Versus Russia.” *IMF Staff Papers* 48(4):171–9.
- [12] Bulte, Erwin, John List, and Mark Strazicich. 2007. “Regulatory Federalism and the Distribution of Air Pollutant Emissions”, *Journal of Regional Science*, 47(1):155–78.



- [13] Charron, Nicholas. 2010. "The Correlates of Corruption in India: Analysis and Evidence from the States." *Asian Journal of Political Science* 18(2):177–94.
- [14] Charron, Nicholas, Lewis Dijkstra, and Victor Lapuente. 2011. "Bringing the Regions Back: European National and Sub-National Variation in Quality of Government and Five Hypotheses." Mimeo, University of Gothenburg.
- [15] Del Monte, Alfredo, and Erasmo Papagni. 2001. "Public Expenditure, Corruption and Economic Growth: The Case of Italy." *European Journal of Political Economy* 17:1–16.
- [16] ——. 2007. "The Determinants of Corruption in Italy: Regional Panel Data Analysis." *European Journal of Political Economy* 23:379–96.
- [17] Dininio, Phyllis, and Robert W. Orttung. 2005. "Explaining Patterns of Corruption in the Russian Regions." *World Politics* 57(4):500–29.
- [18] Dong, Bin, and Benno Torgler. 2010. "The Causes of Corruption: Evidence from China." Fondazione Eni Enrico Mattei working paper no. 461.
- [19] Durante, Ruben, Giovanna Labartino, and Roberto Perotti. 2011. "Academic Dynasties: Decentralization and Familism in the Italian Academia." NBER Working Paper No 17572.
- [20] Fan, C. Simon, Chen Lin, and Daniel Treisman. 2009. "Political Decentralization and Corruption: Evidence from Around the World." *Journal of Public Economics* 93(1-2):14–34.
- [21] Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2):703–45.
- [22] ——. 2011. "Electoral Accountability and Corruption: Evidence from the Audit Reports of Local Governments." *American Economic Review* 101:1274–311.
- [23] Francken, Nathalie, Bart Minten, and Joahn F. M. Swinnen. 2009. "Media, Monitoring, and Capture of Public Funds: Evidence from Madagascar." *World Development* 37(1):242–55.
- [24] Fredriksson, Per and Noel Gaston. 2000. "Environmental Governance in Federal Systems: the Effects of Capital Competition and Lobby Groups." *Economic Inquiry*, 38(3):501–14.

- [25] Gennaioli, Nicola, and Ilija Rainer. 2007. “The Modern Impact of Precolonial Centralization in Africa.” *Journal of Economic Growth* 12(3):185–234.
- [26] Gentzkow, Matthew. 2006. “Television and voter turnout.” *Quarterly Journal of Economics* 121(3):931–72.
- [27] Glaeser, Edward L., and Claudia Goldin. 2006. “Corruption and Reform: an Introduction”, in *Corruption and Reform: Lessons from America’s Economic History*, edited by Edward L. Glaeser and Claudia Goldin. Cambridge, MA: National Bureau of Economic Research.
- [28] Glaeser, Edward L., and Raven E. Saks. 2006. “Corruption in America.” *Journal of Public Economics* 90:1053–72.
- [29] Golden, Miriam A., and Lucio Picci. 2005. “Proposal for a New Measure of Corruption, Illustrated with Italian Data.” *Economics and Politics* 17(1):37–75.
- [30] Inman, Robert and Daniel Rubinfeld. 1997. “Rethinking Federalism.” *Journal of Economic Perspectives*, 11(4):43–64.
- [31] Lindbeck, Assar, and Jörgen W. Weibull. 1987. “Balanced-Budget Redistribution as the Outcome of Political Competition.” *Public Choice* 52:273–97.
- [32] Lockwood, Ben. 2002. “Distributive Politics and the Costs of Centralization.” *Review of Economic Studies* 69:313–37.
- [33] ——. 2006. “The Political Economy of Decentralization.” In *Handbook of Fiscal Federalism*, edited by Ehtisham Ahmad and Giorgio Brosio, 33–60. Cheltenham: Edward Elgar.
- [34] Nannicini, Tommaso, Andrea Stella, Guido Tabellini, and Ugo Troiano. 2010. “Social Capital and Political Accountability.” Mimeo, Bocconi University.
- [35] Oates, Wallace E. 1972. *Fiscal Federalism*. New York, NY: Harcourt Brace Jovanovich.
- [36] ——. 1999. “An Essay on Fiscal Federalism.” *Journal of Economic Literature* 37:1120–49.
- [37] Pashigian, Peter. 1985. “Environmental Regulation: Whose Self Interests Are Being Protected?.” *Economic Inquiry* 23(4):551–84.
- [38] Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.

- [39] Ponzetto, Giacomo A. M. 2011. “Heterogeneous Information and Trade Policy.” CEPR Discussion Paper No. 8726.
- [40] Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.
- [41] Reinikka, Ritva, and Jakob Svensson. 2004. “Local Capture: Evidence from a Central Government Transfer Program in Uganda.” *Quarterly Journal of Economics* 119(2):679–705.
- [42] Seabright, Paul. 1996. “Accountability and Decentralization in Government: An Incomplete Contracts Model.” *European Economic Review* 40:61–89.
- [43] Snyder, James M., and David Strömberg. 2010. “Press Coverage and Political Accountability.” *Journal of Political Economy* 118(2):335–408.
- [44] Strömberg, David. 2004. “Radio’s Impact on Public Spending.” *Quarterly Journal of Economics* 99:189–221.
- [45] Treisman, Daniel. 2007. “What Have We Learned About the Causes of Corruption From Ten Years of Cross-National Empirical Research?” *Annual Review of Political Science* 10:211–44.
- [46] Wallis, John J. 2000a. “American Government Finance in the Long Run: 1790 to 1990.” *Journal of Economic Perspectives*, 14(1):61–82.
- [47] ——. 2000b. “The Concept of Systematic Corruption in American History” in *Corruption and Reform: Lessons from America’s Economic History*, edited by Edward L. Glaeser and Claudia Goldin. Cambridge, MA: National Bureau of Economic Research.
- [48] Wallis, John J., Fishback, Price, and Shawn Kantor. 2006. “Politics, Relief, and Reform: Roosevelt’s Efforts to Control Corruption and Political Manipulation during the New Deal”, in *Corruption and Reform: Lessons from America’s Economic History*, edited by Edward L. Glaeser and Claudia Goldin. Cambridge, MA: National Bureau of Economic Research.
- [49] White, Lawrence J. 1982. *The Regulation of Air Pollutant Emissions from Motor Vehicles*. Washington, DC: American Enterprise Institute for Public Policy Research.

**Table 1 – Media Exposure and Trust in the Federal Government**

Dependent variable: Trust in the Federal Government				
	(1)	(2)	(3)	(4)
Media exposure	-.3423722*	-.4304537**	-.4126927***	-.4685202***
	(.1827157)	(.2031875)	(.1530012)	(.1259238)
Media exposure, square	.1084682***	.1369271***	.0985037***	.1075891***
	(.0420818)	(.0490059)	(.0348852)	(.0302922)
Fixed Effects	—	State	Year	State + Year
Observations	400	400	400	400

  

	(5)	(6)	(7)	(8)
Media exposure	-.5841848***	-.5768846**	-.4500529***	-.4625972***
	(.1459)	(.142168)	(.1566004)	(.1254024)
Media exposure, square	.141691***	.1281098***	.1105005***	.1067344***
	(.0345693)	(.0361978)	(.0350083)	(.0300281)
Population, log	.0361533***	.1347109	.0226981*	.0348617
	(.0126)	(.1148049)	(.0127268)	(.0855645)
Personal income, log	-.1527458***	-.191511***	-.0808362	.2731067*
	(.015765)	(.0261784)	(.0795115)	(.1607037)
Fixed Effects	—	State	Year	State + Year
Observations	400	400	400	400

*Notes:* Trust in the federal government and media exposure are from the American National Election Studies, averaged by state-year. State population is from the U.S. Census. Per capita personal income is from the Bureau of Economic Analysis. Robust standard errors in parentheses. \*\*\*  $p = 1\%$ , \*\*  $p = 5\%$ , \*  $p = 10\%$ .