

CentER 

Discussion Paper

No. 2012-033

**CENTRALIZATION AND ACCOUNTABILITY:
THEORY AND EVIDENCE FROM THE CLEAN AIR ACT**

By

Federico Boffa, Amedeo Piolatto,
Giacomo A. M. Ponzetto

April 16, 2012

ISSN 0924-7815

Centralization and Accountability: Theory and Evidence from the Clean Air Act*

Federico Boffa

Università di Macerata and Barcelona Institute of Economics (IEB)

Amedeo Piolatto

Barcelona Institute of Economics (IEB), Universitat de Barcelona

Giacomo A. M. Ponzetto

CREI, Universitat Pompeu Fabra, and Barcelona GSE

First draft, 15 October 2011 – This draft, 12 April 2012

Abstract

This paper studies fiscal federalism when voter information varies across regions. We develop a model of political agency with heterogeneously informed voters. Rent-seeking politicians provide public goods to win the votes of the informed. As a result, rent extraction is lower in regions with higher information. In equilibrium, electoral discipline has decreasing returns. Thus, political centralization efficiently reduces aggregate rent extraction. The model predicts that a region's benefits from centralization are decreasing in its residents' information. We test this prediction using panel data on pollutant emissions across U.S. states. The 1970 Clean Air Act centralized environmental policy at the federal level. In line with our theory, we find that centralization induced a differential decrease in pollution for uninformed relative to informed states.

*We are grateful to Peter Backus, Vasco Carvalho, Ana Nuevo, Luigi Pascali, Thijs van Rens, Hannes Schwandt, Paolo Surico, and seminar participants at CREI, Erasmus Univeristy Rotterdam, and Toulouse School of Economics 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 most dramatic episode of centralization undertaken by the U.S. federal government achieved a striking decrease in corruption. The New Deal, by introducing federal oversight of welfare spending, eradicated the patronage and political manipulation that had hitherto characterized relief programs managed at the state and local level (Wallis 2000a, b; Wallis, Fishback, and Kantor 2006). International evidence highlights similar instances of a positive impact of political centralization on government accountability. Centralized political institutions in precolonial Africa reduced corruption and fostered the rule of law, causing a long-lasting increase in the provision of public goods that endured into the postcolonial period (Gennaioli and Rainer 2007a,b). Blanchard and Shleifer (2001) argue that China has grown faster than Russia thanks to the greater strength of its central government compared to local politicians. While contemporary cross-country studies of decentralization and perceived corruption have yielded conflicting results (Treisman 2007; Fan, Lin, and Treisman 2009), it is clear empirically that centralization can increase government efficiency and political accountability, at least under the appropriate conditions.

This phenomenon is difficult to understand through the lens of traditional models of fiscal federalism. Oates's (1972, 1999) classic theory does not consider the problem of imperfectly accountable politicians and posits two technological rationales for centralization: economies of scale and benefits from policy coordination. The subsequent literature in political economy has mostly emphasized the advantages of decentralization (Lockwood 2006). In particular, decentralized government is held to be more accountable, thanks to yardstick competition across local jurisdictions (Besley and Case 1995; Besley and Smart 2007).¹

In this paper, on the other hand, we explain how centralization can in fact enhance accountability. Our approach reflects for electoral incentives the fundamental intuition of gains from trade between heterogeneous agents. Voters in different regions are unequally capable of incentivizing self-interested politicians. If the regions are united in a single national polity, the central politician is mainly held accountable by the most capable voters. Hence, his incentives and performance are better than those of the average local politician.

We capture this idea through a model of political agency with imperfectly informed voters. Selfish politicians are entrusted with allocating the government budget between public goods and wasteful private rents. Career concerns induce the incumbent to provide a positive amount of public goods in order to signal ability and win the votes of those citizens who observe public good provision. In equilibrium, we show that politicians extract lower rents

¹Conversely, 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).

if voters are better informed, consistent with the evidence that voter information improves accountability (Besley and Burgess 2002; Ferraz and Finan 2008; Snyder and Strömberg 2010; Ponzetto 2011). Electoral discipline has decreasing returns because the threat of being ousted from office is less costly when rent extraction is already low. National elections, therefore, provide much better incentives and screening than local elections in the least informed regions, and not much worse than in the most informed ones. Centralization then increases overall efficiency by reducing aggregate political rents.

Our model predicts the regional distribution of the efficiency gains from centralization. The central government sets a uniform national policy that entails an identical level of public good provision throughout the country (Oates 1972; Alesina and Spolaore 1997). As a consequence, all regions whose residents have below-average political information derive a strict welfare increase from a switch from decentralized to centralized government. The net impact on better informed regions is instead ambiguous. The average information exploited in electing their rulers falls when they join the less informed parts of the union in a national election. This dilution tends to reduce accountability. On the other hand, if public goods spill over across regions, we show that all voters are keener on the ability of a central than a local politician. This keenness makes national elections more competitive, raising accountability. While the most informed regions may gain or lose from centralization, our theory entails an unambiguous relative prediction. A region's welfare gains from centralization are strictly decreasing in its residents' information.

We test this prediction of our model using panel data on air pollutant emissions across U.S. states. This series enables us to study the effects of one of the most prominent instances of centralization in U.S. history since World War II (Greenstone 2004). With the 1970 Clean Air Act Amendments, the federal government took direct responsibility for air pollution regulation, which until that point had rested primarily with states and local governments. We perform a difference-in-differences analysis to assess whether the federal takeover of environmental policy had a differential impact on states according to their level of information.

The empirical evidence supports our theoretical prediction. The introduction of federal standards is associated with a highly significant differential trend. After 1970, pollutant emissions begin to decline markedly in less informed states relative to more informed ones. The finding obtains for different recorded emissions, and it is robust to specifications that control for simultaneous differences-in-differences based on other potential determinants of pollution, such as income per capita or the size of the manufacturing sector. In line with our theory, the differential reduction in emissions seems to represent a benefit of centralization for the less informed, since we do not observe a simultaneous differential reduction in economic activity.

Both theoretically and empirically, we find that centralization entails systematic distributional consequences that favor regions with less informed voters, whose local government is more dysfunctional. This pattern contrasts with Strömberg’s (2004) finding that discretionary New Deal spending was disproportionately allocated to more informed counties within a state. We conclude by showing that a natural extension of our theoretical framework encompasses and interprets the two opposite phenomena. Under centralization, uniformly provided public goods such as national emission standards benefit the less informed; conversely, discretionary funds are targeted to the more informed. Our theory highlights the importance of striking a balance between the two. Without any uniformity, centralization would be welfare reducing despite the associated reduction in political rents. When instead uniform and discretionary items are in the right proportions, centralization is not only welfare increasing but also Pareto efficient.

2 Theoretical Model

2.1 Public Good Provision and Political Agency

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 $\delta \in (0, 1]$.

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 σ^2 .

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

1. The incumbent politician's past competence shock ε_{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 ε_t .
3. ε_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 θ_i ; with probability $1 - \theta_i$ he remains completely uninformed about $g_{p,t}$.² The arrival of information is independent across voters. No voter can have any direct observation of ε_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 λ_j of voters, who have identical preferences α_p^j , and identical probabilities θ_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 Ω_t^i , individual i has policy preferences

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

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

²This 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 $g_{p,t}$.

incumbent if and only if

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

where Ψ_t and ψ_t^i are independent draws from common-knowledge probability distributions. The common shock Ψ_t to the incumbent's popularity accounts for the aggregate uncertainty in the electoral outcome. The idiosyncratic shock ψ_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 Ψ_t and ψ_t^i is sufficiently wide, and that of the competence shocks $\varepsilon_{p,t}$ sufficiently narrow, that*

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

where

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

2.2 Voter Information and Government Accountability

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 i 's information is described by the set Ω_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

$$\Delta_i(\Omega_t^i) = \sum_{p \in \Omega_t^i} \alpha_p^i \mathbb{E}(\varepsilon_{p,t} | g_{p,t}) = \sum_{p \in \Omega_t^i} \alpha_p^i (\log g_{p,t} - \log \bar{x}_p - \varepsilon_{p,t-1}). \quad (7)$$

Each group j comprises a continuum of agents and the arrival of information is independent across agents, so a share θ_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 ψ^i , the share of members of group j who vote for the incumbent is

$$v_t^j = \frac{1}{2} + \frac{1}{2\bar{\psi}} \left[\theta_j \sum_{p=1}^P \alpha_p^j (\log g_{p,t} - \log \bar{x}_p - \varepsilon_{p,t-1}) - \Psi_t \right], \quad (8)$$

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

$$\pi(\mathbf{x}_t) = \frac{1}{2} + \phi \sum_{j=1}^J \lambda_j \theta_j \sum_{p=1}^P \alpha_p^j (\log x_{p,t} - \log \bar{x}_p) \quad (9)$$

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

$$\mathbf{x} = \arg \max_{\mathbf{x}_t} \left\{ b - \sum_{p=1}^P x_{p,t} + R\pi(\mathbf{x}_t) \right\}, \quad (10)$$

namely

$$x_p = \phi R \sum_{j=1}^J \lambda_j \theta_j \alpha_p^j \text{ for all } p = 1, \dots, P, \quad (11)$$

and thus current rent extraction

$$r = b - \phi \bar{\theta} R. \quad (12)$$

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

$$R = \delta \sum_{t=0}^{\infty} \left(\frac{\delta}{2} \right)^t r = \frac{2\delta}{2-\delta} r. \quad (13)$$

Substituting this relationship into equations 11 and 12, solving the latter, and plugging it into the former, we obtain the equilibrium allocation of the government budget. To economize on notation, we define

$$\tilde{\delta} \equiv \frac{\delta}{2-\delta} \in (0, 1], \quad (14)$$

a convenient rescaling of the politician's discount factor.

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

$$\Psi_t \leq \sum_{j=1}^J \lambda_j \theta_j \sum_{p=1}^P \alpha_p^j \varepsilon_{p,t}. \quad (15)$$

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

$$\hat{\eta}_t = \chi_{t-1} (\varepsilon_{t-1}^I + \varepsilon_t^I) + (1 - \chi_{t-1}) (\varepsilon_{t-1}^C + \varepsilon_t^C), \quad (16)$$

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

$$\mathbb{E}\hat{\eta}_{p,t} = \mathbb{E}(\chi_{t-1}\varepsilon_{p,t-1}). \quad (17)$$

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

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

His expected ability at providing public good p is

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

An increase in $\phi \sum_{j=1}^J \lambda_j \theta_j \alpha_p^j$ induces an increase in $\hat{\eta}_p$ in the sense of first-order stochastic dominance.

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. This result accords with the empirical 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. Rents are also reduced when voters are more likely to be swayed by policy than non-policy considerations ($\partial\rho/\partial\phi < 0$). Finally, 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.

Intuitively, it is impossible for rents to fall to zero, since the politician's incentive to refrain from appropriating the entire current budget is entirely due to his desire to gain reelection and keep pocketing future rents. More broadly, reductions in rent extraction through electoral discipline become progressively more difficult as the equilibrium rent declines. All rent-decreasing factors, namely information $\bar{\theta}$, competitiveness ϕ , and patience δ , have positive but decreasing returns ($\partial^2 \rho / \partial \bar{\theta}^2 > 0$, $\partial^2 \rho / \partial \phi^2 > 0$, $\partial^2 \rho / \partial \delta^2 > 0$).

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 (σ^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 ($\partial \mathbb{E} \hat{\eta}_p / \partial \phi > 0$). These effects raise not only the expectation of ability, but its entire distribution, in the sense of first-order stochastic dominance.

Through improvements in both political incentives and political selection, higher voter information ($\bar{\theta}$) increases the equilibrium welfare of each voter, which is given by

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

for each member of group j . Welfare also rises with voters' keenness on policy outcomes (ϕ), politicians' patience (δ), and the variance of their ability distribution (σ^2).

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

$$\beta_p = \bar{\alpha}_p \equiv \sum_{j=1}^J \lambda_j \alpha_p^j, \quad (19)$$

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 θ^j .

2.3 Efficient Centralization

The economy is divided into L regions, each 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 location l at time t . Politicians are drawn independently across regions, from identical local pools.

In keeping with the classic theory of fiscal federalism (Oates 1972), we admit the possibility of externalities in public-good provision. These externalities are measured by an index $\xi_p^l \in [0, 1]$. A resident of region l derives utility

$$\alpha_{l,p}^l = \left(1 - \frac{L-1}{L}\xi_p^l\right) \alpha_p^l > 0 \quad (20)$$

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

$$\alpha_{m,p}^l = \frac{1}{L}\xi_p^l \alpha_p^l \text{ for } l \neq m \quad (21)$$

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

$$u_t^i = \tilde{u}_t^i + \sum_{p=1}^P \left[(1 - \xi_p^i) \alpha_p^i \log g_{l,p,t} + \frac{1}{L} \xi_p^i \sum_{m=1}^L \alpha_p^i \log g_{m,p,t} \right]. \quad (22)$$

With decentralized government, 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 ρ_l^D .

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

$$g_{l,p,t}^C = \frac{1}{L} e^{\eta_{p,t}} x_{p,t}^C. \quad (23)$$

This formulation implies that there are no economies of scale in public-good provision, which removes the most obvious force pushing 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 another simple 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 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.

The following proposition establishes the beneficial effect of centralization on political accountability.

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 $l \neq m$), 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 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 voter information, this represents a large improvement, because the increase in accountability is powerful given the low starting point of their political accountability under decentralization. For regions with high voter information, 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 also increases political accountability 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). But the classic theory only considers the benefits of coordination, abstracting from any political-economy considerations. Proposition 2 finds 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 the election, 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 poor skills may nonetheless vote for him because of his personal likability or

ideological affinity. He is, however, less likely to do so as the economic stakes in the election rise. If there are public good spillovers, the stakes are indeed higher in a national than in a local election. For $\xi > 0$, every voter is keener on electing a proficient politician at the central rather than at the regional level, because his ability will affect public goods in all regions, and each voter cares about them all. Thus centralization reduces the influence of non-policy preferences on electoral outcomes, improving the monitoring value of elections.

Through the same channel, the screening value of elections also increases. Since voters are more concerned about the ability of a national than a local politician, in equilibrium they select a central government whose average ability ($\mathbb{E}\hat{\eta}_p$) is greater. Thanks to centralization, not only wasteful rents decline, but the efficiency of productive public spending simultaneously rises.

As in Oates (1972), decentralization can be beneficial to avoid the cost of policy uniformity when regions have heterogeneous preferences. However, Proposition 2 establishes new forces that tend to make centralization more efficient than decentralization. Thus, the Decentralization Theorem holds in our model only if voters are homogeneously informed ($\theta_l = \theta$ for all l). With heterogeneous information, centralization is systematically welfare increasing.

Proposition 3 *Suppose that information is homogeneous ($\theta_l = \theta$ for all l).*

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 $l \neq m$). 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. Proposition 3 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 model, 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 their own ability to their own constituents, which they do most effectively by ignoring all externalities. Thus, Proposition 3 reflects an endogenous inability to internalize externalities under decentralization. In addition, Proposition 2 showed that the incentives and screening both improve under centralization when there are externalities in public good provision.

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.

Proposition 3 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 heterogeneous across regions. 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.

2.4 The Distribution of Efficiency Gains

Proposition 2 highlighted the mechanisms through which centralization increases average political accountability and therefore aggregate efficiency. Accordingly, Proposition 3 established that centralization is welfare-increasing from the national perspective when preferences are homogeneous. The benefits of centralization, however, are unevenly distributed across regions. The aggregate improvement in political accountability established by 2 is also a transfer from the more to the less informed.

Proposition 4 *Suppose that preferences are homogeneous and information heterogeneous across regions ($\alpha_p^l = \alpha_p$ and $\xi_p^l = \xi_p$ for all l and p ; $\theta_l \neq \theta_m$ for some $l \neq m$). If residents of region l are less informed than residents of region m , then centralization yields a greater increase in the expected provision of all public goods and thus in welfare in region l than in region m ($\theta_l < \theta_m$ implies $\mathbb{E}(g_{l,p}^C - g_{l,p}^D) > \mathbb{E}(g_{m,p}^C - g_{m,p}^D)$ for all p and $\mathbb{E}(u_l^C - u_l^D) > \mathbb{E}(u_m^C - u_m^D)$).*

If furthermore there are no externalities in public good provision ($\xi_p = 0$ for all p) then region l is better off under centralization than decentralization if and only if its voters are less informed than average ($\theta_l \leq \bar{\theta} \Leftrightarrow \mathbb{E}u_l^C \geq \mathbb{E}u_l^D$).

When voter 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. 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 overall impact of centralization on better informed regions is ambiguous. On the one hand, the quality of their electoral process is diluted when the constituency is expanded to include regions with less informed voters. On the other hand, in the presence of externalities ($\xi_p > 0$) all voters are more likely to act on the basis of whatever information they have in a national election. While the dilution of informed voters blunts the monitoring and screening ability of the electorate, externalities from central policy decisions sharpen it. The net effect is negative in the absence of spillovers, but it can be positive when they are present.

Conversely, regions with below-average information always gains from centralization. The implicit transfer of accountability effected by centralization effectively entails a welfare-increasing transfer of public funds, 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,

the contributions (*b*) are also identical, but less informed regions suffer much higher rent extraction and enjoy considerably lower productive government spending.

In the United States, the data confirm that political corruption varies greatly across 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.³ Proposition 4 then implies that expansion of the federal government should benefit disproportionately the states with lower political accountability, and more precisely lower voter information. The next section tests this prediction with evidence from federal environmental policy.

Suggestive empirical support for our theoretical prediction is also provided by European evidence. The European Union encompasses large disparities in the quality of government across regions and member states (Charron, Dijkstra, and Lapuente 2011). Consistent with our model, Fredriksson and Gaston (2000) conclude that an EU directive introducing uniform standards for packaging waste “was less stringent than the existing German, Danish and Dutch laws, but was significantly stricter than the Greek, Irish and Portuguese requirements.” Italy provides a striking example of large regional disparities in information and accountability (Putnam 1993; Del Monte and Papagni 2001, 2007; Golden and Picci 2005). Durante, Labartino, and Perotti’s (2011) empirical findings on decentralization in the Italian public university system correspond perfectly to our theory. 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.

3 Evidence from the Clean Air Act

To test the fundamental empirical prediction of our model we focus on a clear discontinuity in U.S. environmental policy. Up to the 1960s, air pollution had been primarily regulated at the state and local government. The year 1970 marked a dramatic centralizing intervention by the federal government. Federal involvement rested on two pillars: the establishment of the Environmental Protection Agency (EPA), and the passage of the Clean Air Act of 1970 and subsequent amendments, which phased in national air quality standards for a set of criterion pollutants. National standardization stood in sharp contrast with the previous state-based regulations, which had been adopted only by a few states, imposing very heterogeneous standards (U.S. Senate, 1970). We use the Clean Air Act, and the sharp regulatory shift it

³The average is 2.8 per million and the standard deviation 1.3.

entailed, to test the distributional predictions of our model, set out in Proposition 4.

The core of our analysis considers sulfur dioxide emissions. Sulfur dioxide is a very significant and potentially harmful pollutant. In the United States in 1970, among anthropogenic sources, SO_2 was predominantly released by electric utilities, which accounted for almost 56% of the total emission (90% of which from coal-fueled power plants, and 10% from oil-fueled ones), followed by metals processing (15%) and industrial fuel combustion (14%), released in a variety of industrial processes involving the combustion of high-sulfur fuels. Besides contributing to acid rain along with nitrous oxides, SO_2 at relatively high levels entails significantly adverse health effects, as it harms pulmonary and cardiovascular functions, and, in extreme cases, may lead to premature death. Furthermore, even at much lower levels, it severely damages crops.⁴

Sulfur dioxide was targeted in 1971, immediately after the passage of the Clean Air Act. After the standards were imposed, sulfur dioxide emissions declined drastically. However, the existence of a causal link between the legislation and the downward trend in emissions is still debated (Greenstone 2004). We do not mean to contribute to this debate by estimating the aggregate effect of the policy shift. Our empirical investigation is restricted to the study of the differential impact of the Clean Air Act across states. In particular, we investigate the effect of citizens' information on the differential trends in emission mitigation across states before and after the 1970 reform.

3.1 Data and Empirical Specification

Our main outcome variable is sulfur dioxide (SO_2) emissions for the 48 continental U.S. states; we also consider nitrogen oxides (NO_x) emissions as a robustness check. We use the same data as in Bulte, List, Strazicich (2007). They are part of a series covering the period 1929-1999 of state-level sulfur dioxides and nitrogen oxides emissions in short tons per capita. Bulte, List and Strazicich drew them from National Air Pollutant Emission Trends (U.S. EPA, 1994).

We measure citizens' information with newspapers circulation per capita. We constructed the panel data from annual editions of the *Statistical Abstract of the United States*, which in turn reports data from the *Editor & Publisher International Yearbook*. The figure is available for all states annually, with the exception of 1974. In addition to availability, newspaper circulation is a particularly apt measure of voters' ability to acquire information about government policy. Up to the 1980s, newspapers were Americans' main source of

⁴Sulfur dioxide emissions have been extensively studied, both with reference to the United States (List and Gallet 1999; List and Gerking 2000; Greenstone 2004; List and Sturm 2006; Bulte, List, and Strazicich 2007), and in cross-country analyses (Grossman and Krueger, 1995).

political news. Moreover, newspaper reader are acknowledged to be better informed and more involved in politics than consumers of other media (Graber 1984; Putnam 1993, 2000; Gentzkow 2006; Gentzkow, Shapiro, and Sinkinson 2011).

Our control variables include population, GDP, and value added from the manufacturing sector by state. All of these are from the Bureau of Economic Analysis Regional Accounts.⁵ The state GDP series is only available since 1963, dictating the starting point of our sample. We consider data up to 1980, covering the first decade since the enactment of Clean Air Act as well as most of the prior decade, to estimate the changes in patterns induced by the reform.

Our basic analysis is a difference-in-differences estimate of the effect of information on the reduction in emissions following the imposition of national air quality standards. Our outcome measure for the pollutant's concentration is the emission intensity of state GDP, and our main regression specification is the following:

$$\frac{(SO_2)_{i,t}}{GDP_{i,t}} = \alpha + \eta I_i \times P_t + \theta I_i \times S_t + X'_{i,t} \beta \times P_t + X'_{i,t} \lambda \times S_t + \gamma_t + \delta_i + \zeta_i \times t + \epsilon_{i,t}. \quad (24)$$

We include year fixed effect γ_t , state fixed effects δ_i , and a state-specific linear time trend $\zeta_i \times t$.

The main regressor of interest is information I_i . We define information at the state level, choosing 1970 as the reference year. As a robustness test, we also consider the time average over the entire sample period. We use two different but related measures of information. One is simply newspaper circulation per capita. The second is a binary variable, which equals 1 for states whose circulation is larger than the mean across states, and 0 for those below the mean.

To study the differential trend in emissions after the Clean Air Act, we consider two interactions of I_i . First, with a dummy P_t , which takes up the value 1 if $t > 1970$ and 0 otherwise. This would capture a level break in the series upon the introduction of federal emission standards. Second, with a linear time trend S_t , equal to $t - 1970$ for $t > 1970$ and 0 otherwise. This would capture a break in the trend of emissions after the reform.

Such a gradual impact is to be expected because the regulatory transition was itself gradual and required several years. States retained a role in the implementation and in the enforcement of the national standards, and were allowed until 1975 to comply (U.S. EPA, 1995). Additional evidence of gradualism is provided by the number of operating monitors reading the concentration of air pollutants, a key factor in the enforcement process, which increased gradually throughout the 1970s (Greenstone, 2004).

⁵We compute population density using state land area from the 2000 U.S. Census.

We introduce controls X by treating them exactly like our main independent variable. We begin by allowing a differential impact of the Clean Air Act on the basis not of information but by income. Thus, we introduce GDP per capita as a control. As with newspaper circulation, we consider the GDP of the state in 1970, or a binary variable that equals 1 if state GDP is above the average across states. This measure is then interacted with P_t and S_t , while the baseline is absorbed by the state fixed effect. In our full set of control, we also introduce in the same manner manufacturing value added per capita and population density.

3.2 Results

Before turning to our regression analysis, we can starkly visualize the main result in Figure 1. The graph plots average sulfur emissions for two group of states: those with above-average newspaper circulation in 1970, and those with below-average newspaper circulation. The difference-in-differences emerges clearly: uninformed states have considerably higher average emissions before 1970, and start decreasing them faster than the informed states as soon as national emission standards are introduced by the federal government. The convergence is gradual, with a trend instead of level break, but dramatic.

Table 1 confirms this result in our full regression specification including all controls. Columns (1), (3), and (5) report the results when I_i is measured as a binary variable, as in Figure 1; while in columns (2), (4), and (6) I_i is measured as a continuous variable.

The coefficient on newspapers circulation per capita is statistically significant with the expected sign for the trend effect throughout the six specifications. Indeed, the time trend appears to be the most obvious fit to our natural experiment. As previously discussed, the history of the Clean Air Act suggests that air quality standards were phased in gradually. The binary information measure, in any case, displays a significant level break in addition to its significant trend break.

The inclusion of controls does not qualitatively affect the results. In fact, the estimated coefficients and standard errors show very small quantitative changes, strengthening the significance of the empirical support for the theoretically predicted role of information.

As a robustness test, as anticipated, we repeat the same analysis with a slightly different measure of information. The continuous information variable is now, for each state, the average value of circulation in the 1963-1980 period. The binary variable takes the value 1 if the average value of circulation for the state in the 1963-1980 period exceeds the mean of the average values for all the states in the same period. Figure 2 and Table 2 show the results. As in Table 1, the trend effect for newspaper circulation is significant in all six specification, and the coefficients are only slightly reduced by the inclusion of control variables.

GDP is significant, although now as a trend effect and only under the binary measure of information; its sign is opposite with respect to the baseline specification. Combining the twelve specifications, we observe a robust effect of the information variable (in trends), as opposed to a weaker and less persistent effect of GDP. Notwithstanding concerns of correlation between newspaper circulation and GDP, the inclusion of the GDP in the estimation affects only marginally our estimator of the differential effect of information.

Overall, the results provide a clear empirical support for the notion, predicted by our theory, that the Clean Air Act had a differential impact across states, and that the difference depends on the level of information. Our results highlight that centralization improved the quality of air in states less informed than the average.

An extensive literature has assessed the benefits accruing to the population as a result of the CAA-induced emissions mitigation. Chay and Greenstone (2005) estimate the effect of the improvement in the quality of air on housing prices. By using data on total suspended particulate (TSP) air pollution, they find that better quality of air causes a substantial increase in house prices. Their welfare calculation shows that the mid-1970s TSPs regulation provided a \$45 billion aggregate increase in house values.⁶ Chay and Greenstone (2003) analyze the relation between air pollution and infant mortality, and find a significant positive association. They estimate that a reduction in TSPs by $1 \mu\text{g}/\text{m}^3$ is associated to approximately 200 additional infants surviving to one year of age.

As a further check of the robustness of our results, we repeat the entire analysis using a time series of emissions for nitrogen oxides NO_x . Sources of nitrogen oxides differ remarkably from those of sulfur dioxide. While the latter is released almost exclusively by stationary sources (in 1970, the transportation sector accounted for less than 2% of total emissions), the single main source for the nitrogen oxides is represented by on-road vehicles, which accounted for 35% of total emissions in the United States in 1970, as compared to 23% for electric utilities and 20% for industrial fuel combustion. The adverse health effects of nitrogen oxides primarily involve respiratory and cardiovascular problems, similarly to sulfur dioxides.

National air quality standards for NO_2 were imposed in 1971. However, specific standard for NO_x expressly targeting automobile emissions (consisting in national vehicle emission limits on NO_x emissions) were phased in starting in 1973 (U.S. Environmental Protection Agency, 2000), and applied to an increasingly large portion of the United States vehicle

⁶As their measure of regulation, they use EPA's split of counties into attainment and non-attainment. While the actual criteria used by EPA for the non-attainment designation are not entirely clear (Greenstone, 2004), counties whose emissions exceed the national standards are more likely to be classified as non-attaining. In the context of the CAAs, the non-attainment designation triggers a stricter regulation.

fleet over the years, as old cars were scrapped and replaced by new ones.⁷ Changes in the nitrogen oxides emissions materialized, after 1970, in an even more gradual way than for sulfur dioxide. There are, therefore, reasons to expect a trend effect to be visible after 1970 rather than a level effect.

Figures 3 and 4 show the trend in NO_x emissions intensity of GDP after splitting the states into two groups according to whether their newspaper circulation was above or below average in 1970. The figures suggest that a discontinuity, in this case, might be associated with the 1972 vehicle emission limits rather than with the Clean Air Act itself. Despite this potential concern, Tables 3 and 4 yield results that are remarkably similar to those in Tables 1 and 2 respectively, replacing the sulfur dioxide series with the one for nitrogen oxides. The coefficients on the newspaper circulation trend interactions are statistically significant at the 1% level for all 12 specifications. Again, GDP has a less clear effect.

3.3 Discussion

A legitimate concern with environmental regulation is that clean air may come at the cost of a deterioration in local economic conditions, as plants choose to relocate away from tightly regulated areas. Previous research has shown that air quality regulation indeed affects industrial location and is associated with reductions in employment, investment and shipments at the local level (Henderson 1996; Becker and Henderson 2002). In addition, tight standards determine a significant reduction in total factor productivity for polluting plants (Greenstone, List and Syverson, 2011). We might then suspect that the differential emission reduction following the Clean Air Act could be mirrored by a simultaneous differential negative impact on economic activity in the uninformed states, which recorded the largest emission reductions.

However, our analysis finds no evidence of such a differential effect. Figure 5 shows the series of GDP per capita for the two groups of states, more and less informed than the average across states in 1970. Unlike emissions, GDP per capita does not display any clear pattern of differences-in-differences between informed and uninformed states before and after 1970. A very similar argument can be made for the share of manufacturing in state GDP. Figure 6 shows its evolution for the two groups of states. Federal intervention in environmental regulation does not seem to be associated with any differential impact on the manufacturing share by information.

The pattern of the data then suggest that the differential impact of the Clean Air Act can plausibly be interpreted according to our theoretical model. Before 1970, under state and local governments rule, uninformed states were facing a problem of bad regulation. They

⁷Standards started to be applied to cars produced in the 1973 model year.

were subject to excessively loose standards (or no standards whatsoever), to the point that, when the federal government stepped in and prescribed tighter regulatory requirements, it was able to do so at a very limited (and possibly nil) differential cost for the states involved. Federal standards may have been costly for the entire nation, but they do not appear to have been costlier for the uninformed states that reaped the greatest declines in emissions.

Suggestive evidence of a problem of bad regulation by the uninformed states can also be gleaned from direct measures of local government activity. We can look at two different measures of regulatory inputs: expenditures by state and local governments for air quality control, and actual regulations implemented by states and local governments before 1970.

Consistent with the public good nature of air pollution, Henderson (1996) finds that a larger expenditure in the state on abatement activities is associated with better outcomes in terms of air quality.⁸ We obtain direct state expenditure from the U.S. Department of Commerce yearly report: “Environmental Quality Control” (U.S. Bureau of the Census 1971, 1980). The report publishes the yearly expenditure for air quality control for states, counties, and cities. Since the first published report is for the fiscal year 1969, this source allows us to compare a period before the 1970 Clean Air Act to one after: we consider a decade and compare with data for the fiscal year 1978. By combining state, county and city data, we construct a measure total expenditure for air quality control regulation for each year and state. Splitting the sample into two groups of states, based on average newspaper circulation in 1970, we find that in 1969 spending relative to GDP in uninformed states was on average 71% as much as in informed states. In 1978, spending in uninformed states rose to 86% of spending in informed states. After the introduction of uniform standards, uninformed states closed half the gap with informed states.⁹ The 1970 policy shift may have fostered a convergence, in percentage terms, of state and local government expenditures in air pollution regulation. In line with our theoretical model and with our main empirical result, this suggests a differential effect of the reform on the expenditure levels, based on the level of citizens’ information.

As a second measure of regulatory input, we use the total number of standards implemented at the state level before 1970. Our source are the hearings of the subcommittee on air and water pollution of the United States senate (U.S. Senate, 1970). The document reports the standards adopted by states and local governments before 1970 on ten pollutants.¹⁰ We

⁸He uses the Pollution Abatement Control Expenditures (PACE) series, United States Department of Commerce, Bureau of the Census. The series collects state-level data on firms expenditures on abatement activity. It is available starting in 1973.

⁹The raw correlation between information (measured as a binary variable) and expenditure/GDP declined from .26 (and 10% significance level) in 1969 to 0.10 (and non significant, with a p-value of .4963) in 1978.

¹⁰The ten pollutants are carbon monoxides, beryllium, fluorides, hydrogen sulfide, lead, nitrogen oxides, sulfur dioxides, sulfuric acid, suspended particulates and total oxidants.

count the number of pollutants that each state had regulated prior to 1970, and we then divide the states into two groups according to the usual criterion based on newspaper circulation in 1970. We find that informed states had, on average, adopted four standards, while uninformed states had, on average, adopted three standards only.

The document also reports the number of states that had proposed or adopted emission standards for sulfur dioxide and for total suspended particulate as a result of the provision of the 1967 Air Quality Act.¹¹ We again divide the states into informed and uninformed based on the average level of newspaper circulation in 1970, and find that 50% of the informed states had proposed or adopted emission standards for the two pollutants, as opposed to only 25% among the uninformed. This suggests that more informed states had more standards before the federal intervention. The uniform national standards imposed after the Clean Air Act would *ipso facto* trigger a regulatory convergence, entailing a greater improvement in the standards of the less informed and less regulated states.

4 Centralization without Uniformity

Section 2 presented a model of political and fiscal centralization in which public goods are uniformly provided across regions by the central government, and differentiation only occurs under decentralized local governments. The assumption of a uniform national policy is natural vis à vis the empirical analysis of Section 3, since the very purpose of the 1970 Clean Air Act was to introduce uniform national standards for regulating pollution and its abatement. Moreover, the uniformity assumption has been standard in the literature on fiscal federalism since Oates (1972). Nonetheless, imposing a uniformity constraint on centralized public-good provision is not necessarily realistic in all settings. E.g., discretionary federal spending is not required to be homogeneously spent across states, almost by definition of discretionality. Thus, Lockwood (2002) and Besley and Coate (2003) have modelled fiscal federalism under the alternative hypothesis that the central government can arbitrarily vary the provision of public goods across regions.

A straightforward extension of our theoretical framework includes both public goods whose centralized provision is subject to a uniformity constraint, as in Section 2, and others that the central government can instead provide in different amounts to different regions. Letting the former constitute the set \mathcal{U} and the latter the set \mathcal{D} , the central government

¹¹The 1967 Air Quality Act required that states establish air quality control regions and that the Department of Health, Education, and Welfare promulgate criteria to serve as the basis for setting emission standards. States would then use the HEW information to set air quality standards. Under the Air Quality Act, states retained autonomy in their decision of setting the criteria.

budget constraint becomes

$$b^C = r_t^C + \sum_{p \in \mathcal{U}} x_{l,p,t}^C + \sum_{l=1}^L \sum_{p \in \mathcal{D}} x_{l,p,t}^C, \quad (25)$$

while each region then receives public goods

$$g_{l,p,t}^C = \frac{1}{L} e^{\eta_{p,t}} x_{p,t}^C \text{ for } p \in \mathcal{U} \text{ and } g_{l,p,t}^C = e^{\eta_{p,t}} x_{l,p,t}^C \text{ for } p \in \mathcal{D}. \quad (26)$$

The model then features two opposite distributive patterns. As established by Proposition 4, uniformly provided public goods ($p \in \mathcal{U}$) imply a benefit of centralization for less informed regions. The empirical evidence in Section 3 bears out this prediction, showing that national air quality standards adopted by the EPA benefited disproportionately states with lower information. Conversely, however, unconstrained discretionary spending tends to favor more politically influential group. In our model, political influence stems from information, since more knowledgeable voters provide more of the politicians' incentives. As a consequence, better informed regions benefit disproportionately from non-uniform public goods ($p \in \mathcal{D}$), which are equally financed by all taxpayers. This pattern is consistent with evidence on the regional allocation of discretionary government spending during the New Deal (Strömberg 2004). Responding to voters' information, state governors directed more public funds to counties with a greater share of radio listeners.

The balance between the two countervailing distributional forces depends on the relative importance of the two types of public goods, which is summarized by the welfare weight of public goods that the central government must provide uniformly:

$$\alpha_{\mathcal{U}} \equiv \sum_{p \in \mathcal{U}} \alpha_p \in [0, 1]. \quad (27)$$

Striking the appropriate balance emerges as a crucial requirement of centralization. Its absence is perilous: the welfare gains described by Proposition 3 are no longer assured without it. Its presence is beneficial: by modulating the distribution of the accountability gains between informed and uninformed reason, it can make centralization Pareto efficient even in the absence of externalities, despite Proposition 4.

Proposition 5 *Suppose that preferences are homogeneous, information is heterogeneous, and there are no externalities ($\alpha_p^l = \alpha_p$ and $\xi_p^l = 0$ for all l and p ; $\theta_l \neq \theta_m$ for some $l \neq m$).*

1. *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 $\alpha_{\mathcal{U}} \geq \bar{\alpha}_{\mathcal{U}}$.*

2. *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 first result highlights that a uniformity constraint, which creates costs of centralization in the classic theory of fiscal federalism (Oates 1972), is instead 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. The uniformity constraint is not binding for a benevolent central social planner when preferences are homogeneous

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 ($1 - \alpha_{\mathcal{U}}$) measures the share of the central budget that better informed regions can appropriate. In the limit as $\alpha_{\mathcal{U}} \rightarrow 0$, 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 a Pareto inefficient protectionist bias of trade policy.

Proposition 5 highlights 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 $1 - \alpha_{\mathcal{U}} \geq \rho^C$, 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 ($1 - \alpha_{\mathcal{U}}$) is worth more than a decrease in rent extraction to $\rho_l^D < \rho^C$. For $\rho^C \geq 1 - \alpha_{\mathcal{U}}$, 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 ρ over distributive goods ($1 - \alpha_{\mathcal{U}}$). Hence, if $1 - \alpha_{\mathcal{U}} = \rho^C$ 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 σ^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.

5 Conclusions

Political accountability and the quality of government vary across regions within a country like the United States, and across member states of international organizations like the European Union. In this paper, we have shown that such regional differences imply that centralization increases political accountability.

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. We have shown that electoral discipline has decreasing returns. Therefore, a central politician answerable to the whole national electorate extracts lower rents than a collection of local politicians, some monitored tightly by well-informed voters and some loosely by poorly-informed constituents. Hence, we have found that centralization is beneficial whenever voter information is heterogeneous across regions. This result can help to explain the steady growth of the federal government over the history of the United States, and the sharp increase in the scope and extent of the powers of the European Union since the 1970s (Alesina, Angeloni, and Schuknecht 2005).

Our model also predicts distributional consequences of centralization when regions have different levels of information, even without differences in income. When the central government provides public goods uniformly across the nation, the benefits of centralization are monotone decreasing in voter information. We have tested this prediction by analyzing the differential impact across U.S. states of the Clean Air Act of 1970. Beforehand, environmental regulation was largely in the hands of states and local governments. Beginning in 1970, the federal government took charge and started introducing uniform national standards. We have found significant and robust evidence of differences-in-differences. Consistent with our theoretical model, centralization of environmental policy induced a differential decline in pollution in less informed states, relative to better informed ones.

Our finding that centralization benefits the least informed regions hinges on our focus on uniform policies, both in our baseline model and in empirical evidence from pollutant emission standards. In our theoretical framework, we have shown conversely that if the central politician can differentiate local public good provision across regions, he targets the most informed. While uniform policy entails a transfer of accountability from the informed to the

uninformed, discretionary spending reflects a transfer of power from the uninformed to the informed. Thus we have identified the balancing role of a uniformity requirement for central-government policies. Some uniformity is necessary for centralization to be welfare increasing. A carefully calibrated constraint can even ensure the Pareto efficiency of centralization.

A Derivations and Proofs

A.1. Proof of Proposition 1

The budget allocation and the expectation of the incumbent's ability are derived in the body of the text. The cumulative distribution function of $\hat{\eta}_{p,t}$ is

$$\Psi_t \leq \sum_{j=1}^J \lambda_j \theta_j \sum_{p=1}^P \alpha_p^j \varepsilon_{p,t}. \quad (\text{A1})$$

$$\begin{aligned} \Pr(\hat{\eta}_{p,t} \leq \eta) &= \Pr[\chi_{t-1} (\varepsilon_{p,t-1}^I + \varepsilon_{p,t}^I) + (1 - \chi_{t-1}) (\varepsilon_{p,t-1}^C + \varepsilon_{p,t}^C) \leq \eta] \\ &= \Pr(\chi_{t-1} = 1 \wedge \varepsilon_{p,t-1}^I + \varepsilon_{p,t}^I \leq \eta) + \Pr(\chi_{t-1} = 0 \wedge \varepsilon_{p,t-1}^C + \varepsilon_{p,t}^C \leq \eta) \\ &= \Pr\left(\Psi_{t-1} \leq \sum_{j=1}^J \lambda_j \theta_j \sum_{p=1}^P \alpha_p^j \varepsilon_{p,t-1} \wedge \varepsilon_{p,t-1}^I + \varepsilon_{p,t}^I \leq \eta\right) \\ &\quad + \frac{1}{2} \Pr(\varepsilon_{p,t-1}^C + \varepsilon_{p,t}^C \leq \eta) \\ &= \int_{-\infty}^{\infty} \left(1 + \varepsilon \phi \sum_{j=1}^J \lambda_j \theta_j \alpha_p^j\right) F_\varepsilon(\eta - \varepsilon) f_\varepsilon(\varepsilon) d\varepsilon, \end{aligned} \quad (\text{A2})$$

where $F_\varepsilon(\varepsilon)$ is the cumulative distribution function of $\varepsilon_{p,t}$ and $f_\varepsilon(\varepsilon)$ its probability density function. Since

$$\int_{-\infty}^{\infty} \varepsilon F_\varepsilon(\eta - \varepsilon) f_\varepsilon(\varepsilon) d\varepsilon = \mathbb{E}[\varepsilon F_\varepsilon(\eta - \varepsilon)] < \mathbb{E}\varepsilon \mathbb{E}[F_\varepsilon(\eta - \varepsilon)] = 0, \quad (\text{A3})$$

an increase in $\phi \sum_{j=1}^J \lambda_j \theta_j \alpha_p^j$ induces an increase in $\hat{\eta}_p$ in the sense of first-order stochastic dominance.

A.2. Centralization and Decentralization

Under centralization, Proposition 1 implies that rent extraction is

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\bar{\theta}\right)^{-1}, \quad (\text{A4})$$

and the expected ability of a central politician is

$$\mathbb{E}\hat{\eta}_p^C = \frac{\phi\sigma^2}{L} \sum_{l=1}^L \alpha_p^l \theta_l. \quad (\text{A5})$$

The relative shares of each local public good are

$$\beta_p^C = \frac{1}{L} \sum_{l=1}^L \alpha_p^l \frac{\theta_l}{\theta}. \quad (\text{A6})$$

Welfare in region l is

$$\mathbb{E}u_l^C = \log \frac{b^C}{L} + \log (1 - \rho^C) + \sum_{p=1}^P \alpha_p^l (\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C), \quad (\text{A7})$$

and aggregate welfare is

$$W^C = \log \frac{b^C}{L} + \log (1 - \rho^C) + \sum_{p=1}^P \bar{\alpha}_p (\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C). \quad (\text{A8})$$

Under decentralization, Proposition 1 implies that rent extraction is

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi \left(1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p^l \alpha_p^l \right) \theta_l \right]^{-1}, \quad (\text{A9})$$

and the expected ability of a local politician is

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2 \left(1 - \frac{L-1}{L} \xi_p^l \right) \alpha_p^l \theta_l. \quad (\text{A10})$$

The relative shares of each local public good are

$$\beta_{l,p}^D = \frac{\left(1 - \frac{L-1}{L} \xi_p^l \right) \alpha_p^l}{1 - \frac{L-1}{L} \sum_{q=1}^P \xi_q^l \alpha_q^l}. \quad (\text{A11})$$

Welfare in region l is

$$\mathbb{E}u_l^D = \sum_{p=1}^P \alpha_p^l \left\{ \begin{aligned} & \left(1 - \xi_p^l \right) \left[\log b_l^D + \log (1 - \rho_l^D) + \mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D \right] \\ & + \frac{\xi_p^l}{L} \sum_{m=1}^L \left[\log b_m^D + \log (1 - \rho_m^D) + \mathbb{E}\hat{\eta}_{m,p}^D + \log \beta_{m,p}^D \right] \end{aligned} \right\}, \quad (\text{A12})$$

and aggregate welfare is

$$W^D = \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \left\{ \left[\left(1 - \xi_p^l \right) \alpha_p^l + \frac{1}{L} \sum_{m=1}^L \xi_p^m \alpha_p^m \right] \left[\log b_l^D + \log (1 - \rho_l^D) + \mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D \right] \right\}. \quad (\text{A13})$$

A.3. Proof of Proposition 2

A.3.1 Rent Extraction

Aggregate rent extraction is lower under centralization if and only if

$$\rho^C \leq \frac{1}{L} \sum_{l=1}^L \rho_l^D, \quad (\text{A14})$$

which can be written

$$f\left(\frac{1}{L} \sum_{l=1}^L \theta_l\right) \leq \frac{1}{L} \sum_{l=1}^L f\left(\left(1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p^l \alpha_p^l\right) \theta_l\right) \quad (\text{A15})$$

for

$$f(x) \equiv \frac{1}{1 + 2\tilde{\delta}\phi x} \quad (\text{A16})$$

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

$$f'(x) = -\frac{2\tilde{\delta}\phi}{(1 + 2\tilde{\delta}\phi x)^2} < 0 \text{ and } f''(x) = \frac{(2\tilde{\delta}\phi)^2}{(1 + 2\tilde{\delta}\phi x)^3} > 0. \quad (\text{A17})$$

Thus

$$f\left(\frac{1}{L} \sum_{l=1}^L \theta_l\right) \leq \frac{1}{L} \sum_{l=1}^L f(\theta_l) \leq \frac{1}{L} \sum_{l=1}^L f\left(\left(1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p^l \alpha_p^l\right) \theta_l\right). \quad (\text{A18})$$

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

A.3.2 Ability

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

$$\mathbb{E}\hat{\eta}_p^C \geq \frac{1}{L} \sum_{l=1}^L \mathbb{E}\hat{\eta}_{l,p}^D, \quad (\text{A19})$$

which can be written

$$\sum_{l=1}^L \alpha_p^l \theta_l \geq \sum_{l=1}^L \left(1 - \frac{L-1}{L} \xi_p^l\right) \alpha_p^l \theta_l. \quad (\text{A20})$$

The right-hand side is weakly lower than the left-hand side, and strictly lower if $\xi_p^l > 0$ for some l and p .

A.4. Proof of Proposition 3

A.4.1 The Decentralization Theorem

Suppose that information is homogeneous ($\theta_l = \theta$ for all l). Then under centralization

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\theta\frac{1}{L}\sum_{l=1}^L\alpha_p^l, \text{ and } \beta_p^C = \frac{1}{L}\sum_{l=1}^L\alpha_p^l. \quad (\text{A21})$$

while under decentralization

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\theta\left(1 - \frac{L-1}{L}\sum_{p=1}^P\xi_p^l\alpha_p^l\right)\right]^{-1}, \quad (\text{A22})$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\left(1 - \frac{L-1}{L}\xi_p^l\right)\alpha_p^l, \quad (\text{A23})$$

and

$$\beta_{l,p}^D = \left(1 - \frac{L-1}{L}\xi_p^l\right)\alpha_p^l. \quad (\text{A24})$$

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

$$\rho^C = \rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \mathbb{E}\hat{\eta}_p^C = \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\alpha_p \text{ and } \beta_p^C = \beta_{l,p}^D = \alpha_p. \quad (\text{A25})$$

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

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\theta\alpha_p, \text{ and } \beta_p^C = \alpha_p, \quad (\text{A26})$$

and

$$W^C = \log b + \log(1 - \rho^C) + \sum_{p=1}^P\alpha_p(\mathbb{E}\hat{\eta}_p^C + \log\beta_p^C). \quad (\text{A27})$$

Under decentralization

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\theta\left(1 - \frac{L-1}{L}\sum_{p=1}^P\xi_p\alpha_p\right)\right]^{-1} > \rho^C, \quad (\text{A28})$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p < \mathbb{E}\hat{\eta}_p^C, \quad (\text{A29})$$

$$\beta_{l,p}^D = \frac{\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p}{1 - \frac{L-1}{L}\sum_{q=1}^P\xi_q\alpha_q}, \quad (\text{A30})$$

and

$$W^D = \log b + \log(1 - \rho_l^D) + \sum_{p=1}^P \alpha_p (\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D). \quad (\text{A31})$$

Centralization achieves the optimal allocation of productive expenditures, while decentralization does not unless ξ_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 ξ_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

$$W^C = \log b + \log(1 - \rho^C) + \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l (\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C), \quad (\text{A32})$$

while under decentralization

$$\rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1} = \rho^C, \quad \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\alpha_p^l, \quad \beta_{l,p}^D = \alpha_p^l \quad (\text{A33})$$

and

$$W^D = \log b + \log(1 - \rho_l^D) + \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l (\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D). \quad (\text{A34})$$

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

$$\frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l \mathbb{E}\hat{\eta}_{l,p}^D > \frac{1}{L} \sum_{l=1}^L \sum_{p=1}^P \alpha_p^l \mathbb{E}\hat{\eta}_p^C \quad (\text{A35})$$

because for each public good p

$$\frac{1}{L} \sum_{l=1}^L (\alpha_p^l)^2 > \left(\frac{1}{L} \sum_{l=1}^L \alpha_p^l\right)^2 \quad (\text{A36})$$

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

A.4..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

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\bar{\theta}\right)^{-1}, \quad \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\alpha_p\bar{\theta}, \quad \beta_p^C = \alpha_p, \quad (\text{A37})$$

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

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

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

$$f\left(\frac{1}{L} \sum_{l=1}^L \theta_l\right) \geq \frac{1}{L} \sum_{l=1}^L f(\theta_l) \geq \frac{1}{L} \sum_{l=1}^L f\left(\left(1 - \frac{L-1}{L} \sum_{p=1}^P \xi_p \alpha_p\right) \theta_l\right). \quad (\text{A47})$$

The first inequality follows from convexity by Jensen's inequality and holds strictly if θ_l is heterogeneous across regions. The second 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 ξ_p is homogeneous across public goods.

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

A.5. Proof of Proposition 4

Suppose that preferences are homogeneous ($\alpha_p^l = \alpha_p$ and $\xi_p^l = \xi_p$ for all l). Then public-good provision in region l is

$$g_{l,p,t}^D = x_{p,t}^D \exp(\hat{\eta}_{l,p}^D) = \beta_{l,p}^D b^D (1 - \rho_l^D) \exp(\hat{\eta}_{l,p}^D) \quad (\text{A48})$$

under decentralization, and

$$g_{l,p,t}^C = \frac{1}{L} x_{p,t}^C \exp(\hat{\eta}_p^C) = \beta_p^C b^D (1 - \rho^C) \exp(\hat{\eta}_p^C) \quad (\text{A49})$$

under centralization.

The proof of Proposition 3 has established that $\partial \beta_{l,p}^D / \partial \theta_l = 0$ and $\partial \rho_l^D / \partial \theta_l < 0$. Recalling the proof of Proposition 1, $\hat{\eta}_{l,p}^D$ is increasing in θ_l in the sense of first-order stochastic dominance, so $\partial \mathbb{E} \exp(\hat{\eta}_{l,p}^D) / \partial \theta_l > 0$. Since $g_{l,p,t}^C$ is identical for all regions, it follows that

$$\mathbb{E}(g_{l,p}^C - g_{l,p}^D) > \mathbb{E}(g_{m,p}^C - g_{m,p}^D) \Leftrightarrow \mathbb{E}g_{l,p}^D < \mathbb{E}g_{m,p}^D \Leftrightarrow \theta_l < \theta_m. \quad (\text{A50})$$

Similarly for residents' welfare

$$\begin{aligned} \mathbb{E}(u_l^C - u_l^D) > \mathbb{E}(u_m^C - u_m^D) &\Leftrightarrow \mathbb{E}u_l^D < \mathbb{E}u_m^D \Leftrightarrow \\ \sum_{p=1}^P \alpha_p (1 - \xi_p) [\log(1 - \rho_l^D) + \mathbb{E}\hat{\eta}_{l,p}^D] &< \sum_{p=1}^P \alpha_p (1 - \xi_p) [\log(1 - \rho_l^D) + \mathbb{E}\hat{\eta}_{l,p}^D] \Leftrightarrow \\ &\theta_l < \theta_m \end{aligned} \quad (\text{A51})$$

If furthermore there are no externalities ($\xi_p = 0$ for all p) then under decentralization

$$\rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta_l\right)^{-1}, \quad \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\alpha_p\theta_l, \quad \beta_{l,p}^D = \alpha_p, \quad (\text{A52})$$

and

$$\mathbb{E}u_l^D = \log b^D + \log(1 - \rho_l^D) + \sum_{p=1}^P \alpha_p^l (\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D). \quad (\text{A53})$$

If and only if $\theta_l \leq \bar{\theta}$ then $\rho_l^D \geq \rho^C$, $\mathbb{E}\hat{\eta}_{l,p}^D \leq \mathbb{E}\hat{\eta}_p^C$, $\mathbb{E}u_l^D \leq \mathbb{E}u_l^C$, $\mathbb{E} \exp(\hat{\eta}_{l,p}^D) \leq \mathbb{E} \exp(\hat{\eta}_{l,p}^C)$, and $\mathbb{E}g_{l,p}^D \leq \mathbb{E}g_{l,p}^C$, with joint strict inequalities.

A.6. Proof of Proposition 5

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

$$W^D = \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi \sigma^2 \bar{\theta} \sum_{p=1}^P \alpha_p^2 + \frac{1}{L} \sum_{l=1}^L \log \frac{2\tilde{\delta}\phi\theta_l}{1 + 2\tilde{\delta}\phi\theta_l} \quad (\text{A54})$$

From Proposition 1, rent extraction under centralization is

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\bar{\theta}\right)^{-1} \text{ for } \bar{\theta} = \frac{1}{L} \sum_{l=1}^L \theta_l \quad (\text{A55})$$

and the expected ability of a central politician is

$$\mathbb{E}\hat{\eta}_p^C = \phi \sigma^2 \bar{\theta} \alpha_p. \quad (\text{A56})$$

The relative shares of each local public good are

$$\beta_p^C = \alpha_p \text{ for } p \in \mathcal{U} \text{ and } \beta_{l,p}^C = \frac{1}{L} \frac{\theta_l}{\bar{\theta}} \alpha_p \text{ for } p \in \mathcal{D}. \quad (\text{A57})$$

Welfare in region l is

$$\begin{aligned} \mathbb{E}u_l^C &= \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi \sigma^2 \bar{\theta} \sum_{p=1}^P \alpha_p^2 \\ &\quad + \log \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} + (1 - \alpha_U) (\log \theta_l - \log \bar{\theta}), \end{aligned} \quad (\text{A58})$$

and aggregate welfare is

$$\begin{aligned} W^C &= \log b + \sum_{p=1}^P \alpha_p \log \alpha_p + \phi \sigma^2 \bar{\theta} \sum_{p=1}^P \alpha_p^2 \\ &\quad + \log \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} + (1 - \alpha_U) \left(\frac{1}{L} \sum_{l=1}^L \log \theta_l - \log \bar{\theta} \right). \end{aligned} \quad (\text{A59})$$

Then aggregate welfare is greater under centralization if

$$\alpha_U \geq \frac{\log \left(1 + 2\tilde{\delta}\phi\bar{\theta}\right) - \frac{1}{L} \sum_{l=1}^L \log \left(1 + 2\tilde{\delta}\phi\theta_l\right)}{\log \bar{\theta} - \frac{1}{L} \sum_{l=1}^L \log \theta_l} \equiv \bar{\alpha}_U \in (0, 1). \quad (\text{A60})$$

Region l gains from centralization if

$$\log \frac{1 + 2\tilde{\delta}\phi\theta_l}{1 + 2\tilde{\delta}\phi\bar{\theta}} - \alpha_{\mathcal{U}} \log \frac{\theta_l}{\bar{\theta}} \geq \phi\sigma^2 \sum_{p=1}^P \alpha_p^2 (\theta_l - \bar{\theta}). \quad (\text{A61})$$

The function

$$f(x) \equiv \log \left(1 + 2\tilde{\delta}\phi x \right) - \alpha_{\mathcal{U}} \log x \quad (\text{A62})$$

has a unique minimum

$$f'(x) = \frac{2\tilde{\delta}\phi}{1 + 2\tilde{\delta}\phi x} - \frac{\alpha_{\mathcal{U}}}{x} \geq 0 \Leftrightarrow x \geq \frac{\alpha_{\mathcal{U}}}{2\tilde{\delta}\phi(1 - \alpha_{\mathcal{U}})}, \quad (\text{A63})$$

and convexity

$$f''(x) = - \left(\frac{2\tilde{\delta}\phi}{1 + 2\tilde{\delta}\phi x} \right)^2 + \frac{\alpha_{\mathcal{U}}}{x^2} > 0 \Leftrightarrow \alpha_{\mathcal{U}} > \left(\frac{2\tilde{\delta}\phi x}{1 + 2\tilde{\delta}\phi x} \right)^2 \quad (\text{A64})$$

Hence if

$$\alpha_{\mathcal{U}} = \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} = 1 - \rho^C \quad (\text{A65})$$

the left-hand side of condition A61 is a convex function of $\theta_l \in [0, 1]$ with minimum at $\theta_l = \bar{\theta}$. This also proves that

$$\bar{\alpha}_{\mathcal{U}} < 1 - \rho^C. \quad (\text{A66})$$

Moreover, let

$$m = \arg \min_{l=1, \dots, L} \{ \theta_l : \theta_l > \bar{\theta} \}. \quad (\text{A67})$$

Then for

$$\sigma^2 \leq \frac{1}{\phi \sum_{p=1}^P \alpha_p^2 (\theta_m - \bar{\theta})} \left\{ \log \frac{1 + 2\tilde{\delta}\phi\theta_m}{1 + 2\tilde{\delta}\phi\bar{\theta}} - \frac{2\tilde{\delta}\phi\bar{\theta}}{1 + 2\tilde{\delta}\phi\bar{\theta}} \log \frac{\theta_m}{\bar{\theta}} \right\} \equiv \bar{\sigma}^2 > 0, \quad (\text{A68})$$

centralization Pareto dominates decentralization for $\alpha_{\mathcal{U}} = \rho^C$. 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.

References

- [1] Alesina, Alberto, Ignazio Angeloni, and Federico Etro. 2005. “International Unions.” *American Economic Review* 95(3):602–15.
- [2] Alesina, Alberto, Ignazio Angeloni, and Ludger Schuknecht. 2005. “What Does the European Union Do?” *Public Choice* 123:275–319.
- [3] Alesina, Alberto, and Enrico Spolaore. 1997. “On the Number and Size of Nations.” *Quarterly Journal of Economics* 112(4):1027–56.
- [4] Bardhan, Pranab, and Dilip Mookherjee. 2000. “Capture and Governance at Local and National Levels.” *American Economic Review* 90(2):135–9.
- [5] ———. 2006. “Decentralization, Corruption, and Government Accountability.” In *International Handbook on the Economics of Corruption*, edited by Susan Rose-Ackerman, 161–88. Cheltenham: Edward Elgar.
- [6] Becker, Randy and Vernon Henderson. 2000. “Effects of Air Quality Regulations on Polluting Industries.” *Journal of Political Economy* 108(2):379–421.
- [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, 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.
- [14] Chay, Kenneth, and Michael Greenstone. 2003. “The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession.” *Quarterly Journal of Economics* 118:1121–67.

- [15] ——. 2005. “Does Air Quality Matter: Evidence from the Housing Market.” *Journal of Political Economy* 116:376–424.
- [16] 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.
- [17] ——. 2007. “The Determinants of Corruption in Italy: Regional Panel Data Analysis.” *European Journal of Political Economy* 23:379–96.
- [18] Durante, Ruben, Giovanna Labartino, and Roberto Perotti. 2011. “Academic Dynasties: Decentralization and Familism in the Italian Academia.” NBER Working Paper No 17572.
- [19] 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.
- [20] 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.
- [21] ——. 2011. “Electoral Accountability and Corruption: Evidence from the Audit Reports of Local Governments.” *American Economic Review* 101:1274–311.
- [22] 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.
- [23] Gennaioli, Nicola, and Ilija Rainer. 2007a. “The Modern Impact of Precolonial Centralization in Africa.” *Journal of Economic Growth* 12(3):185–234.
- [24] ——. 2007b. “Precolonial Centralization and Institutional Quality in Africa.” In *Institutions and Norms in Economic Development*, edited by Mark Gradstein and Kai Konrad. Cambridge, MA: MIT Press.
- [25] Gentzkow, Matthew. 2006. “Television and voter turnout.” *Quarterly Journal of Economics* 121(3):931–72.
- [26] Gentzkow, Matthew and Jesse Shapiro. 2010. “What Drives Media Slant? Evidence from U.S. Daily Newspapers.” *Econometrica*, 78(1): 35–71.
- [27] Gentzkow, Matthew, Jesse Shapiro, and Michael Sinkinson. 2011. “The Effect of Newspaper Entry and Exit on Electoral Politics” *American Economic Review* 101:2980–3018.
- [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] Graber, Doris A. 1984. *Processing the News: How People Tame the Information Tide*. New York, NY: Longman.
- [31] Greenstone, Michael. 2002. “The impacts of environmental regulations on industrial activity: evidence from the 1970 and 1977 clean air act amendments and the census of manufacturers.” *Journal of Political Economy*. 110:1175–1219.
- [32] ——. 2004. “Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?” *Journal of Environmental Economics and Management*. 47:585-611.
- [33] Greenstone, Michael, John List, and Chad Syverson. 2011 “The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing”, mimeo.
- [34] Grossman, Gene, and Alan Krueger. 1995. “Economic growth and the environment.” *Quarterly Journal of Econom* 110:353-377.
- [35] Henderson, J. Vernon. 1996. “Effects of air quality regulation.” *American Economic Review* 86:789–813.
- [36] Lindbeck, Assar, and Jörgen W. Weibull. 1987. “Balanced-Budget Redistribution as the Outcome of Political Competition.” *Public Choice* 52:273–97.
- [37] List, John, and Craig Gallet. 1999. “The environmental Kuznets curve: does one size fit all?” *Ecological Economics* 31: 409–423.
- [38] List, John, and Shelby Gerking. 2000. “Regulatory Federalism and Environmental Protection in the United States.” *Journal of Regional Sciences*. 40:453-471.
- [39] List, John, and Daniel Sturm. 2006. “How Elections Matter: Theory and Evidence from Environmental Policy.” *Quarterly Journal of Economics*. 121:1249-1281.
- [40] Lockwood, Ben. 2002. “Distributive Politics and the Costs of Centralization.” *Review of Economic Studies* 69:313–37.
- [41] ——. 2006. “The Political Economy of Decentralization.” In *Handbook of Fiscal Federalism*, edited by Ehtisham Ahmad and Giorgio Brosio, 33–60. Cheltenham: Edward Elgar.
- [42] Nannicini, Tommaso, Andrea Stella, Guido Tabellini, and Ugo Troiano. 2010. “Social Capital and Political Accountability.” Mimeo, Bocconi University.
- [43] Oates, Wallace E. 1972. *Fiscal Federalism*. New York, NY: Harcourt Brace Jovanovich.
- [44] ——. 1999. “An Essay on Fiscal Federalism.” *Journal of Economic Literature* 37:1120–49.
- [45] Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.

- [46] Ponzetto, Giacomo A. M. 2011. “Heterogeneous Information and Trade Policy.” CEPR Discussion Paper No. 8726.
- [47] Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.
- [48] ——. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York, NY: Simon & Schuster.
- [49] Seabright, Paul. 1996. “Accountability and Decentralization in Government: An Incomplete Contracts Model.” *European Economic Review* 40:61–89.
- [50] Snyder, James M., and David Strömberg. 2010. “Press Coverage and Political Accountability.” *Journal of Political Economy* 118(2):335–408.
- [51] Strömberg, David. 2004. “Radio’s Impact on Public Spending.” *Quarterly Journal of Economics* 99:189–221.
- [52] 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.
- [53] U.S. Department of Commerce. 1971. “Environmental Quality Control.” *State and Local Governments Special Studies*. U.S. Bureau of the Census.
- [54] U.S. Department of Commerce. 1980. “Environmental Quality Control.” *State and Local Governments Special Studies*. U.S. Bureau of the Census.
- [55] U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. 1995. “National Air Pollutant Emission Trends, 1900–1994.” *EPA-454/R-95-011*. Research Triangle Park, NC.
- [56] U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. 2000. “National Air Pollutant Emission Trends, 1900–1998.” *EPA-454/R-95-011*. Research Triangle Park, NC.
- [57] U.S. Senate. 1970. “Hearings of the subcommittee on air and water pollution of the Committee on Public Works.” U.S. Government Printing Office.
- [58] Wallis, John J. 2000a. “American Government Finance in the Long Run: 1790 to 1990.” *Journal of Economic Perspectives*, 14(1):61-82.
- [59] ——. 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.
- [60] Wallis, John J., Price Fishback, 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.

Table 1 – Information and the Effects of the Clean Air Act on SO₂ Emissions

Dependent variable: SO ₂ Intensity of State GDP						
	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper	1.322**	-1.370	2.120***	1.932	1.871***	1.611
× after 1970	(0.554)	(3.782)	(0.740)	(4.602)	(0.632)	(4.218)
Newspaper	0.448***	3.853***	0.400**	3.884***	0.425***	3.078***
× t since 1970	(0.117)	(0.836)	(0.156)	(0.992)	(0.134)	(0.905)
State GDP			-1.715**	-0.141	-1.433*	-0.138
× after 1970			(0.714)	(0.099)	(0.775)	(0.103)
State GDP			0.104	-0.001	-0.014	-0.022
× t since 1970			(0.149)	(0.020)	(0.168)	(0.021)
Manufacturing					1.354**	0.307**
× after 1970					(0.560)	(0.155)
Manufacturing					-0.235**	0.070**
× t since 1970					(0.111)	(0.033)
Pop. density					-1.561***	-0.002***
× after 1970					(0.532)	(0.001)
Pop. density					0.465***	0.001***
× t since 1970					(0.104)	(0.000)
State fixed Effects	YES	YES	YES	YES	YES	YES
Year fixed Effects	YES	YES	YES	YES	YES	YES
State time trend	YES	YES	YES	YES	YES	YES
Measures	Bin.	Cont.	Bin.	Cont.	Bin.	Cont.
Observations	864	864	864	864	864	864
R ²	0.967	0.967	0.967	0.967	0.968	0.967

Notes: Robust standard errors in parentheses. Asterisks denote significance respectively at the 10%*, 5%** , and 1%*** confidence level. Emissions are from the EPA; newspaper circulation per capita from the *Statistical Abstract of the United States*; state GDP, manufacturing value added and population density from the BEA Regional Economic Accounts. For all four independent variables, binary measures in odd-numbered columns are dummies for a 1970 value above the mean across states; continuous measures in even-numbered columns are 1970 values.

Table 2 – Information and the Effects of the Clean Air Act on SO₂ Emissions Groups Defined by Time Averages

Dependent variable: SO₂ Intensity of State GDP

	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper × after 1970	0.492 (0.533)	-1.128 (3.962)	0.773 (0.478)	1.331 (4.690)	0.852** (0.433)	1.234 (4.255)
Newspaper × t since 1970	0.593*** (0.111)	3.994*** (0.871)	0.399*** (0.098)	3.570*** (0.994)	0.386*** (0.091)	2.330*** (0.888)
State GDP × after 1970			-0.481 (0.478)	-0.103 (0.092)	-0.132 (0.483)	-0.101 (0.096)
State GDP × t since 1970			0.334*** (0.098)	0.018 (0.018)	0.285*** (0.100)	0.009 (0.019)
Manufacturing × after 1970					0.240 (0.452)	0.261* (0.140)
Manufacturing × t since 1970					-0.027 (0.098)	0.069** (0.030)
Pop. density × after 1970					-1.349*** (0.351)	-0.002*** (0.001)
Pop. density × t since 1970					0.183*** (0.069)	0.000*** (0.000)
State fixed Effects	YES	YES	YES	YES	YES	YES
Year fixed Effects	YES	YES	YES	YES	YES	YES
State time trend	YES	YES	YES	YES	YES	YES
Measures	Bin.	Cont.	Bin.	Cont.	Bin.	Cont.
Observations	864	864	864	864	864	864
R ²	0.967	0.967	0.968	0.967	0.968	0.967

Notes: Robust standard errors in parentheses. Asterisks denote significance respectively at the 10%^{*}, 5%^{**}, and 1%^{***} confidence level. Emissions are from the EPA; newspaper circulation per capita from the *Statistical Abstract of the United States*; state GDP, manufacturing value added and population density from the BEA Regional Economic Accounts. For all four independent variables, binary measures in odd-numbered columns are dummies for a 1963-1980 state average above the mean of state averages; continuous measures in even-numbered columns are 1963-1980 state averages.

Table 3 – Information and the Effects of the Clean Air Act on NO_x Emissions

Dependent variable: NO_x Intensity of State GDP

	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper × after 1970	-0.060 (0.218)	-2.524 (1.724)	-0.225 (0.345)	-5.114 (3.154)	-0.156 (0.300)	-4.153* (2.455)
Newspaper × t since 1970	0.177*** (0.044)	1.872*** (0.360)	0.188*** (0.069)	2.602*** (0.635)	0.163*** (0.060)	2.183*** (0.492)
State GDP × after 1970			0.354 (0.338)	0.110 (0.077)	0.602 (0.401)	0.135 (0.086)
State GDP × t since 1970			-0.024 (0.068)	-0.031** (0.015)	-0.067 (0.081)	-0.041** (0.016)
Manufacturing × after 1970					-0.017 (0.189)	-0.074 (0.083)
Manufacturing × t since 1970					0.060 (0.037)	0.067*** (0.016)
Pop. density × after 1970					-0.695*** (0.201)	-0.001** (0.000)
Pop. density × t since 1970					0.089** (0.039)	-0.000 (0.000)
State fixed Effects	YES	YES	YES	YES	YES	YES
Year fixed Effects	YES	YES	YES	YES	YES	YES
State time trend	YES	YES	YES	YES	YES	YES
Measures	Bin.	Cont.	Bin.	Cont.	Bin.	Cont.
Observations	864	864	864	864	864	864
R ²	0.950	0.951	0.950	0.952	0.951	0.955

Notes: Robust standard errors in parentheses. Asterisks denote significance respectively at the 10%*, 5%** , and 1%*** confidence level. Emissions are from the EPA; newspaper circulation per capita from the *Statistical Abstract of the United States*; state GDP, manufacturing value added and population density from the BEA Regional Economic Accounts. For all four independent variables, binary measures in odd-numbered columns are dummies for a 1970 value above the mean across states; continuous measures in even-numbered columns are 1970 values.

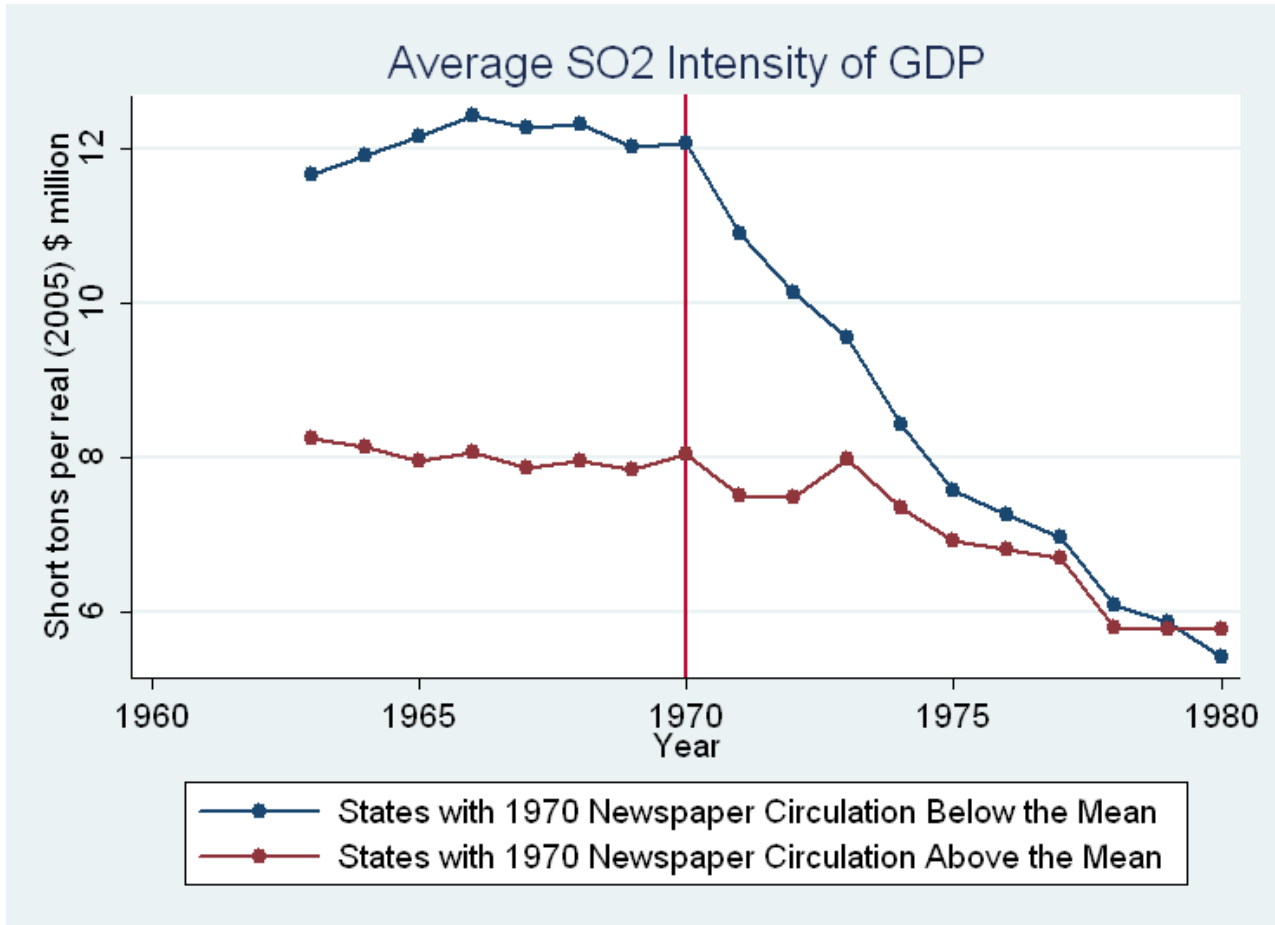
Table 4 – Information and the Effects of the Clean Air Act on NO_x Emissions Groups Defined by Time Averages

Dependent variable: NO_x Intensity of State GDP

	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper × after 1970	-0.032 (0.204)	-2.408 (1.754)	-0.206 (0.387)	-5.284 (3.252)	-0.061 (0.328)	-4.032 (2.521)
Newspaper × t since 1970	0.206*** (0.041)	1.980*** (0.368)	0.272*** (0.078)	3.296*** (0.651)	0.241*** (0.066)	2.702*** (0.502)
State GDP × after 1970			0.299 (0.387)	0.120 (0.085)	0.549 (0.451)	0.129 (0.083)
State GDP × t since 1970			-0.113 (0.078)	-0.055*** (0.017)	-0.162* (0.091)	-0.059*** (0.016)
Manufacturing × after 1970					-0.303 (0.221)	-0.057 (0.068)
Manufacturing × t since 1970					0.075* (0.044)	0.058*** (0.013)
Pop. density × after 1970					-0.522*** (0.175)	-0.001** (0.000)
Pop. density × t since 1970					0.088*** (0.034)	-0.000 (0.000)
State fixed Effects	YES	YES	YES	YES	YES	YES
Year fixed Effects	YES	YES	YES	YES	YES	YES
State time trend	YES	YES	YES	YES	YES	YES
Measures	Bin.	Cont.	Bin.	Cont.	Bin.	Cont.
Observations	864	864	864	864	864	864
R ²	0.950	0.951	0.951	0.955	0.951	0.957

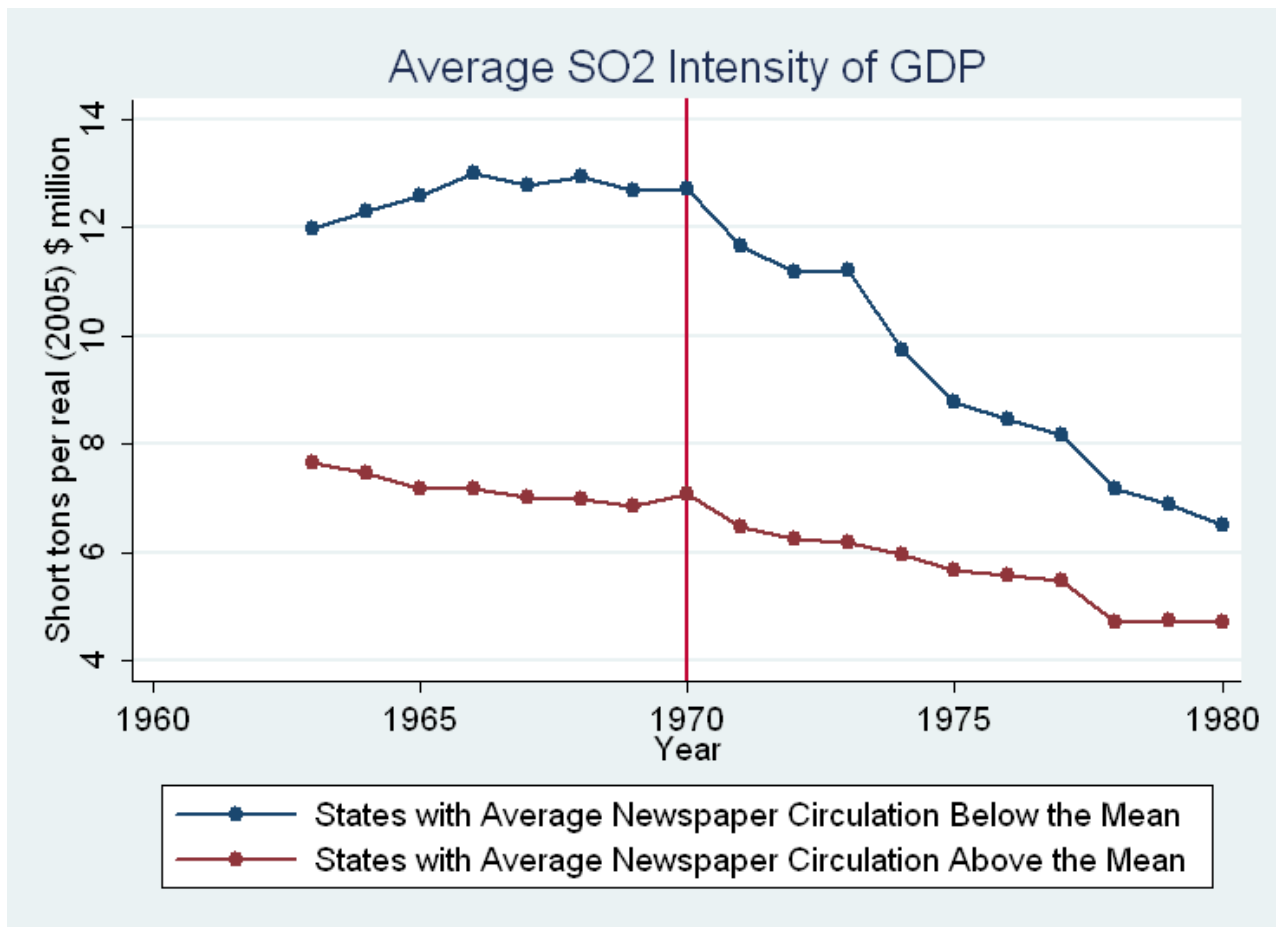
Notes: Robust standard errors in parentheses. Asterisks denote significance respectively at the 10%^{*}, 5%^{**}, and 1%^{***} confidence level. Emissions are from the EPA; newspaper circulation per capita from the *Statistical Abstract of the United States*; state GDP, manufacturing value added and population density from the BEA Regional Economic Accounts. For all four independent variables, binary measures in odd-numbered columns are dummies for a 1963-1980 state average above the mean of state averages; continuous measures in even-numbered columns are 1963-1980 state averages.

Figure 1 – Information and the Effects of the Clean Air Act on SO₂ Emissions



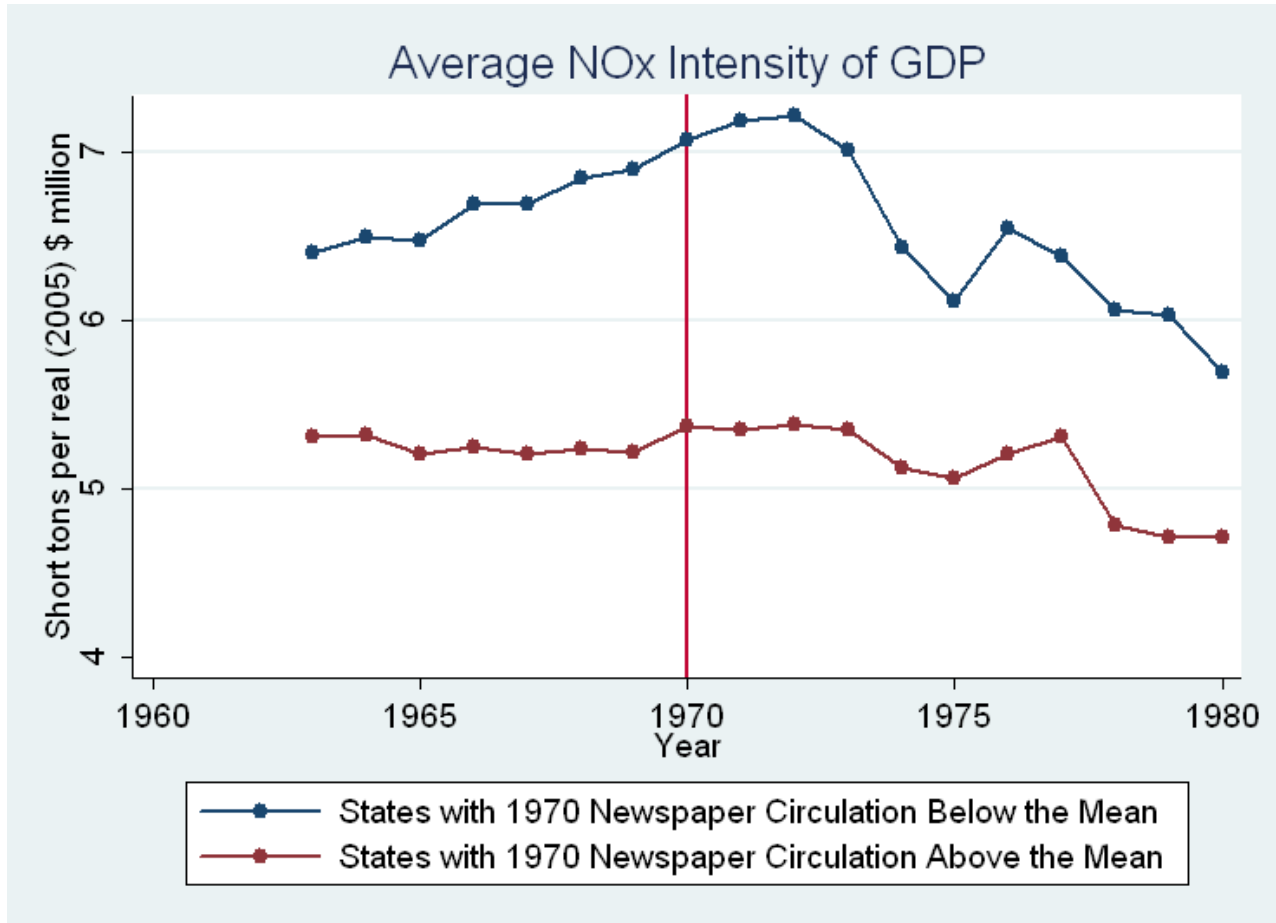
Sources: Emissions are from the EPA, newspaper circulation from the *Statistical Abstract of the United States*, and GDP from the BEA Regional Economic Accounts.

Figure 2 – Information and the Effects of the Clean Air Act on SO₂ Emissions Groups Defined by Time Averages



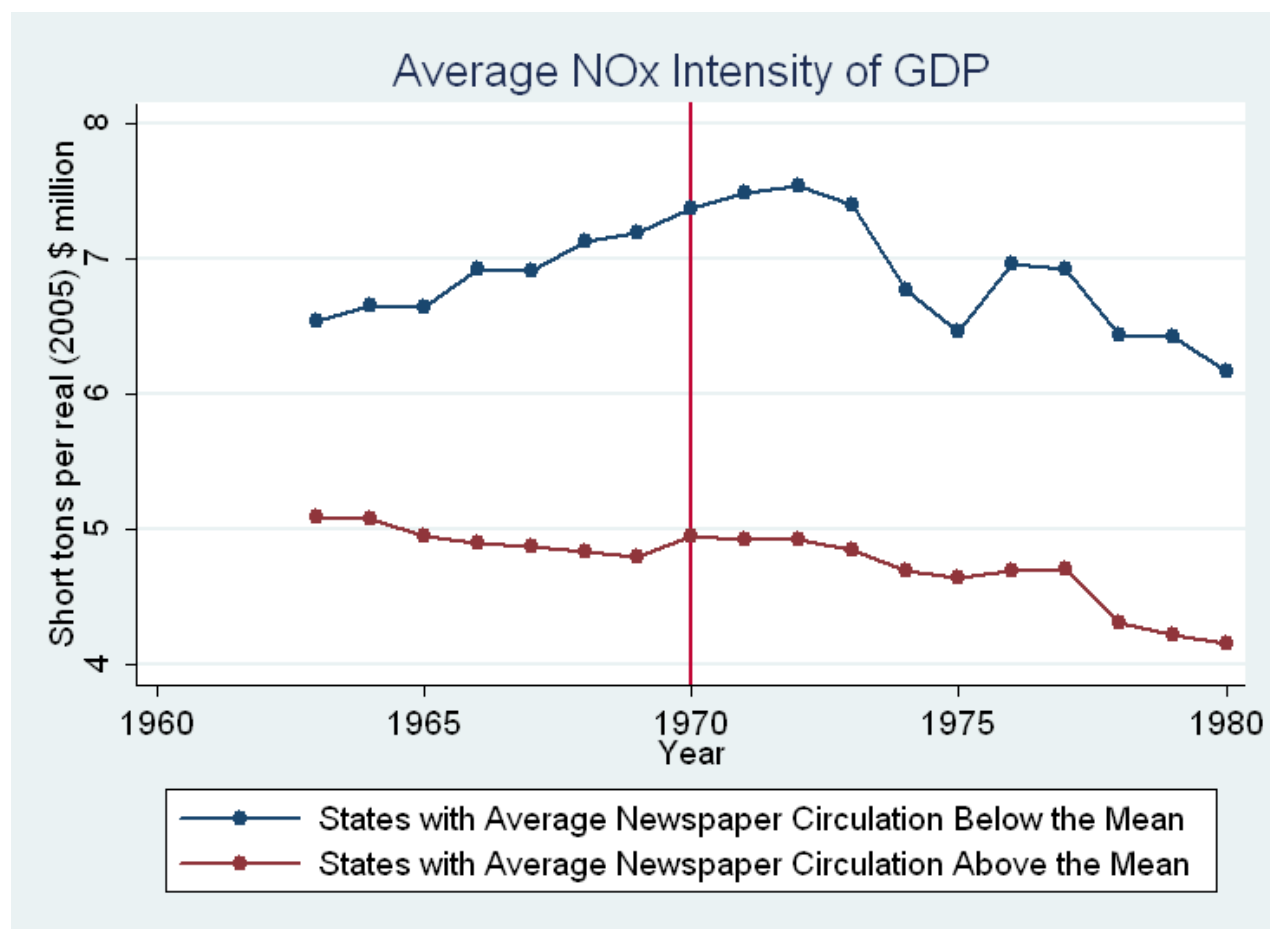
Sources: Emissions are from the EPA, newspaper circulation from the *Statistical Abstract of the United States*, and GDP from the BEA Regional Economic Accounts.

Figure 3 – Information and the Effects of the Clean Air Act on NO_x Emissions



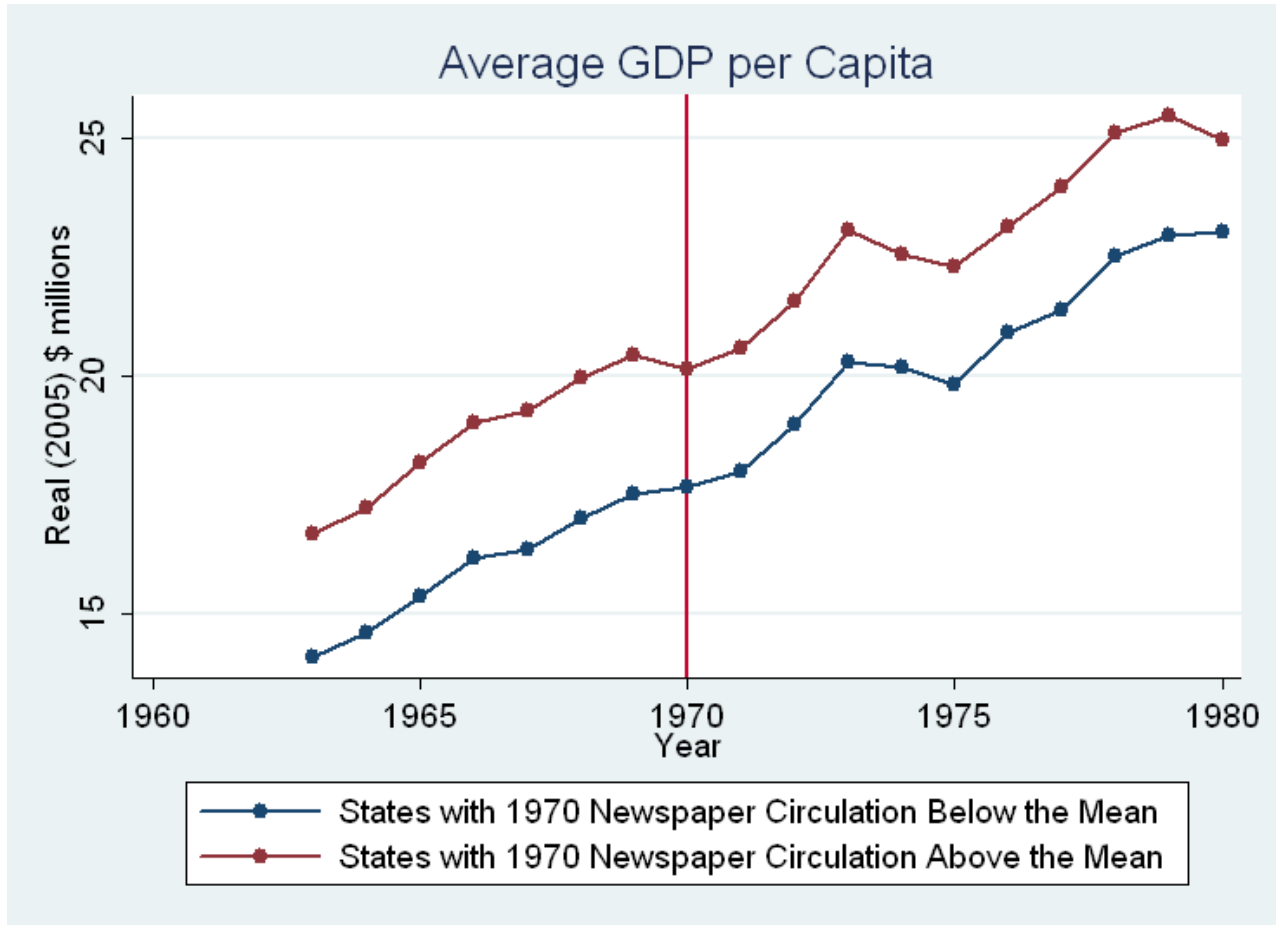
Sources: Emissions are from the EPA, newspaper circulation from the *Statistical Abstract of the United States*, and GDP from the BEA Regional Economic Accounts.

Figure 4 – Information and the Effects of the Clean Air Act on NO_x Emissions Groups Defined by Time Averages



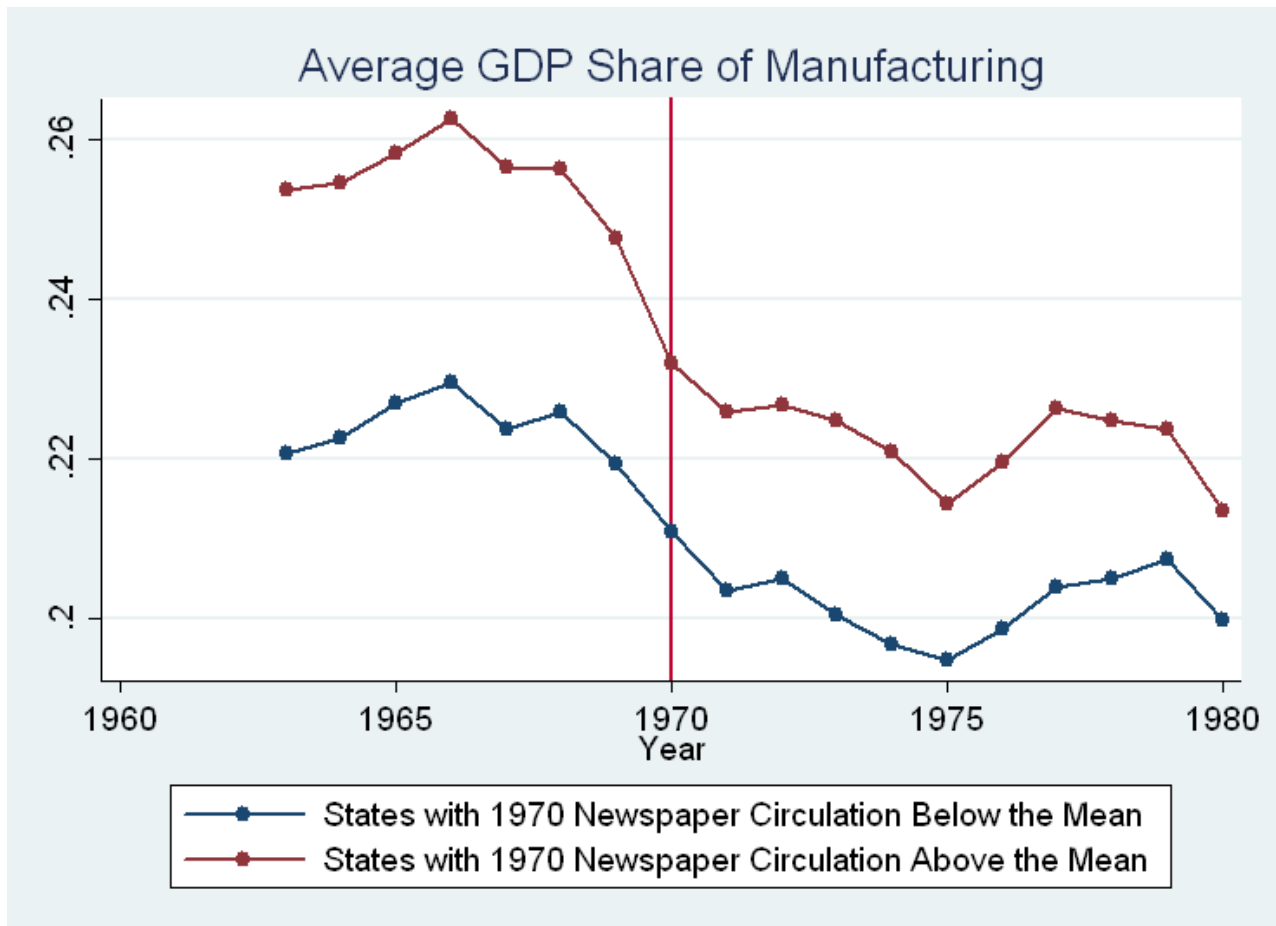
Sources: Emissions are from the EPA, newspaper circulation from the *Statistical Abstract of the United States*, and GDP from the BEA Regional Economic Accounts.

Figure 5 – Information and State GDP per Capita



Sources: Newspaper circulation is from the *Statistical Abstract of the United States* and GDP from the BEA Regional Economic Accounts.

Figure 6 – Information and the Share of Manufacturing in State GDP



Sources: Newspaper circulation is from the *Statistical Abstract of the United States*, GDP and value added in manufacturing from the BEA Regional Economic Accounts.