serie ad

WP-AD 2013-03

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

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

Ivie

Working papers

Working papers

Working papers

Los documentos de trabajo del Ivie ofrecen un avance de los resultados de las investigaciones económicas en curso, con objeto de generar un proceso de discusión previo a su remisión a las revistas científicas. Al publicar este documento de trabajo, el Ivie no asume responsabilidad sobre su contenido.

Ivie working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication. Ivie's decision to publish this working paper does not imply any responsibility for its content.

La Serie AD es continuadora de la labor iniciada por el Departamento de Fundamentos de Análisis Económico de la Universidad de Alicante en su colección "A DISCUSIÓN" y difunde trabajos de marcado contenido teórico. Esta serie es coordinada por Carmen Herrero.

The AD series, coordinated by Carmen Herrero, is a continuation of the work initiated by the Department of Economic Analysis of the Universidad de Alicante in its collection "A DISCUSIÓN", providing and distributing papers marked by their theoretical content.

Todos los documentos de trabajo están disponibles de forma gratuita en la web del Ivie http://www.ivie.es, así como las instrucciones para los autores que desean publicar en nuestras series.

Working papers can be downloaded free of charge from the Ivie website http://www.ivie.es, as well as the instructions for authors who are interested in publishing in our series.

Versión: [iþÆÍ 2013 / Version: éþÆÍ 2013

Edita / Published by:
Instituto Valenciano de Investigaciones Económicas, S.A.
C/ Guardia Civil, 22 esc. 2 1º - 46020 Valencia (Spain)

WP-AD 2013-03

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

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

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 reduces aggregate rent extraction. When the central government provides public goods uniformly across space, 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 and newspaper circulation across the United States. The 1970 Clean Air Act centralized environmental policy at the federal level. In line with our theory, we find that centralization induced a faster decrease in pollution in less informed states.

Keywords: Political centralization, Government accountability, Imperfect information, Interregional heterogeneity, Elections, Environmental policy, Air pollution.

JEL classification numbers: D72, D82, H73, H77, Q58.

* We are grateful to Peter Backus, Paula Bustos, Vasco Carvalho, Edward Glaeser, Matthew Kahn, Gianmarco León, Ana Nuevo-Chiquero, Luigi Pascali, Diego Puga, Thijs van Rens, Kurt Schmidheiny, Hannes Schwandt, Paolo Surico, Yanos Zylberberg, and seminar participants at CEMFI, CREI, Erasmus University Rotterdam, IMT Lucca, Toulouse School of Economics, and UPF 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 ECO-2012-37131), and the Generalitat de Catalunya (grant 2009 SGR 102). Ponzetto acknowledges financial support from the Spanish Ministry of Science and Innovation (grants ECO-2011-25624 and Juan de la Cierva JCI-2010-08414), the Barcelona GSE Research Network and the Generalitat de Catalunya (2009 SGR 1157).

** F. Boffa: Università di Macerata and Barcelona Institute of Economics (IEB) (federico.boffa@unimc.it). A. Piolatto: IEB, Universitat de Barcelona (piolatto@ub.edu). G.A.M. Ponzetto: CREI, Universitat Pompeu Fabra, and Barcelona GSE (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 2000, 2006; 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 post-colonial period (Gennaioli and Rainer 2007a,b). Fiscal centralization was a key element in the modernization of European states: it proved a necessary step for the consolidation of state capacity, which was in turn a critical determinant of economic development (Dincecco 2011; Gennaioli and Voth 2011; Dincecco and Katz 2012). In recent decades, Blanchard and Shleifer (2001) argue that China grew 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 supposed 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 may misallocate personal effort and government resources,

[1] Conversely, 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).

extracting wasteful private rents instead of providing public goods. Such rent-seeking behavior is constrained by electoral discipline. Career concerns induce the incumbent to provide public goods in order to signal ability and win the votes of those citizens who observe public goods provision. In equilibrium, we show that politicians extract lower rents if voters are better informed, consistent with empirical evidence that voter information improves accountability (Besley and Burgess 2002; Adserà, Boix, and Payne 2003; Ferraz and Finan 2008; Snyder and Strömberg 2010; Ponzetto 2011). Our model proves that the link between voter information and political accountability is subject to 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 theory also accounts for the regional distribution of the efficiency gains from centralization. When the central government provides public goods uniformly throughout the union, lower-information regions enjoy a transfer of accountability from their more informed partners. We prove that a region's welfare gains from centralization are then strictly decreasing in its residents' information. Conversely, if the central government can provide public goods dishomogeneously across space, higher-information regions enjoy a transfer of power from their less informed peers. The central government then targets spending to the informed voters who monitor it most closely, consistent with Strömberg's (2004) evidence that discretionary New Deal funds were disproportionately allocated to more informed counties within each state. Our theory highlights the importance of striking a balance between uniform and discretionary public goods provision at the central level. Without any uniformity, centralization would be welfare reducing despite the associated reduction in political rents. Given the right balance between uniform and discretionary items, instead, centralization can be not only welfare increasing but also Pareto efficient.

Our model predicts that each region should benefit in inverse proportion to its residents' information when the central government is empowered to set a uniform policy for the whole union. We test this prediction by studying one of the most prominent instances of centralization in the history of the United States since World War II (Greenstone 2004). The Clean Air Act of 1970 transferred responsibility for pollution regulation from the state and local governments to the federal Environmental Protection Agency. We perform a difference-in-differences analysis using panel data on pollution and newspaper circulation across states.

The empirical evidence supports our theoretical prediction: the federal takeover of environmental policy had a significant differential impact on states according to their level of information. After national air-quality standards were introduced in 1970, pollutant emissions begin to decline, relative to pre-existing trends, considerably faster in states with lower newspaper readership. This finding obtains for different pollutant, and it is robust to a wide range of controls for state-specific economic structure and political ideology. In line with our theory, the differential reduction in emissions seems to represent a benefit of centralization for the uninformed, since we find no evidence of displacement effects across states. Information does not account for a differential contraction in economic activity after 1970, neither in the aggregate nor for specific polluting industries.

2 Theoretical Model

2.1 Political Agency and Public Goods Provision

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

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

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 goods provision, treating $\tilde{u}_t^i$ as an exogenous shock.

Public goods are produced by the government with technology

$$g_{p,t} = e^{\eta_{p,t}} x_{p,t}. \tag{2}$$

The production technology has constant returns to scale: $x_{p,t}$ measure per capita investment in each public good p. Thus, we do not consider economies of scale in public goods provision, which would provide an immediate technological rationale for efficient centralization.

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

$$\eta_{p,t} = \varepsilon_{p,t} + \varepsilon_{p,t-1}. \tag{3}$$

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.

Politicians are self-interested rent-seekers. Each period, the incumbent extracts a rent

$$r_t = b - \sum_{p=1}^{P} x_{p,t}. \tag{4}$$

The simplest interpretation of rent extraction is pecuniary. The incumbent allocates a given government budget b, invariant over time and subject to a balanced-budget constraint. Each period he spends an amount $x_{p,t}$ on the provision of each public good, and devotes the remainder r_t to socially unproductive ends ranging from party finance to outright embezzlement (Persson and Tabellini 2000). While this reading is perhaps the most intuitive, the model can identically represent slacking instead of stealing. The incumbent enjoys an invariant exogenous reward b from holding office, including compensation and perks as well as the "ego rent" of being in power. However, he incurs a cost $x_{p,t}$ from exerting effort to provide each public good. Rent extraction r_t then captures his failure to work diligently in his constituents' interest.

The politician's objective is to maximize rent extraction over his term in office. He discounts future rents by the discount factor $\delta \in (0, 1]$. He faces election at the end of each period, and if ousted he will never return to power. Politicians lack the ability to make credible policy commitment, so the election is not based on campaign promises, but rather on retrospective evaluation of the incumbent's track record. Current rent-extraction is disciplined by career concerns: the incumbent invests in providing public goods because their provision signals his ability and thereby increases his chances of re-election (Holmström [1982] 1999; Persson and Tabellini 2000). In the standard model of political career concerns, all voters observe the level of public goods provision. We relax the assumption of perfect information, and assume instead that each voter i reaches the election with rational expectations based on incomplete and heterogeneous information according to the following timeline.

1. The incumbent politician's past competence shock ε_{t-1} becomes common knowledge.

2. The incumbent chooses investments $\mathbf{x}_t$, 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. Each voter i observes $\mathbf{g}_t$ with probability θ_i; with probability $1 - \theta_i$ he remains completely uninformed.[2] The arrival of information is independent across voters. No voter has any direct observation of ε_t, r_t, or $\mathbf{x}_t$.

[2] This is not inconsistent with knowledge of one's own utility u_t^i. The exogenous component $\tilde{u}_t^i$ may include a stochastic shock, and uninformed voters are unable to distinguish between the effects of the shock and those of $\mathbf{g}_t$.

4. 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 $\boldsymbol{\alpha}^j$, and identical probabilities θ_j of information acquisition. We allow for an intensive margin of political support, following the probabilistic voting approach (Lindbeck and Weibull 1987). Each voter's preferences consist of two independent elements.

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

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

where $g_{p,t+1}^I$ denotes public goods 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 individual likability or the long-standing ideology of their party. Thus, voter i votes for the incumbent if and only if

$$\Delta_i\left(\Omega_t^i\right) \geq \Psi_t + \psi_t^i, \tag{6}$$

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 is i.i.d. across agents. Both variables have uniform distributions with supports respectively $[-1/(2\phi), 1/(2\phi)]$ and $\left[-\bar{\psi}, \bar{\psi}\right]$, sufficiently wide that in a rational expectations equilibrium neither any voter's ballot nor the outcome 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*

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

where

$$\bar{\theta} = \sum\nolimits_{j=1}^{J} \lambda_j \theta_j$$

denotes voters' average information.

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 $\Omega_t^i \in \{0, 1\}$, a binary variable that describes whether he observed public goods provision $\mathbf{g}_t$.

Since no information about the challenger is available, all voters have rational expectations

$$\mathbb{E}\left[\sum_{p=1}^{P} \alpha_p^i \left(\log g_{p,t+1}^C\right) |\Omega_t^i\right] = \sum_{p=1}^{P} \alpha_p^i \log \bar{x}_p. \tag{7}$$

Uninformed voters also have no way of assessing the incumbent's skill innovation $\boldsymbol{\varepsilon}_t$, and thus his future ability $\boldsymbol{\eta}_{t+1}$. Hence they rationally perceive the incumbent and the challenger as identical from the perspective of future public goods provision:

$$\Delta_i (0) = 0. \tag{8}$$

Informed voters, instead, can infer from $\mathbf{g}_t$ the incumbent's competence $\eta_{p,t}$. Their policy preferences are therefore

$$\Delta_i (1) = \sum_{p=1}^{P} \alpha_p^i \mathbb{E} \left(\varepsilon_{p,t} | g_{p,t}\right) = \sum_{p=1}^{P} \alpha_p^i \left(\log g_{p,t} - \log \bar{x}_p - \varepsilon_{p,t-1}\right). \tag{9}$$

In a rational expectations equilibrium their inference turns out to be perfect, accurately revealing $\boldsymbol{\varepsilon}_t$.

Each group j comprises a continuum of agents and the arrival of information is independent across agents, so a share θ_j of its members have observed public goods provision $\mathbf{g}_t$, 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

$$v_t^j = \frac{1}{2} + \frac{1}{2\bar{\psi}} \left[\theta_j \sum_{p=1}^{P} \alpha_p^j \left(\log g_{p,t} - \log \bar{x}_p - \varepsilon_{p,t-1}\right) - \Psi_t\right], \tag{10}$$

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

$$\pi \left(\mathbf{x}_t\right) = \frac{1}{2} + \phi \sum_{j=1}^{J} \lambda_j \theta_j \sum_{p=1}^{P} \alpha_p^j \left(\log x_{p,t} - \log \bar{x}_p\right) \tag{11}$$

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 (endogenous) expected present value. The trade-off between current and future rent extraction leads to policy choices

$$\mathbf{x} = \arg\max_{\mathbf{x}_t} \left\{ b - \sum_{p=1}^{P} x_{p,t} + R\pi\left(\mathbf{x}_t\right) \right\}, \tag{12}$$

namely

$$x_p = \phi R \sum_{j=1}^{J} \lambda_j \theta_j \alpha_p^j \text{ for all } p = 1, ..., P, \tag{13}$$

and thus current rent extraction

$$r = b - \phi\bar{\theta}R. \tag{14}$$

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

$$R = \delta \sum_{t=0}^{\infty} \left(\frac{\delta}{2}\right)^t r = \frac{2\delta}{2-\delta} r. \tag{15}$$

Substituting this relationship into equations 13 and 14, solving the latter, and plugging it into the former, we obtain the equilibrium allocation of the government budget. To economize on notation, we define

$$\tilde{\delta} \equiv \frac{\delta}{2-\delta} \in (0, 1], \tag{16}$$

a convenient rescaling of the politician's discount factor.

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

$$\Psi_t \leq \sum_{j=1}^{J} \lambda_j \theta_j \sum_{p=1}^{P} \alpha_p^j \varepsilon_{p,t}. \tag{17}$$

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

$$\hat{\eta}_t = \chi_{t-1}\left(\varepsilon_{t-1}^I + \varepsilon_t^I\right) + \left(1 - \chi_{t-1}\right)\left(\varepsilon_{t-1}^C + \varepsilon_t^C\right), \tag{18}$$

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

$$\mathbb{E}\hat{\eta}_{p,t} = \mathbb{E}\left(\chi_{t-1}\varepsilon_{p,t-1}\right). \tag{19}$$

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

Proposition 1 *In equilibrium, the politician extract rents $r = \rho b$, with rent extraction*

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

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

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

His expected ability at providing public good p is

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

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.

The key result in the proposition is that rent extraction declines as voters are more informed ($\partial\rho/\partial\bar{\theta} < 0$). Intuitively, information allows voters to monitor the politician more tightly. This result accords with the empirical finding that government performance improves with media scrutiny (Besley and Burgess 2002; Adserà, Boix, and Payne 2003; Ferraz and Finan 2008; Snyder and Strömberg 2010; Ponzetto 2011). Nevertheless, this link between greater information and better governance does not suffice to create incentives for agents to acquire political information, due to the paradox of the rational voter. Since each voter has a negligible (in the model, precisely nil) chance of determining the outcome of the election, he also has vanishing incentives to improve his monitoring ability. Thus, the decision to acquire information $\bar{\theta}$ is not endogenous to the election game, but derives from exogenous characteristics of the electorate. On these lines, Putnam (1993) argues that newspaper readership reflects an individual's civic involvement and social capital.

Rent extraction is also reduced when voters are more likely to be swayed by policy than non-policy considerations ($\partial\rho/\partial\phi < 0$). Then politicians realize their chances of re-election depend more on their perceived competence and less on their likability, and accordingly strive to provide more public goods. Finally, rent extraction declines when patience is greater ($\partial\rho/\partial\delta < 0$). Then politicians are more willing to sacrifice current benefits for a higher probability of remaining in office in the future.

It is impossible to reduce rent extraction to zero because the incumbent's incentive to refrain from extreme rent extraction ($\mathbf{x}_t = 0$) is entirely due to his desire to gain reelection and keep extracting rents in the future. More broadly, reductions in rent extraction through electoral discipline become progressively more difficult as the equilibrium rent declines. All rent-decreasing factors, namely information $\bar{\theta}$, voters' keenness of competence 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. Average government productivity 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 goods 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. While the unconditional probability of re-election is always equal to 1/2, it becomes monotonically more likely that capable incumbents are retained and incompetent ones thrown out. Instead, in the limit case of no information ($\bar{\theta} = 0$), the probability of re-election reflects the popularity shock Ψ_t alone. Then it equals 1/2 not only unconditionally, but also conditional on any realization of $\boldsymbol{\varepsilon}_t$.

Through improvements in both political incentives and political selection, higher average information ($\bar{\theta}$) increases Equilibrium utility is given by

$$\mathbb{E}u_j = \sum_{p=1}^{P} \alpha_p^j \mathbb{E}\log g_{p,t} = \log b + \log(1-\rho) + \sum_{p=1}^{P} \alpha_p^j \left(\mathbb{E}\hat{\eta}_p + \log\beta_p\right) \tag{20}$$

for each member of group j. An equiproportional increase in the information of each voter (θ_j) unambiguously raises the welfare of all. Both political incentives and political selection improve ($\partial\rho/\partial\bar{\theta} < 0$ and $\partial\mathbb{E}\hat{\eta}_p/\partial\theta_j > 0$), while investment shares $\boldsymbol{\beta}$ are unchanged. Welfare also rises with voters' keenness on policy outcomes (ϕ), politicians' patience (δ), and the variance of their ability distribution (σ^2).

Finally, proposition 1 shows that the share β_p of each public good reflects differences in both preferences and information across voters. A utilitarian social welfare planner would set

$$\beta_p = \bar{\alpha}_p \equiv \sum\nolimits_{j=1}^{J} \lambda_j \alpha_p^j, \tag{21}$$

reflecting average preferences for each public good. However, politicians are only imperfectly accountable, and specifically they are held accountable only by informed voters. Thus, their choices deviate from social welfare optimization even beyond the extraction of a rent. The allocation of investment across public goods obey a social welfare function in which each group's preferences are weighted by its relative level of information ($\theta_j/\bar{\theta}$). This result mirrors the finding that information translates into a group's power over political issues such as the allocation of discretionary expenditure (Strömberg 2004), the influence of religion in politics (Glaeser, Ponzetto, and Shapiro 2005), or the structure of trade policy (Ponzetto 2011)

2.3 Centralization and Rent Extraction

To examine the effect of political centralization or decentralization on rent extraction and public goods provision, we consider and economy divided into L regions 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 pools.

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

$$\alpha_{l,p}^{l} = \left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p^l > 0 \tag{22}$$

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

$$\alpha_{m,p}^{l} = \frac{1}{L}\xi_p \alpha_p^l \text{ for } l \neq m \tag{23}$$

from public goods provided in any other region. Thus we can write the utility of individual i in region l as

$$u_t^i = \tilde{u}_t^i + \sum_{p=1}^{P} \alpha_p^l \left[\left(1-\xi_p\right) \log g_{l,p,t} + \frac{\xi_p}{L} \sum_{m=1}^{L} \log g_{m,p,t} \right]. \tag{24}$$

With decentralized government, in each region l a local politician with ability $\eta_{l,p,t}^D$ independently invests in the provision of public goods $x_{l,p,t}^D$ and extracts rent

$$r_{l,t}^D = b - \sum_{p=1}^{P} x_{l,p,t}^D. \tag{25}$$

Following proposition 1, we measure rent extraction in region l under decentralization by $\rho_l^D = r_l^D/b$. This can be immediately interpreted as the fraction of the regional budget b that the politician misallocates. Identically, as discussed above, it could measure the extent to which the local politician enjoys the rewards of office b without exerting effort $x_{l,p,t}^D$.

Centralization means that a single politician with ability $\eta_{p,t}^C$ chooses investment in public goods $x^C_{l,p,t}$ for all l. and extracts rents

$$r^C_t = bL - \sum_{l=1}^{L}\sum_{p=1}^{P} x^C_{l,p,t}. \tag{26}$$

If we interpret rent rent-extraction simply as stealing, the size of the central politician's rent follows immediately because he controls under centralization the same aggregate budget (bL) as the sum of local politicians under decentralization. Under the alternative slacking interpretation, this expression corresponds to the additional assumption that the perks and ego rents of office are similarly additive. In either case, rent extraction under centralization can be measured by $\rho^C = r^C/(bL)$, which has the same normalization as ρ^D_l.

Centralization may also require the central government to provide public goods uniformly across regions ($x^C_{l,p,t} = x^C_{p,t}$ for all l, implying $g^C_{l,p,t} = g^C_{p,t}$ given the common competence $\eta^C_{p,t}$). The literature has typically assumed such a uniformity constraint, and underlined that it provides a simple rationale for decentralization given heterogeneous preferences across regions (Oates 1972; Alesina and Spolaore 1997; Alesina, Angeloni, and Etro 2005). Nonetheless, imposing a uniformity constraint on centralized public goods provision is not necessarily realistic in all settings. Discretionary federal spending is not required to be allocated uniformly across states, almost by definition of discretionality. 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. We encompass both cases by assuming that there is a set $\mathcal{U}$ of public goods whose centralized provision is exogenously subject to a uniformity constraint. The complementary set $\mathcal{D}$ consists of public goods that the central government can instead provide in different amounts to different regions.

The following proposition establishes the beneficial effect of centralization on political accountability, which is independent of the presence of a uniformity constraint.

Proposition 2 *Aggregate rent extraction is lower under centralization than decentralization ($\rho^C \leq \sum_{l=1}^{L} \rho^D_l / L$). It is strictly lower if information is heterogeneous across regions ($\theta_l \neq \theta_m$ for some $l \neq m$), or if there are externalities in public goods provision ($\xi_p > 0$ for some p).*

Average efficiency in providing each public good is higher under centralization than decentralization ($\mathbb{E}\hat{\eta}^C_p \geq \sum_{l=1}^{L} \mathbb{E}\hat{\eta}^D_{l,p}/L$). It is strictly higher if there are externalities in public goods provision ($\xi_p > 0$).

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 for policy coordination. When preferences are identical across regions ($\boldsymbol{\alpha}^l = \boldsymbol{\alpha}$ for all l), all voters agree on the optimal allocation across public goods. Benevolent planners would then attain the first best even if the maximized non-cooperatively each region's welfare. With imperfect political agency, however, spillovers imply a benefit of centralization due to reduced rent extraction rather than inter-regional coordination.

In the election, informed citizens are 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. This is less likely, however, when the economic stakes in the election rise. If there are public goods spillovers, the stakes are higher in a national than in a local election. The ability of a local politician influences only local public goods; the ability of a central politician influences both local public goods and spillovers from other regions. A voter who cares about spillovers ($\xi_p > 0$) will, therefore, be keener on electing a proficient politician at the central than at the regional level. Hence, 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.

Our results can be contrasted with Oates's (1972) Decentralization Theorem. In his classic analysis, centralization is useful to internalize cross-regional externalities. Conversely, decentralization is beneficial to avoid the cost of policy uniformity when regions have heterogeneous preferences. If we follow the standard assumption that all public goods provided by the central government are subject to the uniformity constraint ($\mathcal{D} = \varnothing$), our framework replicates the standard results if and only if voters are homogeneously informed ($\theta_l = \theta$ for all l). With heterogeneous information, Proposition 2 establishes new forces that tend to make centralization welfare increasing.

Proposition 3 *Suppose that all public goods provided by the central government are subject to the uniformity constraint ($\mathcal{D} = \varnothing$) and that information is homogeneous across regions ($\theta_l = \theta$ for all l).*

1. *If there are no externalities and preferences are homogeneous ($\alpha_p^l = \alpha_p$ for all l and $\xi_p = 0$ for all p), then centralization and decentralization yield identical outcomes.*
2. *If there are externalities and preferences are homogeneous ($\alpha_p^l = \alpha_p$ for all l, while $\xi_p > 0$ for some p), then centralization yields higher welfare than decentralization.*
3. *If there are no externalities and preferences are heterogeneous ($\xi_p = 0$ for all 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 all public goods provided by the central government are subject to the uniformity constraint ($\mathcal{D} = \varnothing$) and that information is heterogeneous across regions ($\theta_l \neq \theta_m$ for some $l \neq m$). If preferences are homogeneous ($\alpha_p^l = \alpha_p$ for all l), centralization yields higher welfare than decentralization regardless of externalities ($\xi_p \geq 0$).

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 rent. 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 are exogenously incapable of cooperating. 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 goods provision.

The third point highlights the standard cost of a binding uniformity constraint. When regions have heterogeneous preferences,uniformity implies 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 reduces rent extraction. Lower rents imply greater public goods provision, and accordingly Proposition 3 established that centralization is welfare-increasing from the national perspective when preferences are homogeneous and public goods are uniformly provided. A more thorough assessment of the consequences of centralization must take into account the uneven distribution of its benefits across regions. The distinction between uniform and discretionary public goods provision by the central government moves to the forefront in this analysis.

A uniformity constraint implies that centralization transfers accountability from the more to the less informed regions. Thus, the welfare gains described by Proposition 3 accrue disproportionately to the uninformed.

Proposition 4 *Suppose that all public goods provided by the central government are subject to the uniformity constraint ($\mathcal{D} = \varnothing$), and that preferences are homogeneous but* *information is heterogeneous across regions ($\alpha_p^l = \alpha_p$ for all l, while $\theta_l \neq \theta_m$ 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 $\mathbb{E}\left(g_{l,p}^C - g_{l,p}^D\right) > \mathbb{E}\left(g_{m,p}^C - g_{m,p}^D\right)$ for all p and $\mathbb{E}\left(u_l^C - u_l^D\right) > \mathbb{E}\left(u_m^C - u_m^D\right)$).

If furthermore there are no externalities in public goods 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 average information exploited in electing their rulers falls when they join the less informed parts of the union in a national election. On the other hand, if public goods spill over across regions ($\xi_p > 0$) 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.

Instead, regions with below-average information clearly gain from centralization. The implicit transfer of accountability effected by centralization effectively entails a welfare-increasing transfer of public resources, although every region contributes an equal amount (b) to 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.

Proposition 4 then implies that empowering the federal government to set a uniform nation-wide policy should benefit disproportionately the states with lower voter information. The next section tests this prediction with evidence from environmental policy in the United States. 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 2013). 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.

The distributional consequences of centralization are completely inverted if the uniformity constraint is relaxed. Unconstrained discretional policies favor more politically influential groups. 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 ($p \in \mathcal{D}$). 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. Without a uniformity constraint, centralization transfers power from the less to the more informed regions.

The balance between the two countervailing distributional forces depends on the relative importance of the two types of public goods. When preferences are homogeneous, the balance can be summarized by the welfare weight of public goods subject to the uniformity constraint:

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

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 regions, 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 ($\alpha_p^l = \alpha_p$for all l, $\theta_l \neq \theta_m$ for some $l \neq m$, and $\xi_p = 0$ for all p).*

1. *There exists a threshold $\bar{\alpha}_{\mathcal{U}} \in \left(0, 1-\rho^C\right)$ 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 interregional 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: $\beta^C_{l,p}/\beta^C_{m,p} = \theta_l/\theta_m$ 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 $\alpha_{\mathcal{U}} \to 0$, uninformed 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.

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 $\rho^D_l < \rho^C$. 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 $\rho^D_l > \rho^C$ is worth more than the loss of control 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.

3 Evidence from the Clean Air Act

A clear discontinuity in U.S. environmental policy allows us to test the fundamental empirical prediction of our model. Up to the 1960s, air pollution had been regulated primarily by state and local governments. 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 polluters. National standardization stood in sharp contrast with the previous state-based regulations, which had been adopted only by a few states, imposing very heterogenous standards (U.S. Senate, 1970). We use the Clean Air Act, and the sharp regulatory shift it entailed, to test the distributional predictions of our model for a uniform nation-wide policy, as derived in Proposition 4.

We consider emissions of sulfur dioxide and nitrogen oxides, two very significant and harmful pollutants. At relatively high concentrations, NO_x and especially SO_2 pollution has serious adverse health effects. It harms respiratory and cardiovascular functions and is a cause of premature death. Even at much lower levels, it severely damages crops and contributes to acid rain. Both SO_2 and NO_x were among the pollutants immediately targeted by the Clean Air Act, starting in 1971. After the National Ambient Air Quality Standards were imposed, emissions for both sulfur dioxides and nitrogen oxides declined drastically. While this decline has been extensively studied, the extent of the causal link between the legislation and the downward trend in emissions is still debated (List and Gallet 1999; List and Gerking 2000; Greenstone 2002, 2004; List and Sturm 2006; Bulte, List, and Strazicich 2007; Greenstone, List, and Syverson 2012).

Our empirical investigation focuses on the differential impact of the Clean Air Act across states. We do not aim at estimating the aggregate effect of the policy shift. Instead, we test whether the decline in emissions after the federal takeover in 1970 was faster in states with less informed voters, as predicted by our theory.

The effect of the Clean Air Act on SO_2 and NO_x emissions is an appropriate natural experiment to test the predictions of Proposition 4. The Act introduced uniform regulation for the entire country, with national air quality standards and a single federal regulator. At the same time, the consequences of SO_2 and NO_x pollution are mostly localized, in contrast to the case of CO_2 and global warming. Therefore, emissions abatement generated benefits primarily at the local level, with at most moderate spillovers across states.

Hence, our model predicts that until 1970 states with uninformed voters suffered from bad environmental regulation. They applied ineffective standards, or no standards whatsoever, because the local government in charge failed to invest money and regulatory effort on air quality control. The introduction of uniform federal requirements starting in 1971 should have yielded differential benefits that are monotone decreasing in voter information. Our testable hypothesis is that the lower the level of information in a state, the more rapid the decline in pollutant emissions in that state after the enactment of the Clean Air Act, relative to the pre-1970 baseline. In the robustness analysis, we also test that this faster reduction in pollution corresponds to an improvement in technique, rather than a change in the composition or scale of economic activity, whose welfare consequences would be more ambiguous.

3.1 Empirical Specification

Our econometric analysis consists of a difference-in-differences estimate of the effect of information on the reduction in emissions following the imposition of national air quality standards in 1970.

We use a balanced panel of the 48 contiguous United States from 1960 to 1981. We choose this time horizon because in 1981 Ronald Reagan took office. In the following years, his environmental policy choices weakened the EPA and curtailed its budget and staff. Moreover, the Reagan administration championed devolution and a general expansion of the role of the states. Therefore, developments after 1981 may have entailed a roll-back of the centralization episode that is the focus of our analysis.

Our baseline regression specification is the following:

$$p_{i,t} = \alpha + \delta_t + \zeta_i + \theta_i t + d_t x_i' \beta + d_t x_i' \gamma \left(t - 1970\right) + \varepsilon_{i,t}. \tag{28}$$

We include year fixed effects δ_t, state fixed effects ζ_i, and state-specific linear time trends θ_i. We cluster the standard errors $\varepsilon_{i,t}$ by state to account for serial correlation of state-specific shocks. Since spatial correlation is also likely to be present, in our main specification we allow for two-way clustering by year as well as by state (Cameron, Gelbach, and Miller 2011). A concern with this specification is that the number of clusters in the time dimension is relatively small (22 years). Therefore, we also report in the appendix all our results with one-way clustering by state only. Since the standard errors are very close in the two cases, we are confident in the validity of two-way clustered standard errors in our application.

As the first difference, we compare pollution $p_{i,t}$ before the Clean Air Act ($d_t = 0$ from 1960 to 1970) and after its enactment ($d_t = 1$ from 1971 to 1981). The difference in differences explores differential changes depending on a vector of state characteristics x_i. Both our key explanatory variable and all additional controls are measured taking

1970 as the reference year. We consider two interactions of the regressors x_i. First, the interaction with the indicator variable d_t would capture a level break (β) in the series upon the introduction of federal emission standards. Second, we add an interaction with the time elapsed since the creation of the EPA ($d_t (t - 1970)$). This would capture a break (γ) in the trend of emissions after the reform.

The switch to federal regulation should have an impact on the trend rather than the level of emissions. The effect of regulatory changes is necessarily gradual because the object of regulation is durable capital that is only gradually scrapped and replaced. In fact, the standards introduced by the Clean Air Act and its subsequent amendments stipulated more stringent regulation on new pollution sources than on pre-existing ones. Motor vehicles arguably provide the starker example: increasingly strict requirements were mandated for successive model-years, so over time tighter standards applied to a steadily growing share of the U.S. vehicle fleet (Kahn 1996). The same pattern holds for stationary sources: a particularly significant case is differential regulation of old and newly built power plants, the main source of SO_2 emissions (Nelson, Tietenberg, and Donihue 1993).

In addition to the technological constraint of natural capital turnover, the implementation of the Clean Air Act was also somewhat gradual. The National Ambient Air Quality Standards were defined in 1970, but compliance was expected to be achieved by 1975. The number of operating monitors reading the concentration of air pollutants, a key factor in enforcement, increased steadily throughout the 1970s (Greenstone, 2004).

3.2 Data

The outcomes of interest are sulfur dioxide and nitrogen oxides emissions. We use the same data as Bulte, List, Strazicich (2007), who obtained from the EPA panel data on SO_2 and NO_x emissions in the 48 contiguous United States from 1929 to 1999. The availability of pollution data dictates the level of disaggregation of our analysis. States are the finest geographic unit for which we have emissions data both before and after the Clean Air Act. At the county level, pollution data are completely unavailable before 1969; even after the creation of the EPA, emissions were monitored in a very small subset of countries until the second half of the 1970s.[3]

Our preferred measure of pollution is the emission intensity of income, measured in tons per real dollar. This choice reflects our focus on improvements in technique, and represents our first step in controlling for a potential reduction in the size of economic activity due to a tightening of environmental regulation. In Tables 7 and 8 we check that our results are robust to alternative scalings of the dependent variable, including

[3] Sulfur dioxide monitors were initially operating in 16 counties, and the sample did not grow to 100 until 1974.

emissions per capita and the density of emission per square mile.

We proxy citizens' information with average daily newspaper circulation per capita. This is a particularly apt measure of voters' ability to acquire information about government policy. Up to the 1980s, newspapers were Americans' main source of political news. Moreover, newspaper readers are better informed and more involved in politics than consumers of other media (Graber 1984; Putnam 1993, 2000; Gentzkow 2006; Gentzkow, Shapiro, and Sinkinson 2011). Even in recent years, newspapers remain an influential source of information on politics and environmental issues. When they are subject to greater newspaper coverage, members of Congress more actively support their constituents' interests (Snyder and Strömberg 2010), and polluting plants reduce the amount of their toxic emissions (Campa 2013). We obtain circulation data from Gentzkow, Shapiro, and Sinkinson's (2011) dataset. Since their data are only available for presidential election years, we select as our measure the average of newspaper circulation per capita in 1968 and 1972.

Figure 1 shows the pattern of newspaper circulation per capita across the United States around 1970. The colors depict four bands whose boundaries are at the cross-state mean and one standard deviation above and below it. Some geographic clustering of high- and low-information states is apparent. Beyond allowing for arbitrary spatial correlation of the residuals through our two-way clustering strategy, we also check that the results are not driven by region-specific common factors. We can focus on the role of information within each Census Region by adding to the vector of explanatory variables x_i a full set of dummies for the four regions.

Standard economic variables are from the BEA Regional Economic Accounts. We use average personal income both to construct the dependent variable and as a control variable.[4]. We use Census population estimates to compute both newspaper circulation per capita and population density.[5] Moreover, we exploit the decomposition of state value added by major industry groups (2-digit SIC sector).

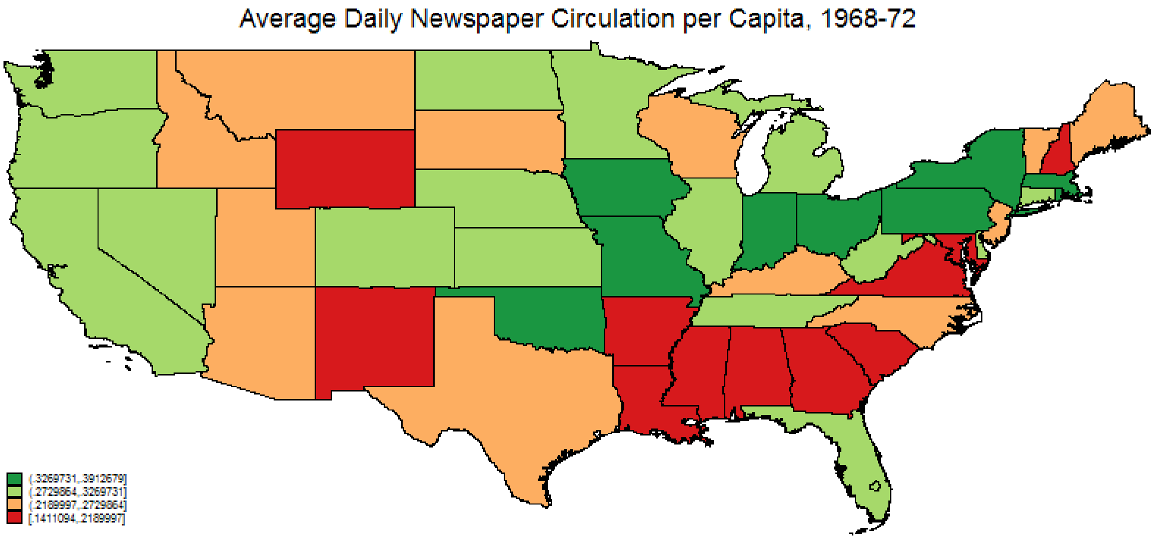
Following Greenstone's (2002) methodology, we categorize a sector as a heavy polluter on the basis of EPA estimates of its contribution to total emissions by industry as a whole. This procedure clearly identifies five polluting manufacturing industries: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), and Primary metal industries (SIC 33). Each of these industry groups accounts for more than 10% of the aggregate industrial emissions of either of the two pollutants we consider. Instead, every other 2-digit SIC sector contributes less than 5% of both SO_2 and NO_x emissions.

In addition to the share of value added contributed by the sum of these five polluting

[4]We transformed amounts originally in current dollars into constant real dollars by using the U.S. GDP deflator.

[5]Land area for each state is taken from the 2000 Census.

Figure 1 – Information across the United States



Source: Gentzkow, Shapiro, and Sinkinson (2011) industries, our controls include the share of manufacturing and the share of Electric, gas, and sanitary services (SIC 49). The latter is relevant because it includes power generation. Electric utilities were responsible for almost 56% of anthropogenic SO_2 emissions in the United States in 1970 (90% of which from coal-fueled power plants), while the entire industrial sector accounted for 29%—half due to metals processing and the other by the combustion of high-sulfur fuels in a variety of industrial processes. For NO_x emissions, the contribution of electric utilities was 23%, compared to 20% for industrial fuel combustion. The single main source of nitrogen oxides was instead on-road vehicles, which accounted for 35% of the total (but for less than 2% of sulfur dioxide emissions).

Since both SO_2 and NO_x are primarily released by the combustion of fossil fuels, we construct additional control variables using data on state energy consumption provided by the State Energy Data System (SEDS) database of the Energy Information Administration (EIA). We compute the fossil-fuel intensity of income as the ratio of aggregate consumption of all fossil fuels (in Btu) to aggregate personal income (in real dollars). Additional controls are the shares of coal and of motor gasoline in total fossil-fuel consumption. The former captures the dominant role of coal combustion in SO_2 emissions, and the latter the importance of motor vehicles for NO_x emissions.

A further set of control variables captures the political environment of each state, beyond our main focus on voter information. We measure ideology by the average DW-Nominate score of the state's two U.S. senators. The score ranks each member of Congress according to his ideology (left to right), based on legislative roll-call voting behavior (Poole and Rosenthal, 1985). It proxies for the ideological orientation of the voters the senators represent.

We consider two measures of partisanship. First, we compute the share of political offices controlled by Republicans in mid-1970. We consider a total of six positions: the governorship, the two U.S. Senate seats, the majority leaderships in the state senate and in the state house of representatives, and finally the majority in the state delegation to the U.S. House of Representatives. We collected data on state legislatures from the *Book of the States*, and on all other institutions from *Wikipedia*.

Our second measure of partisanship is the share of Republican votes in gubernatorial elections. Gentzkow, Shapiro, and Sinkinson (2011) provide the data for elections that took place in the presidential years 1968 and 1972. We collected data on gubernatorial elections from 1969 to 1971 from *Wikipedia* and the website www.ourcampaigns.com. Our measure of partisanship around 1970 is the average share of the vote won by the Republican party in all gubernatorial elections in the state from 1968 to 1972. We use the same data to construct a standard measure of electoral evenness. For every gubernatorial election, the index is defined by $1 - |\%\text{ Republican} - 0.5|$. Hence, it equals zero if either party wins 100% of the votes, and one if the two parties are exactly tied. Again, our overall measure is the average of the indices for all gubernatorial elections in the state from 1968 to 1972.

We also control for proximity to the election of the state's U.S. senators. The U.S. Senate is divided into three classes of senators serving overlapping six-year terms, so that a third of the seats are scheduled for re-election every two years. In the period of interest, Class 2 senators were up for re-election in 1972, Class 3 senators in 1974, and Class 1 senators in 1976. We include among our controls a full set of dummies for the three possible cases, i.e., senators of Classes 1 and 2, 1 and 3, or 2 and 3. Finally, we include as a control variable the share of white population (from the U.S. Census), to check for the possibility of environmental racism.

3.3 Results

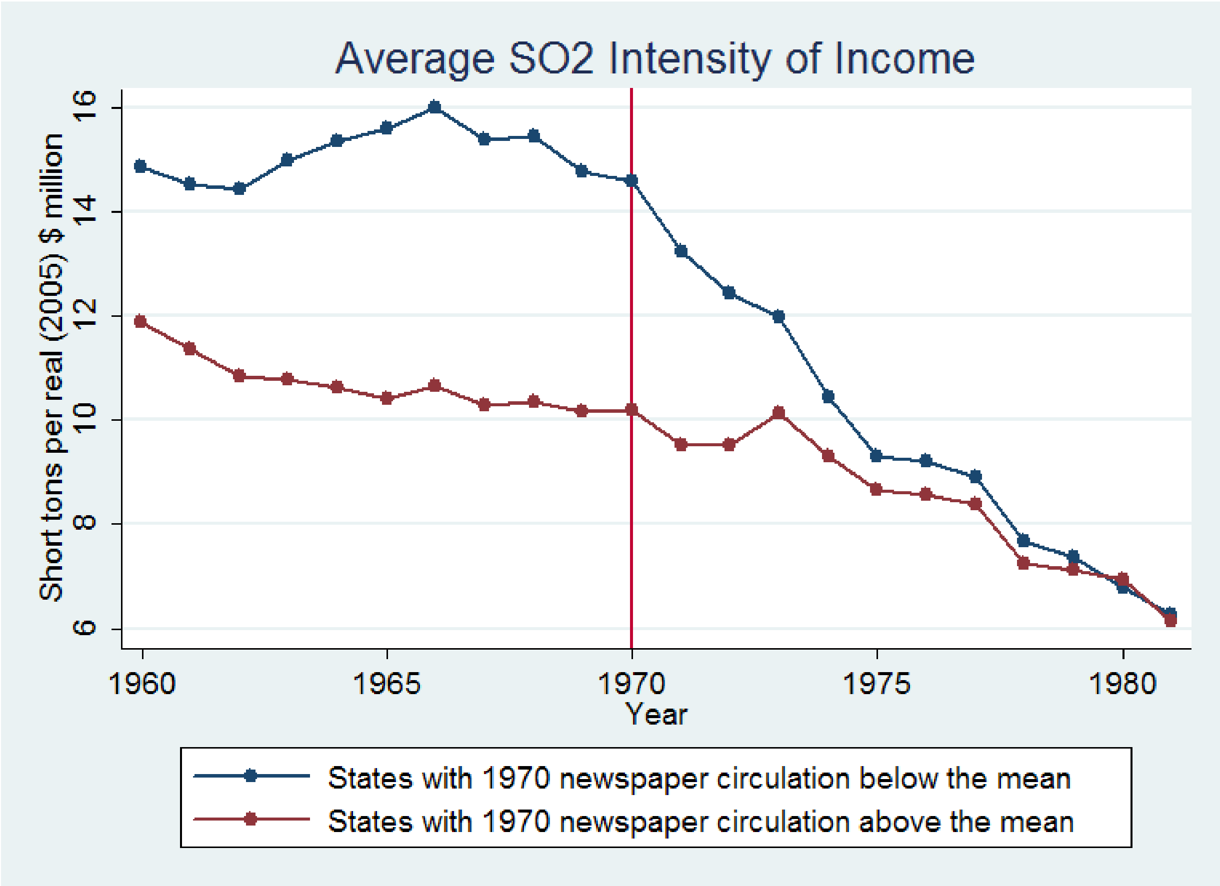
Before turning to our regression analysis, we can starkly visualize the main result in Figure 2. The graph plots average sulfur dioxide emissions for two group of states: those with above-average newspaper circulation in 1970, and those with below-average newspaper circulation. The difference-in-differences 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, but the break in the trend is dramatic.

Table 1 confirms this result in our full regression specification. Column (1), with no controls, shows a differential break in the trend of SO_2 emissions after 1970, which is significant at the 10% level. The impact of the Clean Air Act on pollution is a function of the level of information in each state. As expected, the break in levels is not significant, but the trends change differently as a consequence of the phasing in of federal air quality standards. Consistent with our theoretical prediction, the speed of emissions abatement following the Clean Air Act is inversely proportional to newspaper circulation. If newspaper circulation in 1970 differed across two states by an amount equal to the standard deviation of the cross-state distribution (.05 copies per person), the rate of decline of SO_2 emissions in the 1970s compared to the 1960s was higher in the less informed state by 1.7 percentage points per year.

The following columns sequentially add a set of controls. Their inclusion never has a material impact on our key finding. The point estimate for the coefficient on newspaper circulation is remarkably stable. Moreover, its precision increases as we add controls, reaching the 5% and even 1% significance threshold. These results strengthen the empirical support for the prediction of Proposition 4, and indicate that the effect of newspaper circulation is not due to its correlation with an omitted variable.

Our controls are meant to test for alternative determinants of a heterogeneous impact of the Clean Air Act on sulfur dioxide abatement across states. The main concern is that some drivers of differential effects would suggest that the uniform environmental

Figure 2 – Information and the Impact of the CAA on SO_2 Emissions



Sources: Emissions are from the EPA, newspaper circulation from Gentzkow, Shapiro, and Sinkinson (2011), and personal income from the BEA Regional Economic Accounts.

Table 1 – Information and the Effects of the CAA on SO_2 Emissions

Dependent Variable: SO_2 Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	-1.641	-0.155	-0.158	-0.092	0.430	-0.102	0.392
× after 1970	(1.095)	(0.799)	(0.774)	(0.836)	(1.127)	(0.773)	(1.100)
Newspaper Circ.	0.319*	0.389*	0.390**	0.424**	0.517**	0.357**	0.431***
× t since 1970	(0.179)	(0.192)	(0.185)	(0.191)	(0.211)	(0.141)	(0.132)
Income (log)		-0.767	-0.535	-0.776	-0.337	-0.740	-0.198
× after 1970		(0.587)	(0.549)	(0.595)	(0.640)	(0.576)	(0.609)
Income (log)		-0.099	-0.130	-0.104	-0.018	-0.115	-0.048
× t since 1970		(0.087)	(0.091)	(0.085)	(0.105)	(0.069)	(0.069)
Pop. Density (log)		-0.048	-0.099	-0.046	-0.019	-0.039	-0.053
× after 1970		(0.057)	(0.073)	(0.056)	(0.062)	(0.052)	(0.059)
Pop. Density (log)		0.010	0.017*	0.012	-0.001	0.005	-0.002
× t since 1970		(0.008)	(0.009)	(0.007)	(0.011)	(0.006)	(0.007)
Polluting Mfg. %			3.278*				2.108
× after 1970			(1.611)				(1.246)
Polluting Mfg. %			-0.441				0.056
× t since 1970			(0.261)				(0.164)
DW-Nominate				0.066			0.001
× after 1970				(0.190)			(0.164)
DW-Nominate				0.036			-0.004
× t since 1970				(0.037)			(0.018)
Midwest					0.428**		0.394*
× after 1970					(0.201)		(0.195)
Midwest					-0.022		0.007
× t since 1970					(0.024)		(0.021)
South					0.600**		0.553**
× after 1970					(0.240)		(0.215)
South					0.029		0.038
× t since 1970					(0.035)		(0.026)
West					0.454		0.410
× after 1970					(0.286)		(0.285)
West					-0.026		0.002
× t since 1970					(0.044)		(0.034)
SO_2 Intensity (log)						0.080	0.032
× after 1970						(0.061)	(0.052)
SO_2 Intensity (log)						-0.049***	-0.050***
× t since 1970						(0.007)	(0.006)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.939	0.941	0.942	0.941	0.943	0.946	0.949

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Independent variables are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

regulation imposed by the federal government is inefficient or even harmful for some states. This would be inconsistent with our interpretation of pollution abatement as a public good.

Column (2) adds the basic controls for income and population density. Controlling for income addresses concerns related to the environmental Kuznets curve. According to this line of reasoning, air quality is a luxury good, so poorer states desire a higher level of pollution (List and Gallet, 1999). The Clean Air Act might then have imposed an inefficient emissions reduction on poor states, compelling them to reduce emissions to a level suitable only for richer regions. However, the coefficient on income is insignificant and has a negative point estimate, which seems to rule out this interpretation.

Controlling for population density speaks to a similar concern. Since pollution is more harmful in more densely populated areas, due to its localized adverse health effects, the Clean Air Act might have suboptimally compelled a greater reduction of emissions in low-density areas. This does not seem to be the case, since there is no significant difference in differences based on population density—even if the point estimate is positive in this case, and marginally significant in column (3).

Column (3) also finds some evidence that the introduction of federal regulation had differential effects depending on the concentration of polluting manufacturing industries. Surprisingly, the regression suggests that the Clean Air Act immediately induced a greater decline in pollution in states with a lower GDP share of polluting manufacturing. We should not overemphasize this result: the coefficient becomes insignificant when further controls are added in column (7). Moreover, it is paired with a large albeit imprecisely estimated coefficient on the trend, so that any differential effect would seem to disappear before the end of the decade. In any case, the absence of a significant negative coefficient dispels the concern that the federal regulations might have forced overly stringent regulations on regions with a specialization in polluting manufacturing industries.

The result also suggests that the effects of the Clean Air Act, and local policy before its enactment, were not determined primarily by industry lobbying. Under this interpretation, polluting industries would have pressured state politicians into adopting overly lax regulations. This alternative scenario would also have implied a faster emissions reduction after 1970 in states with a greater specialization in polluting manufacturing, albeit with associated welfare gains. The data, however, do not point to a sharper incidence of federal regulation on states with a greater concentration of polluting industries.[6]

Column (4) introduces ideology. We might be concerned that the states most affected by federal environmental regulation are those that ideologically oppose it. We proxy voters' ideology with the ideology of their elected representatives, and precisely with

[6] Incentives for polluting industries to lobby against environmental regulation have been recently questioned by Ryan (2012), who shows that incumbent polluters benefit from stricter regulation, which softens competition by increasing the cost of entry.

the average DW-Nominate score of each state's U.S. senators. We find no statistically discernible effect. This suggests that differences in the impact of the Clean Air Act are driven by differences in accountability, consistent with our model, rather than in ideology.

Column (5) includes region fixed effects, addressing the concern that the groups of more and less informed states have an imperfect geographic balance, as seen in Figure 1. We find that newspaper readership is an even stronger and more significant determinant of the differential impacts of the Clean Air Act within than across regions.

Column (6) controls for SO_2 intensity in 1970. As expected, this control is highly significant and hardly affected by the inclusion of additional controls in column (7). States that started with higher emissions have to reduce them more rapidly as they converge to a uniform national level of pollution (Kahn 1997). Nonetheless, the effect of information persists almost unchanged, and is more precisely estimated.

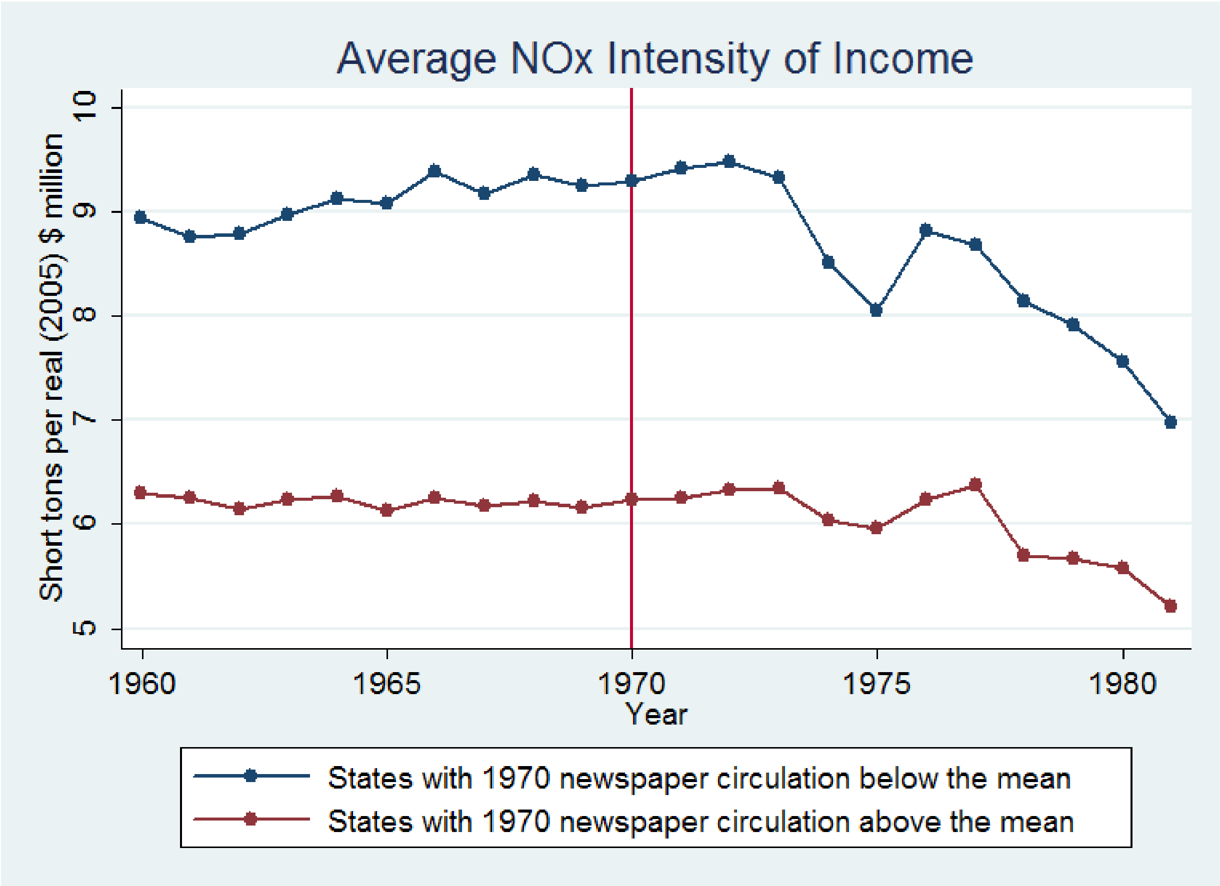
Figure 2 might have suggested that our finding was purely driven by a correlation between information and the level of pollution before 1970. In fact, the results are stronger than a mere cross-section for the period before 1970. The difference-in-differences specification allows us to test for state fixed effects and constant state-specific trends. In column (6) we also establish that less informed states converge faster to the federal standard regardless of their initial distance. This combined evidence supports the reading that newspaper circulation explains not simply 1970 pollution as a whole, but more precisely a specific inefficient component of the emissions level. Consistent with Proposition 4, lower voter information implied greater inefficiency of state regulation, and thus a faster improvement after the switch to federal regulation.

Figure 3 and Table 2 present the results for nitrogen oxides. These are essentially the same as those for sulfur dioxide. Once again there is a significant, robust and stable differential break in the emissions trend after 1970 depending on the level of newspaper circulation. The effect is smaller, though in general more precisely estimated. The point estimate in column (1) implies that, for two states whose 1970 newspaper circulation differed by one standard deviation (.05 copies per person), the rate of decline of NO_x emissions in the 1970s compared to the 1960s was higher in the less informed state by 0.9 percentage points per year.

When controls are added in columns (2) to (7), the initial level of emissions is strongly significant while others are typically insignificant, just as in Table 1. In column (3), all coefficients except the one on information are now statistically indistinguishable from zero. This strengthens our conclusion that the Clean Air Act did not have differential effects depending on a state's specialization in polluting manufacturing industries.

If there is a difference between the patterns of abatement for SO_2 and NO_x emissions, it is that for the latter, once initial pollution is controlled for, population density and income significantly predict a faster reduction. This difference may plausibly reflect the fact that motor vehicles are responsible for releasing a large share of NO_x (35%) but

Figure 3 – Information and the Impact of the CAA on NO_x Emissions



Sources: Emissions are from the EPA, newspaper circulation from Gentzkow, Shapiro, and Sinkinson (2011), and personal income from the BEA Regional Economic Accounts.

Table 2 – Information and the Effects of the CAA on NO_x Emissions

Dependent Variable: NO_x Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	0.041	0.121	0.120	0.150	0.557	0.165	0.552
× after 1970	(0.256)	(0.316)	(0.296)	(0.335)	(0.395)	(0.325)	(0.392)
Newspaper Circ.	0.160**	0.190**	0.190**	0.187**	0.189*	0.164**	0.157**
× t since 1970	(0.067)	(0.082)	(0.082)	(0.085)	(0.097)	(0.059)	(0.065)
Income (log)		0.020	0.087	0.016	0.055	0.142	0.138
× after 1970		(0.173)	(0.177)	(0.180)	(0.210)	(0.183)	(0.233)
Income (log)		-0.005	-0.006	-0.004	-0.032	-0.078***	-0.081***
× t since 1970		(0.026)	(0.028)	(0.026)	(0.036)	(0.022)	(0.016)
Pop. Density (log)		-0.015	-0.030	-0.014	-0.009	-0.008	-0.013
× after 1970		(0.020)	(0.026)	(0.020)	(0.025)	(0.020)	(0.032)
Pop. Density (log)		-0.003	-0.003	-0.003	0.001	-0.008***	-0.008**
× t since 1970		(0.004)	(0.005)	(0.004)	(0.005)	(0.002)	(0.003)
Polluting Mfg. %			0.939				0.640
× after 1970			(0.586)				(0.454)
Polluting Mfg. %			-0.013				0.132
× t since 1970			(0.116)				(0.091)
DW-Nominate				0.031			0.004
× after 1970				(0.058)			(0.061)
DW-Nominate				-0.004			-0.004
× t since 1970				(0.015)			(0.011)
Midwest					0.010		-0.007
× after 1970					(0.038)		(0.034)
Midwest					0.001		0.016*
× t since 1970					(0.009)		(0.008)
South					0.115*		0.084
× after 1970					(0.063)		(0.055)
South					-0.006		0.014
× t since 1970					(0.014)		(0.010)
West					0.105		0.091
× after 1970					(0.067)		(0.064)
West					0.013		0.018
× t since 1970					(0.017)		(0.014)
NO_x Intensity (log)						0.071*	0.042
× after 1970						(0.039)	(0.028)
NO_x Intensity (log)						-0.043***	-0.050***
× t since 1970						(0.011)	(0.009)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.973	0.973	0.973	0.973	0.974	0.976	0.976

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Independent variables are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

a negligible share of SO_2 (2%). As a result, a substantial fraction of nitrogen oxides emissions may be abated through policies aimed at individual car owners. Table 2, columns (6) and (7) suggest that such policies are more effective in richer and more densely populated states. This pattern hints at the efficiency of federal regulation, since richer individuals and urban residents are likely to find it easiest to accelerate the decline in pollution by replacing their vehicles more rapidly or using them more sparingly. At the same time, the decline is faster in less informed states, consistent with the view that they provided lower fruit for the EPA to pick due to the inefficiency of their pre-existing local regulation, as predicted by Proposition 4.

3.4 Robustness

Tables 3 and 4 explore the robustness of our results, respectively for SO_2 and NO_x emissions, to different controls for the industrial structure of each state and for its energy consumption. Both the point estimates and the significance of the coefficient on newspaper circulation hardly change across specifications.

Column (1) in these tables replicates column (3) of the baseline Tables 1 and 2. Just as with polluting manufacturing industries, we might be concerned that the Clean Air Act affected primarily states that specialized in manufacturing as a whole, and that as a consequence had optimally adopted non-restrictive standards before its passage. Column (2) finds no evidence of such a pattern, which would imply a significant negative coefficient on the share of manufacturing. Analogously, column (3) shows that there is no differential effect due to specialization in power generation, the single most polluting sector (accounting for 56% of SO_2 and 23% of NO_x emissions, as opposed to 29% and 20%, respectively, for manufacturing).

Columns (4) to (6) control for total consumption of fossil fuels, relative to income, and for its breakdown by primary energy source. The main finding is that the Clean Air Act induced a differentially faster reduction of SO_2 emissions in states with a greater reliance of coal, and a faster reduction of NO_x in states with higher overall fuel consumption. These results are unsurprising, since combustion of coal is the principal source of sulfur dioxide, while total consumption of fossil fuels is the main determinant of nitrogen oxides emissions. As a consequence, there is a tight correspondence between Table 3, column (5), and Table 1, column (6); and analogously between Table 4, column (4), and Table 1, column (6).

Tables 5 and 6 expand the set of political control beyond the measure of ideology included in column (4) of the baseline Tables 1 and 2, which is replicated here as column (1). Our baseline results are again remarkably robust.

Column (2) controls for the share of white population. The concern arising from environmental justice model is that, absent uniform prescriptive standards, non-whites

Table 3 – SO_2 Emissions: Alternative Controls for Economic Conditions

Dependent Variable: SO_2 Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	-0.158	-0.251	-0.164	-0.073	-0.183	-0.257	-0.264
× after 1970	(0.774)	(0.805)	(0.835)	(0.845)	(0.837)	(0.827)	(0.818)
Newspaper Circ.	0.390**	0.419**	0.389*	0.391*	0.458**	0.447**	0.486**
× t since 1970	(0.185)	(0.179)	(0.188)	(0.190)	(0.187)	(0.199)	(0.186)
Income (log)	-0.535	-0.647	-0.756	-0.254	-0.205	-0.785	-0.710
× after 1970	(0.549)	(0.620)	(0.625)	(0.607)	(0.605)	(0.762)	(0.819)
Income (log)	-0.130	-0.135	-0.098	-0.089	-0.119	0.073	-0.006
× t since 1970	(0.091)	(0.092)	(0.079)	(0.095)	(0.091)	(0.121)	(0.096)
Density (log)	-0.099	-0.087	-0.047	-0.031	-0.045	-0.099	-0.162
× after 1970	(0.073)	(0.078)	(0.058)	(0.051)	(0.058)	(0.078)	(0.105)
Density (log)	0.017*	0.022**	0.011	0.011	0.019**	0.031***	0.034**
× t since 1970	(0.009)	(0.010)	(0.008)	(0.007)	(0.007)	(0.011)	(0.016)
Polluting Mfg. %	3.278*						1.598
× after 1970	(1.611)						(1.559)
Polluting Mfg. %	-0.441						-0.230
× t since 1970	(0.261)						(0.277)
Manufacturing %		0.719					0.571
× after 1970		(0.728)					(0.785)
Manufacturing %		-0.217					0.024
× t since 1970		(0.139)					(0.162)
Utilities %			0.952				-8.200
× after 1970			(9.497)				(10.469)
Utilities %			0.057				-0.216
× t since 1970			(1.957)				(1.985)
Fossil Fuels (log)				0.375**	0.321**	-0.200	-0.229
× after 1970				(0.153)	(0.135)	(0.329)	(0.414)
Fossil Fuels (log)				0.007	0.040	0.183*	0.162*
× t since 1970				(0.036)	(0.029)	(0.089)	(0.083)
Coal %					0.354		-0.004
× after 1970					(0.320)		(0.275)
Coal %					-0.215***		-0.181**
× t since 1970					(0.043)		(0.074)
Motor Gas. %						-3.280	-3.332
× after 1970						(1.982)	(2.320)
Motor Gas. %						1.000*	0.660
× t since 1970						(0.547)	(0.547)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.942	0.941	0.940	0.942	0.944	0.943	0.945

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = $10\%^{*}$, $5\%^{**}$, $1\%^{***}$ All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. Manufacturing %, Polluting Manufacturing %, and Utilities % are shares of state GDP. Polluting manufacturing industries are: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). Fossil Fuels is the log of consumption relative to income. Coal % and Motor Gasoline % are shares of fossil-fuel consumption.

Table 4 – NO_x Emissions: Alternative Controls for Economic Conditions

Dependent Variable: NO_x Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	0.120	0.102	0.163	0.139	0.130	0.161	0.237
× after 1970	(0.296)	(0.300)	(0.314)	(0.319)	(0.324)	(0.326)	(0.311)
Newspaper Circ.	0.190**	0.189**	0.189**	0.182***	0.187**	0.176**	0.182**
× t since 1970	(0.082)	(0.081)	(0.079)	(0.063)	(0.069)	(0.067)	(0.067)
Income (log)	0.087	0.044	-0.033	0.138	0.142	0.201	0.136
× after 1970	(0.177)	(0.179)	(0.165)	(0.195)	(0.199)	(0.204)	(0.185)
Income (log)	-0.006	-0.003	-0.003	-0.056*	-0.058*	-0.073	-0.088**
× t since 1970	(0.028)	(0.027)	(0.023)	(0.029)	(0.029)	(0.046)	(0.039)
Density (log)	-0.030	-0.023	-0.018	-0.011	-0.012	-0.003	-0.026
× after 1970	(0.026)	(0.025)	(0.019)	(0.019)	(0.020)	(0.023)	(0.029)
Density (log)	-0.003	-0.004	-0.003	-0.005*	-0.004	-0.007	-0.009*
× t since 1970	(0.005)	(0.004)	(0.004)	(0.002)	(0.003)	(0.005)	(0.005)
Polluting Mfg. %	0.939						0.961
× after 1970	(0.586)						(0.604)
Polluting Mfg. %	-0.013						0.284**
× t since 1970	(0.116)						(0.128)
Manufacturing %		0.140					0.020
× after 1970		(0.213)					(0.205)
Manufacturing %		0.009					-0.054
× t since 1970		(0.047)					(0.047)
Utilities %			-4.499				-7.437*
× after 1970			(3.117)				(3.815)
Utilities %			0.146				0.606
× t since 1970			(0.664)				(0.387)
Fossil Fuels (log)				0.086*	0.082*	0.154	0.126
× after 1970				(0.050)	(0.045)	(0.117)	(0.126)
Fossil Fuels (log)				-0.038**	-0.035**	-0.056	-0.075**
× t since 1970				(0.014)	(0.014)	(0.040)	(0.035)
Coal %					0.029		-0.066
× after 1970					(0.084)		(0.083)
Coal %					-0.015		-0.032
× t since 1970					(0.018)		(0.022)
Motor Gas. %						0.386	0.218
× after 1970						(0.577)	(0.660)
Motor Gas. %						-0.104	-0.146
× t since 1970						(0.190)	(0.171)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.973	0.973	0.973	0.975	0.975	0.975	0.976

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. Manufacturing %, Polluting Manufacturing %, and Utilities % are shares of state GDP. Polluting manufacturing industries are: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). Fossil Fuels is the log of consumption relative to income. Coal % and Motor Gasoline % are shares of fossil-fuel consumption.

Table 5 – SO_2 Emissions: Alternative Controls for Political Conditions

Dependent Variable: SO_2 Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	-0.092	1.707*	-0.132	0.320	0.272	-0.206	1.678
× after 1970	(0.836)	(0.966)	(0.783)	(0.750)	(0.780)	(0.848)	(1.221)
Newspaper Circ.	0.424**	0.484**	0.397*	0.375*	0.421*	0.381*	0.513**
× t since 1970	(0.191)	(0.228)	(0.197)	(0.203)	(0.205)	(0.185)	(0.235)
Income (log)	-0.776	-0.188	-0.727	-0.502	-0.641	-0.830	-0.452
× after 1970	(0.595)	(0.509)	(0.751)	(0.629)	(0.607)	(0.591)	(0.611)
Income (log)	-0.104	-0.069	-0.085	-0.106	-0.089	-0.068	-0.064
× t since 1970	(0.085)	(0.098)	(0.085)	(0.094)	(0.092)	(0.093)	(0.103)
Density (log)	-0.046	-0.136**	-0.052	-0.062	-0.057	-0.047	-0.140**
× after 1970	(0.056)	(0.065)	(0.054)	(0.057)	(0.057)	(0.055)	(0.065)
Density (log)	0.012	0.006	0.009	0.011	0.010	0.011	0.008
× t since 1970	(0.007)	(0.010)	(0.008)	(0.007)	(0.008)	(0.008)	(0.009)
DW-Nominate	0.066						-0.311
× after 1970	(0.190)						(0.206)
DW-Nominate	0.036						0.048
× t since 1970	(0.037)						(0.042)
White Pop. %		-2.741***					-3.789**
× after 1970		(0.948)					(1.418)
White Pop. %		-0.139					-0.065
× t since 1970		(0.156)					(0.180)
% R. in Office			-0.038				0.416
× after 1970			(0.253)				(0.262)
% R. in Office			-0.013				-0.041*
× t since 1970			(0.022)				(0.024)
% R. Votes				-0.749*			-0.434
× after 1970				(0.379)			(0.544)
% R. Votes				0.022			0.294**
× t since 1970				(0.102)			(0.135)
Elect. Evenness					-0.586		0.721
× after 1970					(0.357)		(0.475)
Elect. Evenness					-0.043		-0.290
× t since 1970					(0.130)		(0.173)
Sen. Class 1+3						0.047	-0.041
× after 1970						(0.130)	(0.100)
Sen. Class 1+3						0.006	0.012
× t since 1970						(0.019)	(0.021)
Sen. Class 2+3						-0.009	-0.126
× after 1970						(0.101)	(0.118)
Sen. Class 2+3						0.034	0.036
× t since 1970						(0.023)	(0.023)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.941	0.943	0.941	0.941	0.942	0.941	0.945

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. The DW-Nominate score is the average for the state's U.S. senators. The share of Republicans in office is computed considering the governor, the leaders of the two state legislatures, the two U.S. senators, and the majority of the state delegation to the U.S. House of Representatives. The share of Republican votes is the average in gubernatorial elections from 1968 to 1972. Election evenness is the average of $1-|\%\text{ R. Votes } -0.5|$ in gubernatorial elections from 1968 to 1972.

Table 6 – NO_x Emissions: Alternative Controls for Political Conditions

Dependent Variable: NO_x Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	0.150	0.427	0.188	0.235	0.273	0.111	0.478
× after 1970	(0.335)	(0.321)	(0.296)	(0.294)	(0.290)	(0.339)	(0.361)
Newspaper Circ.	0.187**	0.188*	0.192**	0.217**	0.204**	0.173**	0.152
× t since 1970	(0.085)	(0.099)	(0.082)	(0.091)	(0.086)	(0.072)	(0.089)
Income (log)	0.016	0.115	0.138	0.084	0.065	-0.006	0.156
× after 1970	(0.180)	(0.192)	(0.195)	(0.218)	(0.192)	(0.171)	(0.220)
Income (log)	-0.004	-0.006	-0.002	0.010	-0.001	-0.017	-0.003
× t since 1970	(0.026)	(0.026)	(0.027)	(0.020)	(0.030)	(0.025)	(0.020)
Density (log)	-0.014	-0.030	-0.026	-0.018	-0.018	-0.015	-0.030
× after 1970	(0.020)	(0.024)	(0.021)	(0.021)	(0.020)	(0.019)	(0.023)
Density (log)	-0.003	-0.003	-0.003	-0.004	-0.003	-0.003	-0.002
× t since 1970	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
DW-Nominate	0.031						0.118*
× after 1970	(0.058)						(0.068)
DW-Nominate	-0.004						0.000
× t since 1970	(0.015)						(0.018)
White Pop. %		-0.451					-0.136
× after 1970		(0.267)					(0.247)
White Pop. %		0.004					0.047
× t since 1970		(0.048)					(0.060)
% R. in Office			-0.112				-0.140
× after 1970			(0.069)				(0.084)
% R. in Office			-0.003				-0.004
× t since 1970			(0.011)				(0.013)
% R. Votes				-0.181			0.108
× after 1970				(0.151)			(0.160)
% R. Votes				-0.043			-0.099
× t since 1970				(0.040)			(0.091)
Elect. Evenness					-0.210		-0.200
× after 1970					(0.134)		(0.130)
Elect. Evenness					-0.019		0.078
× t since 1970					(0.040)		(0.105)
Sen. Class 1+3						0.009	0.009
× after 1970						(0.035)	(0.032)
Sen. Class 1+3						0.015*	0.013
× t since 1970						(0.008)	(0.008)
Sen. Class 2+3						-0.015	-0.016
× after 1970						(0.024)	(0.026)
Sen. Class 2+3						0.004	0.004
× t since 1970						(0.009)	(0.009)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.973	0.973	0.973	0.973	0.973	0.973	0.974

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = $10\%^{*}$, $5\%^{**}$, $1\%^{***}$ All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. The DW-Nominate score is the average for the state's U.S. senators. The share of Republicans in office is computed considering the governor, the leaders of the two state legislatures, the two U.S. senators, and the majority of the state delegation to the U.S. House of Representatives. The share of Republican votes is the average in gubernatorial elections from 1968 to 1972. Election evenness is the average of $1-|\%$ R. Votes $-0.5|$ in gubernatorial elections from 1968 to 1972.

face disproportionately more pollution due to their lower political clout. If so, the Clean Air Act should disproportionately reduce emissions in states with a lower percentage of white residents. We find instead that SO_2 emissions declined more rapidly after 1970 in states with a higher share of white population, while there is no differential effect for nitrogen oxides. The result is in line with Gray and Shadbegian's findings (2004). They study air and water pollution for 409 U.S. pulp and paper mills between 1985 and 1997, and conclude that pollution is in fact lower in minority areas.

Columns (3) and (4) control for partisanship, with the same rationale we discussed for ideology. We check if the impact of the Clean Air Act varies across states depending on their political leanings. Overall, we find no evidence that the EPA had a disproportionate impact in more Republican states, whether measured by vote shares in the gubernatorial election or by the number of state-wide elections won by the Republican party. All coefficients are insignificant for nitrogen oxides in Table 6. For sulfur dioxide, Table 5 finds conflicting results within column (7) as well as across columns (4) and (7). In any case, a measure of partisanship seems less appropriate than our baseline measure of ideology. First, we should not overlook that the creation of the EPA was proposed by a Republican president, Richard Nixon. Second, party affiliation around 1970 masked substantial ideological differences between the very conservative Southern Democrats, and the more liberal Democrats in the rest of the country.

Column (5) controls for a differential effect on states whose elections are more evenly split between the two parties. The absence of a significant effect is consistent with our theoretical model, in which the margin of victory in the election is determined by random shocks rather than by the underlying level of political accountability.

Column (6) considers the proximity of U.S. Senate elections. Table 6 finds that the decline in NO_x emissions is slower in states whose U.S. senators are furthest from re-election in 1971, i.e., who belong to Classes 1 and 3. This finding might suggest that federally mandated pollution abatement is popular with voters. However, we should not read too much into a coefficient that is not robust to the inclusion of additional controls in column (7), and whose significance is limited to one of the pollutants for no discernible economic reason.

Finally, Tables 7 and 8 consider alternative dependent variables. For each we report two regressions: one controlling for income and population density alone, the other including all the controls from our baseline Tables 1 and 2.

Our preferred measure of pollution is the emissions intensity of income, which has the advantage of capturing the environmental efficiency of production, and inherently controlling for the potential displacement of economic activity. However, measuring emissions per capita or emissions density over space may provide a more accurate estimate of the adverse environmental and health consequences of pollution, which depend on the concentration of pollutants and the number of people exposed to them. In every case

Table 7 – SO_2 Emissions: Alternative Dependent Variables

	SO_2/cap.	SO_2/cap.	SO_2/sq.m.	SO_2/sq.m.	SO_2/$	SO_2/$
Income (log) at *t*					-0.720*	-0.690**
					(0.416)	(0.328)
Pop. Den. (log) at *t*					-0.820	-0.525
					(1.132)	(1.270)
Newspaper Circ.	-0.096	0.352	-0.048	0.402	-0.073	0.391
× after 1970	(0.741)	(1.055)	(0.729)	(1.039)	(0.750)	(1.066)
Newspaper Circ.	0.392*	0.443***	0.374*	0.449***	0.377*	0.442***
× t since 1970	(0.194)	(0.144)	(0.200)	(0.156)	(0.196)	(0.150)
Income (log)	-0.834	-0.221	-0.884	-0.294	-0.856	-0.252
× after 1970	(0.580)	(0.579)	(0.568)	(0.558)	(0.578)	(0.565)
Income (log)	-0.076	-0.035	-0.099	-0.066	-0.101	-0.055
× t since 1970	(0.090)	(0.072)	(0.094)	(0.078)	(0.103)	(0.096)
Pop. Density (log)	-0.066	-0.074	-0.062	-0.066	-0.058	-0.063
× after 1970	(0.056)	(0.057)	(0.054)	(0.053)	(0.054)	(0.053)
Pop. Density (log)	0.007	-0.004	0.002	-0.009	0.004	-0.006
× t since 1970	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.007)
Polluting Mfg. %		1.948		1.901		1.973
× after 1970		(1.245)		(1.219)		(1.234)
Polluting Mfg. %		0.108		0.122		0.099
× t since 1970		(0.179)		(0.177)		(0.173)
DW-Nominate		0.015		0.049		0.028
× after 1970		(0.158)		(0.156)		(0.160)
DW-Nominate		-0.006		-0.004		-0.004
× t since 1970		(0.017)		(0.018)		(0.018)
Midwest		0.415*		0.414**		0.408*
× after 1970		(0.200)		(0.198)		(0.198)
Midwest		0.006		0.001		0.004
× t since 1970		(0.023)		(0.024)		(0.022)
South		0.559**		0.547**		0.551**
× after 1970		(0.210)		(0.208)		(0.209)
South		0.038		0.035		0.037
× t since 1970		(0.027)		(0.027)		(0.026)
West		0.383		0.380		0.390
× after 1970		(0.274)		(0.268)		(0.274)
West		0.009		0.012		0.008
× t since 1970		(0.034)		(0.035)		(0.035)
SO_2 Intensity (log)		0.036		0.037		0.035
× after 1970		(0.051)		(0.049)		(0.051)
SO_2 Intensity (log)		-0.050***		-0.050***		-0.050***
× t since 1970		(0.006)		(0.006)		(0.006)
Observations	1056	1056	1056	1056	1056	1056
R^2	0.936	0.945	0.974	0.977	0.941	0.949

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = $10\%^{*}$, $5\%^{**}$, $1\%^{***}$ All regressions include time f.e., state f.e., and state-specific linear time trends. Dependent variables are logarithms. Independent variables interacted with the 1970 break are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

Table 8 – NO_x Emissions: Alternative Dependent Variables

	NO_x/cap.	NO_x/cap.	NO_x/sq.m.	NO_x/sq.m.	NO_x/$	NO_x/$
Income (log) at t					-1.101***	-1.007***
					(0.160)	(0.166)
Pop. Den. (log) at t					-1.031***	-0.933***
					(0.313)	(0.324)
Newspaper Circ.	0.179	0.497	0.227	0.536	0.235	0.533
× after 1970	(0.236)	(0.350)	(0.234)	(0.336)	(0.230)	(0.334)
Newspaper Circ.	0.193***	0.179***	0.175***	0.189***	0.175***	0.188***
× t since 1970	(0.066)	(0.058)	(0.053)	(0.054)	(0.049)	(0.055)
Income (log)	-0.047	0.105	-0.097	0.020	-0.105	0.025
× after 1970	(0.151)	(0.197)	(0.129)	(0.166)	(0.124)	(0.168)
Income (log)	0.017	-0.054***	-0.006	-0.077***	-0.004	-0.075***
× t since 1970	(0.020)	(0.018)	(0.013)	(0.020)	(0.018)	(0.019)
Pop. Density (log)	-0.033	-0.037	-0.029	-0.032	-0.031	-0.032
× after 1970	(0.020)	(0.031)	(0.017)	(0.027)	(0.018)	(0.028)
Pop. Density (log)	-0.006*	-0.008**	-0.012***	-0.011***	-0.012***	-0.011***
× t since 1970	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Polluting Mfg. %		0.547		0.549		0.548
× after 1970		(0.464)		(0.473)		(0.477)
Polluting Mfg. %		0.142		0.137		0.137
× t since 1970		(0.105)		(0.106)		(0.106)
DW-Nominate		0.016		0.049		0.047
× after 1970		(0.052)		(0.047)		(0.046)
DW-Nominate		-0.006		-0.003		-0.003
× t since 1970		(0.010)		(0.010)		(0.010)
Midwest		0.019		0.022		0.022
× after 1970		(0.034)		(0.032)		(0.031)
Midwest		0.011		0.003		0.004
× t since 1970		(0.008)		(0.008)		(0.007)
South		0.094*		0.087		0.088
× after 1970		(0.051)		(0.052)		(0.052)
South		0.008		0.003		0.003
× t since 1970		(0.009)		(0.010)		(0.009)
West		0.067		0.066		0.066
× after 1970		(0.055)		(0.050)		(0.052)
West		0.024*		0.026*		0.026*
× t since 1970		(0.013)		(0.013)		(0.014)
NO_x Intensity (log)		0.033*		0.020		0.020
× after 1970		(0.019)		(0.014)		(0.019)
NO_x Intensity (log)		-0.036***		-0.028**		-0.028**
× t since 1970		(0.010)		(0.010)		(0.013)
Observations	1056	1056	1056	1056	1056	1056
R^2	0.971	0.973	0.995	0.996	0.977	0.979

Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Dependent variables are logarithms. Independent variables interacted with the 1970 break are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

our baseline results are fully confirmed. For nitrogen oxides, the estimates become, if anything, more precise.

In the last two columns we consider a fully general specification that leaves emission intensity of income on the left-hand side, but adds contemporaneous controls for income and population density to the right-hand side:

$$\log \frac{P_{i,t}}{Y_{i,t}} = \kappa \log \frac{Y_{i,t}}{L_{i,t}} + \lambda \log \frac{L_{i,t}}{T_i} + \alpha + \delta_t + \zeta_i + \theta_i t + d_t x_i' \beta + d_t x_i' \gamma \left(t - 1970\right) + \varepsilon_{i,t}, \quad (29)$$

where $P_{i,t}$ denotes tons of polluting emissions, $Y_{i,t}$ aggregate personal income in real dollars, $L_{i,t}$ population and T_i land area. This specification has the appeal of nesting all the alternative we previously considered, but the drawback of including potentially endogenous time-varying regressors. The robustness of our baseline results is confirmed yet again.

Moreover, we find suggestive evidence of an environmental Kuznets curve. The emissions intensity of income declines when income rises. In addition, nitrogen oxides are significantly lower in more densely populated states, suggesting that urban areas record lower emissions per unit of income with respect to rural areas, possibly as a result of the lower incidence of driving.

3.5 Discussion

Overall, our empirical evidence provides robust support for the theoretical prediction that the Clean Air Act induced a faster decline of pollution in states with lower newspaper circulation. This pattern can be plausibly interpreted according to the model of Section 2. The differential emissions abatement represents the welfare gains that the uninformed reap from centralization. Federal intervention solves a regulatory failure in states with low newspaper readership, whose local governments failed to invest efficiently in environmental regulation.

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

We obtain expenditure data 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 by states, counties, and cities. Since the first published report is for the fiscal year 1969, this source allows a rough comparison between one year before the 1970 Clean Air Act to one after: we consider a decade and compare data for the fiscal year 1969 and 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. Hence, the 1970 policy shift may have fostered convergence of state and local government expenditures in air pollution regulation, with less informed states gradually catching up to more informed ones.

Our second measure of regulatory input is 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 U.S. Senate (U.S. Senate, 1970). The document reports the standards adopted by states and local governments before 1970 on ten pollutants.[7] Counting the number of pollutants that each state had regulated prior to 1970, we find that states with above-average newspaper circulation had adopted four standards on average, while uninformed states had adopted three standards only. The hearings also report 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.[8] 50% of the states with high newspaper circulation had proposed or adopted emission standards for the two pollutants, as opposed to only 25% of the states with low newspaper circulation. This suggests that informed states were putting more effort in the development of air quality standards, while uninformed states were in greater need of federal intervention.

The benefits associated with environmental regulation are large and well documented. Chay and Greenstone (2005) estimate the hedonic value of the improvements in air quality induced by the Clean Air Act through their impact on housing prices. By using data on total suspended particulates, they find that better air quality caused a substantial increase in house prices, which amounted in the aggregate to $45 billion (in 2001 dollars). Bayer, Kehoane and Timmins (2009) provide much larger estimates of the benefits of emissions abatement. They model explicitly households' location choice, recognizing the costs of moving. Considering U.S. metropolitan areas in the 1990s, they estimate a marginal willingness to pay for a better air quality that is as much as seven times as large as Chay and Greenstone's (2005) estimate. Luechinger (2009) behavioral economics analysis also suggests that Chay and Greenstone's (2005) estimates are conservative. Comparing self-reported survey perceptions of life satisfaction and house prices in Germany from

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

[8] The 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.

1985 to 2003, he estimates that only 16% of the total willingness to pay for clean air is capitalized in the housing market. Finally, air pollution is shown to be positivelly related to infant mortality. Chay and Greenstone (2003) estimate that a reduction in suspended particulates by 1 $\mu g/m^3$ is associated to approximately 200 additional infants per year surviving to one year of age in the United States. Currie and Schmieder (2009) show that the abatement of toluene, cadmium and lead emissions that occurred in the late 1980s and 1990s has determined a reduction of more than 220 infants deaths in 2000.

On the other hand, a legitimate concern is that clean air may come at the cost of a deterioration in local economic conditions. Plants may choose to relocate away from tightly regulated areas, implying an inefficiency at least for a subset of states. At the local level, air quality regulation does affect industrial location and causes reductions in employment, investment and shipments (Henderson 1996; Becker and Henderson 2000). Tight environmental standards affect most sharply heavily polluting industries, which experience a reduction in employment, output, capital stock, and total factor productivity (Greenstone, 2002; Greenstone, List and Syverson, 2012).

Our finding that the differential effects of the Clean Air Act are not determined by industrial composition suggests a limited role for displacement. We test for displacement more directly by examining differences in differences for outcomes related to economic activity and industrial composition. In Table 9 we run the difference-in-differences specification from equation 28, replacing emissions on the left-hand side with eight variables that may have been affected, directly or indirectly, by abatement efforts mandated by the Clean Air Act. On the right-hand side, in addition to newspaper circulation, we include our fundamental controls: income and population density.

Most important, we find no significant differential effect of information on income or population density after 1970. Figures 4 and 5 depict the result graphically. After the passage of the Clean Air Act, states with lower newspaper circulation experienced a faster reduction in pollution, but neither slower income growth nor slower population growth. This finding lends support to the view that federal intervention helped the uninformed without harming them, neither by reducing their income in the short run, nor by reducing the appeal of the state and inducing outmigration in the longer run. The differential reduction in pollution was derived without any impact on the scale of economic activity.

There is some evidence that more informed states had a slower decline in the share of polluting manufacturing industries after 1970. This suggests that at least some displacement may have occurred, although Figure 6 raises the suspicion that the pattern may be driven by differential trends in the early 1960s. This could be the reason why the significant coefficient is on a break in the level of the series, rather than on its trend, where it would be expected due to the gradual phase-in of environmental regulation.[9] In

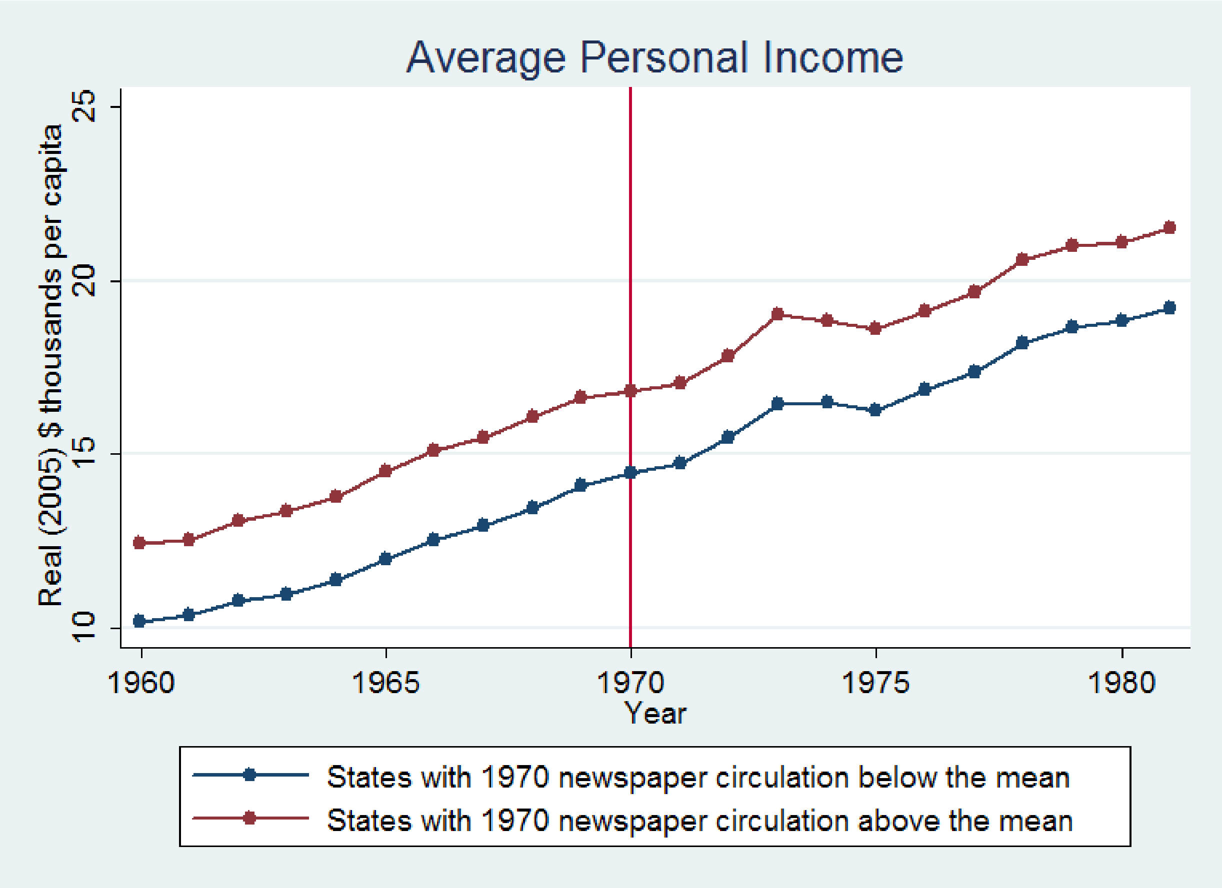
[9] Moreover, the statistical significance of this coefficient uniquely depends on the choice of two-way or

Table 9 – Channels of Pollution Abatement

	Income	Pop. Den.	Poll. %	Mfg. %	Util. %	F. Fuel	Coal %	Gas. %
Newspaper Circ.	0.059	0.048	0.020*	0.022	0.011	-0.147	0.223	-0.017
× after 1970	(0.172)	(0.037)	(0.010)	(0.034)	(0.007)	(0.274)	(0.150)	(0.036)
Newspaper Circ.	0.002	-0.018	0.000	-0.009	0.001	0.097*	0.041	-0.013
× t since 1970	(0.025)	(0.024)	(0.005)	(0.009)	(0.002)	(0.052)	(0.025)	(0.008)
Income (log)	-0.067	-0.050*	-0.009	-0.018	0.003	0.216	-0.031	-0.035
× after 1970	(0.058)	(0.028)	(0.006)	(0.010)	(0.003)	(0.175)	(0.093)	(0.029)
Income (log)	0.022**	-0.023**	0.001	0.012***	-0.000	-0.008	0.001	0.003
× t since 1970	(0.010)	(0.010)	(0.002)	(0.004)	(0.001)	(0.019)	(0.013)	(0.004)
Pop. Den. (log)	-0.018**	0.004	-0.000	-0.004**	0.001**	-0.004	-0.015	0.004
× after 1970	(0.007)	(0.003)	(0.001)	(0.002)	(0.000)	(0.018)	(0.010)	(0.003)
Pop. Den. (log)	-0.003**	-0.005***	0.000	0.000	0.000	-0.001	-0.003**	0.001**
× t since 1970	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.002)	(0.001)	(0.000)
Observations	1056	1056	912	912	912	1056	1056	1056
R^2	0.989	1.000	0.987	0.991	0.929	0.983	0.979	0.975

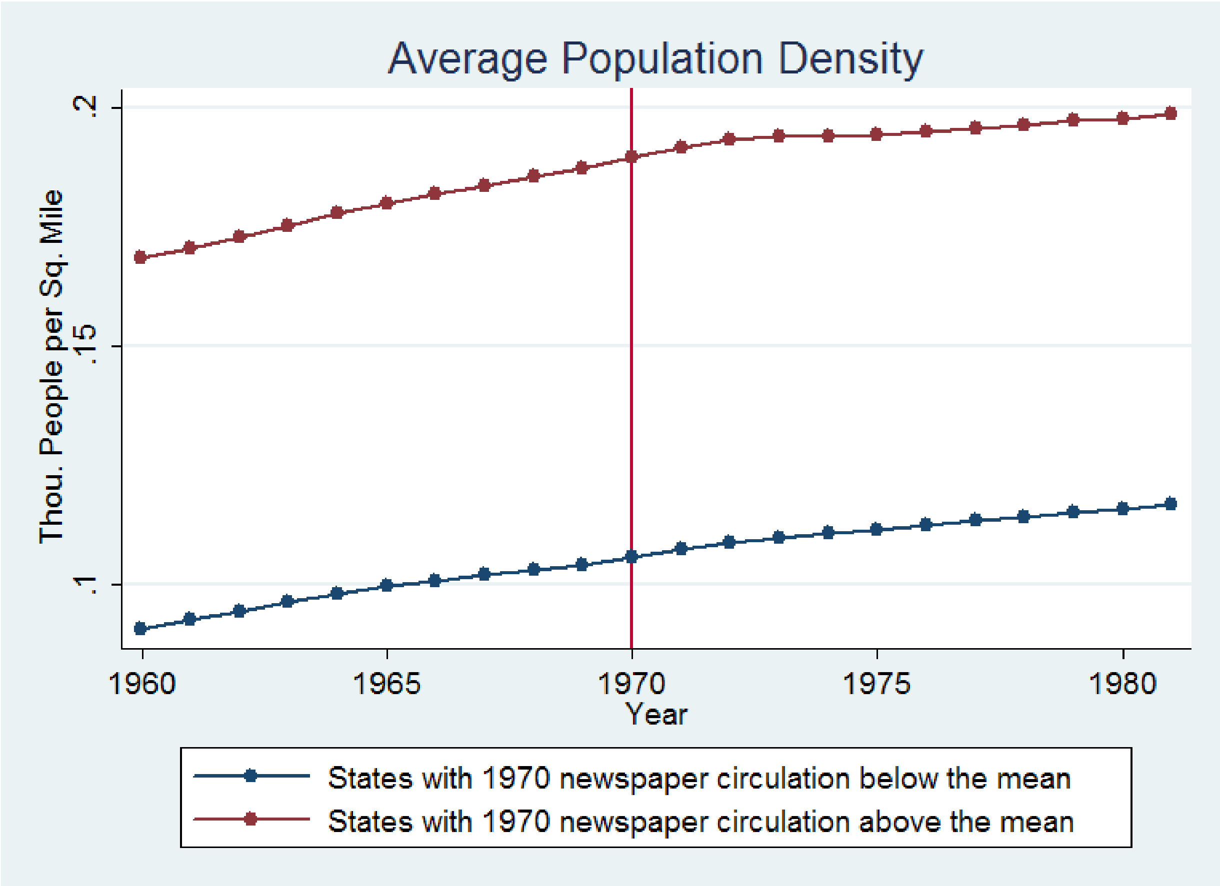
Notes: Robust standard errors in parentheses, with two-way clustering by state and by year; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Independent variables interacted with the 1970 break are 1970 values. Income and Population Density are logarithms. Manufacturing %, Polluting Manufacturing %, and Utilities % are shares of state GDP. Polluting manufacturing industries are: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). Fossil Fuels is the log of consumption relative to income. Coal % and Motor Gasoline % are shares of fossil-fuel consumption.

Figure 4 – Information and the Impact of the CAA on Personal Income



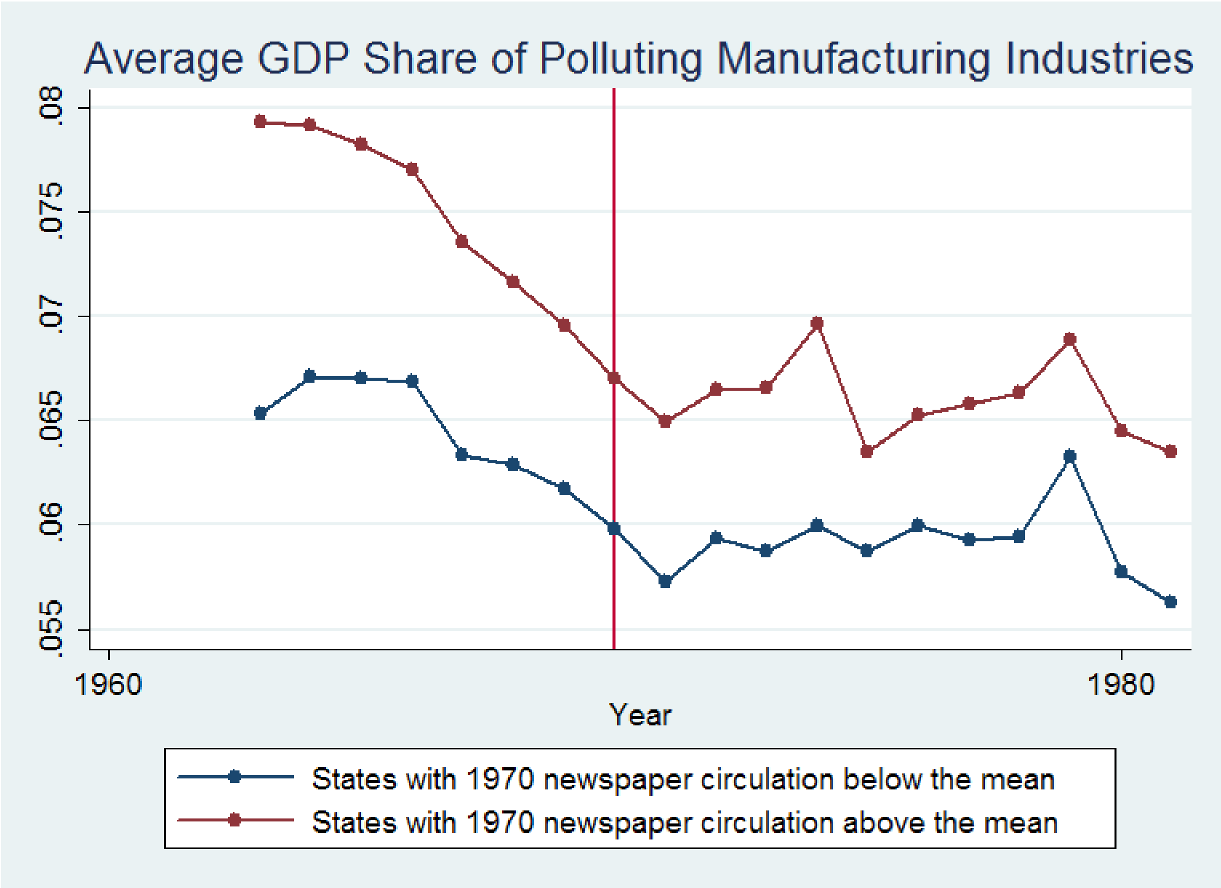
Sources: Newspaper circulation is from Gentzkow, Shapiro, and Sinkinson (2011), and personal income from the BEA Regional Economic Accounts.

Figure 5 – Information and the Impact of the CAA on Population Density



Sources: Newspaper circulation is from Gentzkow, Shapiro, and Sinkinson (2011), and population density from the U.S. Census.

Figure 6 – Information and the Impact of the CAA on Polluting Industries



Sources: Newspaper circulation is from Gentzkow, Shapiro, and Sinkinson (2011), and value added from the BEA Regional Economic Accounts. Polluting manufacturing industries are: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33).

any case, the economic impact of such displacement is quantitatively minimal. Considering two states whose 1970 newspaper circulation differed by one standard deviation (.05 copies per person), the rate of decline of pollutant emissions in the 1970s compared to the 1960s was higher in the less informed state by 1-2 percentage points per year. At the same time, its share of value added in polluting manufacturing industries would suffer a one-time decline by 0.1 percentage points.

There is no evidence of displacement for manufacturing as a whole, nor for the single most polluting sector, Electric, gas, and sanitary services (SIC 49, which is not a manufacturing industry). If the differential reduction in pollution in uninformed states after 1970 was accompanied by a change in the composition of economic activity, this change was of very limited size, and restricted in scope to highly polluting manufacturing industries.

Having ruled out changes in scale and substantial changes in composition, we conclude that the differential impact of the Clean Air Act operated through differential improvement in technique. Our baseline analysis showed that uninformed states achieved faster reductions in the emission intensity of income after 1970. Table 9 explores the channel in greater detail and finds evidence in particular of an impact on the fuel intensity of income. As Figure 7 also shows, after the creation of the EPA aggregate consumption of fossil fuels fell more quickly in states with lower newspaper circulation. In this case, too, the impact appears to be entirely on technique and not on composition, since there is no differential effect on the share of coal or of motor gasoline in total fuel consumption.

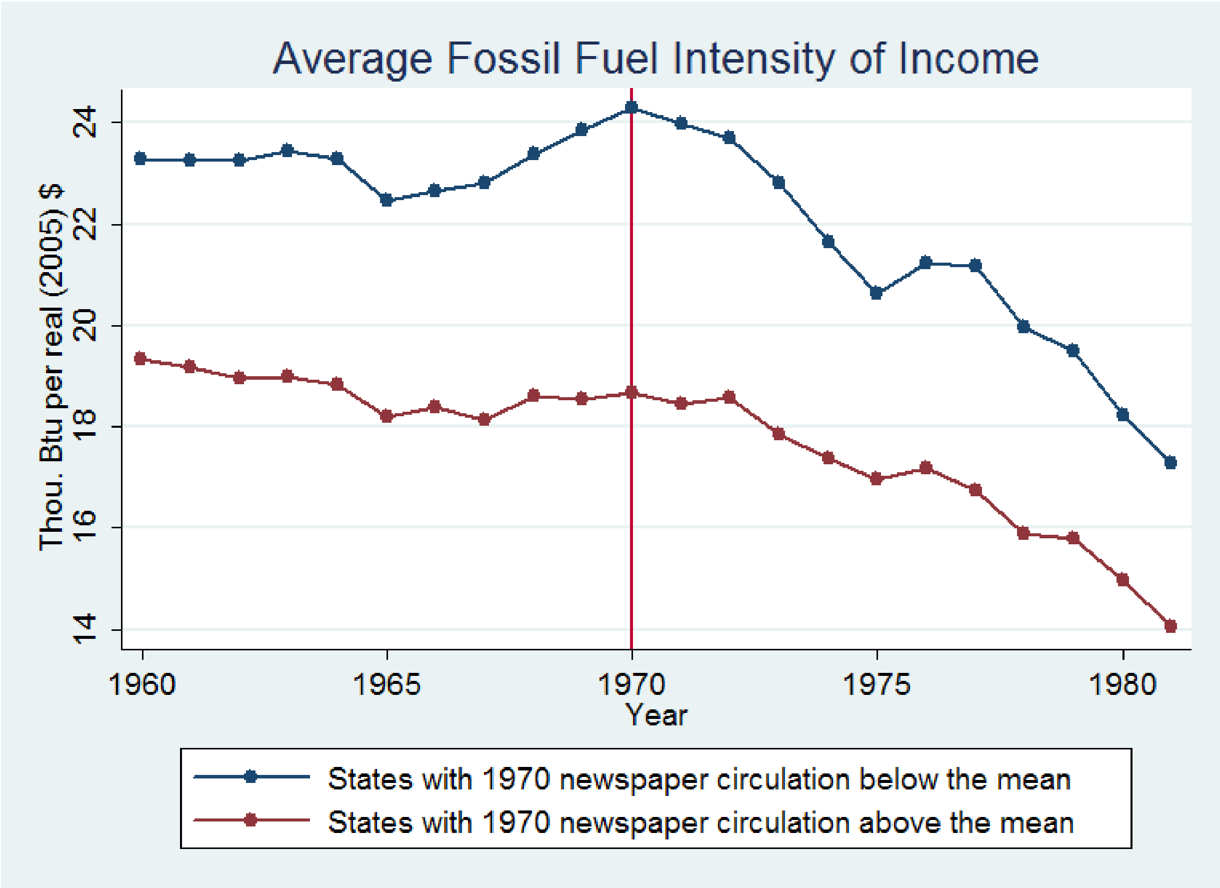
4 Conclusions

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 reduces rent extraction 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

one-way clustering. If we rely, conservatively, on the larger of the two standard errors, the null hypothesis of no differential impact cannot be rejected (Table A9).

Figure 7 – Information and the Impact of the CAA on Fuel Consumption



Sources: Newspaper circulation is from Gentzkow, Shapiro, and Sinkinson (2011), fossil fuel consumption from the EIA State Energy Data System, and personal income from the BEA Regional Economic Accounts.

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. When the central government provides public goods uniformly across the union, the benefits of centralization are monotone decreasing in voter information. We have tested this prediction by analyzing the differential impact of the 1970 Clean Air Act across the United States. Beforehand, environmental regulation was in the hands of state and local governments. 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 faster decline in pollution in states with lower newspaper readership.

The finding that centralization benefits the least informed regions hinges on uniform public goods provision, as in the case of national air quality standards for pollutant emissions. In our model, we have shown conversely that if the central politician can differentiate local public goods 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 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

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

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

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

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

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. Equilibrium with Many Regions

A.2..1 Decentralization

By Proposition 1, under decentralization rent extraction is

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\left(1 - \frac{L-1}{L}\sum_{p=1}^{P}\xi_p \alpha_p^l\right)\theta_l\right]^{-1}, \tag{A4}$$

the expected ability of a local politician is

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p^l \theta_l, \tag{A5}$$

and the relative shares of each local public good are

$$\beta_{l,p}^D = \frac{\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p^l}{1 - \frac{L-1}{L}\sum_{q=1}^{P}\xi_q \alpha_q^l}. \tag{A6}$$

Welfare in region l is

$$\mathbb{E}u_l^D = \log b + \sum_{p=1}^{P} \alpha_p^l \left\{ \begin{array}{l} \left(1-\xi_p\right)\left[\log\left(1-\rho_l^D\right)+\mathbb{E}\hat{\eta}_{l,p}^D+\log\beta_{l,p}^D\right] \\ +\frac{\xi_p}{L}\sum_{m=1}^{L}\left[\log\left(1-\rho_m^D\right)+\mathbb{E}\hat{\eta}_{m,p}^D+\log\beta_{m,p}^D\right] \end{array} \right\}, \tag{A7}$$

and aggregate welfare is

$$W^D = \log b + \frac{1}{L}\sum_{l=1}^{L}\sum_{p=1}^{P}\left[\left(1-\xi_p\right)\alpha_p^l+\xi_p\bar{\alpha}_p\right]\left[\log\left(1-\rho_l^D\right)+\mathbb{E}\hat{\eta}_{l,p}^D+\log\beta_{l,p}^D\right]. \tag{A8}$$

for

$$\bar{\alpha}_p = \frac{1}{L}\sum_{l=1}^{L}\alpha_p^l. \tag{A9}$$

A.2..2 Centralization

By Proposition 1, under centralization rent extraction is

$$\rho^C = \left(1+2\tilde{\delta}\phi\bar{\theta}\right)^{-1} \text{ for } \bar{\theta} = \frac{1}{L}\sum_{l=1}^{L}\theta_l, \tag{A10}$$

the expected ability of a central politician is

$$\mathbb{E}\hat{\eta}_p^C = \frac{\phi\sigma^2}{L}\sum_{l=1}^{L}\theta_l\alpha_p^l, \tag{A11}$$

and the relative shares of each local public good are

$$\beta_p^C = \frac{1}{L^2}\sum_{l=1}^{L}\frac{\theta_l}{\bar{\theta}}\alpha_p^l \text{ for } p\in\mathcal{U} \tag{A12}$$

and

$$\beta_{l,p}^C = \frac{1}{L}\left[\left(1-\xi_p\right)\frac{\theta_l}{\bar{\theta}}\alpha_p^l+\frac{\xi_p}{L}\sum_{m=1}^{L}\frac{\theta_m}{\bar{\theta}}\alpha_p^m\right] \text{ for } p\in\mathcal{D}. \tag{A13}$$

Welfare in region l is

$$\begin{aligned}\mathbb{E}u_l = \log\left(Lb\right)+\log\left(1-\rho^C\right)+\sum_{p=1}^{P}\alpha_p^l\mathbb{E}\hat{\eta}_p^C+\sum_{p\in\mathcal{U}}\alpha_p^l\log\beta_p^C \\ +\sum_{p\in\mathcal{D}}\alpha_p^l\left[\left(1-\xi_p\right)\log\beta_{l,p}^C+\frac{\xi_p}{L}\sum_{m=1}^{L}\log\beta_{m,p}^C\right]\end{aligned} \tag{A14}$$

and aggregate welfare is

$$W^C = \log(Lb) + \log\left(1-\rho^C\right) + \sum_{p=1}^{P} \bar{\alpha}_p \mathbb{E}\hat{\eta}_p^C + \sum_{p\in\mathcal{U}} \bar{\alpha}_p \log \beta_p^C$$
$$+ \frac{1}{L}\sum_{l=1}^{L}\sum_{p\in\mathcal{D}} \alpha_p^l \left[\left(1-\xi_p\right)\log\beta_{l,p}^C + \frac{\xi_p}{L}\sum_{m=1}^{L}\log\beta_{m,p}^C\right]. \quad \text{(A15)}$$

A.3. Proof of Proposition 2

Aggregate rent extraction is lower under centralization if and only if

$$\rho^C \leq \frac{1}{L}\sum_{l=1}^{L}\rho_l^D, \quad \text{(A16)}$$

which can be written

$$f\left(\frac{1}{L}\sum_{l=1}^{L}\theta_l\right) \leq \frac{1}{L}\sum_{l=1}^{L} f\left(\left(1-\frac{L-1}{L}\sum_{p=1}^{P}\xi_p\alpha_p^l\right)\theta_l\right) \quad \text{(A17)}$$

for

$$f(x) \equiv \frac{1}{1+2\tilde{\delta}\phi x} \quad \text{(A18)}$$

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

$$f'(x) = -\frac{2\tilde{\delta}\phi}{\left(1+2\tilde{\delta}\phi x\right)^2} < 0 \text{ and } f''(x) = \frac{\left(2\tilde{\delta}\phi\right)^2}{\left(1+2\tilde{\delta}\phi x\right)^3} > 0. \quad \text{(A19)}$$

Thus

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

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. Ability

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

$$\mathbb{E}\hat{\eta}_p^C \geq \frac{1}{L}\sum_{l=1}^{L}\mathbb{E}\hat{\eta}_{l,p}^D \iff \xi_p \geq 0, \quad \text{(A21)}$$

with joint strict inequality.

A.4. Proof of Proposition 3

A.4..1 The Decentralization Theorem

Suppose that all public goods provided by the central government are subject to the uniformity constraint ($\mathcal{D} = \varnothing$) and that information is homogeneous across regions ($\theta_l = \theta$ for all l). Then under centralization

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \ \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\theta\bar{\alpha}_p, \text{ and } \beta_p^C = \frac{1}{L}\bar{\alpha}_p. \tag{A22}$$

while under decentralization

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\theta\left(1 - \frac{L-1}{L}\sum_{p=1}^{P}\xi_p\alpha_p^l\right)\right]^{-1}, \tag{A23}$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p^l, \tag{A24}$$

and

$$\beta_{l,p}^D = \left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p^l. \tag{A25}$$

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

$$\rho^C = \rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \ \mathbb{E}\hat{\eta}_p^C = \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\alpha_p \text{ and } L\beta_p^C = \beta_{l,p}^D = \alpha_p. \tag{A26}$$

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

$$\rho^C = \left(1 + 2\tilde{\delta}\phi\theta\right)^{-1}, \ \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\theta\alpha_p, \text{ and } L\beta_p^C = \alpha_p, \tag{A27}$$

and

$$W^C = \log b + \log\left(1 - \rho^C\right) + \sum_{p=1}^{P}\alpha_p\left(\mathbb{E}\hat{\eta}_p^C + \log\alpha_p\right). \tag{A28}$$

Under decentralization

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\theta\left(1 - \frac{L-1}{L}\sum_{p=1}^{P}\xi_p\alpha_p\right)\right]^{-1} > \rho^C, \tag{A29}$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p < \mathbb{E}\hat{\eta}_p^C, \tag{A30}$$

$$\beta_{l,p}^D = \frac{\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p}{1 - \frac{L-1}{L}\sum_{q=1}^{P}\xi_q\alpha_q}, \tag{A31}$$

and

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

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 ($\xi_p = 0$ for all p, while $\alpha_p^l \neq \alpha_p^m$ for some $l \neq m$ and p), then under centralization

$$W^C = \log b + \log\left(1-\rho^C\right) + \frac{1}{L}\sum_{l=1}^{L}\sum_{p=1}^{P} \alpha_p^l \left[\mathbb{E}\hat{\eta}_p^C + \log\left(L\beta_p^C\right)\right], \tag{A33}$$

while under decentralization

$$\rho_l^D = \left(1+2\tilde{\delta}\phi\theta\right)^{-1} = \rho^C,\ \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\theta\alpha_p^l,\ \beta_{l,p}^D = \alpha_p^l \tag{A34}$$

and

$$W^D = \log b + \log\left(1-\rho_l^D\right) + \frac{1}{L}\sum_{l=1}^{L}\sum_{p=1}^{P} \alpha_p^l \left(\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D\right). \tag{A35}$$

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

$$\frac{1}{L}\sum_{l=1}^{L}\sum_{p=1}^{P} \alpha_p^l \mathbb{E}\hat{\eta}_{l,p}^D > \frac{1}{L}\sum_{l=1}^{L}\sum_{p=1}^{P} \alpha_p^l \mathbb{E}\hat{\eta}_p^C \tag{A36}$$

because for each public good p

$$\frac{1}{L}\sum_{l=1}^{L}\left(\alpha_p^l\right)^2 > \left(\frac{1}{L}\sum_{l=1}^{L}\alpha_p^l\right)^2 \tag{A37}$$

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

A.4..2 Efficient Centralization

Suppose that all public goods provided by the central government are subject to the uniformity constraint ($\mathcal{D} = \varnothing$) and that preferences are homogeneous ($\alpha_p^l = \alpha_p$ for all l). Then under centralization

$$\rho^C = \left(1+2\tilde{\delta}\phi\bar{\theta}\right)^{-1},\ \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\alpha_p\bar{\theta},\ \beta_p^C = \alpha_p, \tag{A38}$$

and

$$W^C = \log b + \log\left(1-\rho^C\right) + \sum_{p=1}^{P} \alpha_p \left(\mathbb{E}\hat{\eta}_p^C + \log \beta_p^C\right). \tag{A39}$$

Under decentralization

$$\rho_l^D = \left[1 + 2\tilde{\delta}\phi\left(1 - \frac{L-1}{L}\sum_{p=1}^{P}\xi_p\alpha_p\right)\theta_l\right]^{-1}, \tag{A40}$$

$$\mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p\theta_l, \tag{A41}$$

$$\beta_{l,p}^D = \frac{\left(1 - \frac{L-1}{L}\xi_p\right)\alpha_p}{1 - \frac{L-1}{L}\sum_{q=1}^{P}\xi_q\alpha_q}, \tag{A42}$$

and

$$W^D = \frac{1}{L}\sum_{l=1}^{L}\left[\log b + \log\left(1 - \rho_l^D\right) + \sum_{p=1}^{P}\alpha_p\left(\mathbb{E}\hat{\eta}_{l,p}^D + \log\beta_{l,p}^D\right)\right]. \tag{A43}$$

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

$$\log\left(1 - \rho^C\right) > \frac{1}{L}\sum_{l=1}^{L}\log\left(1 - \rho_l^D\right) \tag{A44}$$

which can be written

$$f\left(\frac{1}{L}\sum_{l=1}^{L}\theta_l\right) > \frac{1}{L}\sum_{l=1}^{L}f\left(\left(1 - \frac{L-1}{L}\sum_{p=1}^{P}\xi_p\alpha_p\right)\theta_l\right) \tag{A45}$$

for

$$f(x) \equiv \log x - \log\left(1 + 2\tilde{\delta}\phi x\right) \tag{A46}$$

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

$$f'(x) = \frac{1}{x\left(1 + 2\tilde{\delta}\phi x\right)} > 0 \text{ and } f''(x) = -\frac{1 + 4\tilde{\delta}\phi x}{\left[x\left(1 + 2\tilde{\delta}\phi x\right)\right]^2} < 0. \tag{A47}$$

Thus

$$f\left(\frac{1}{L}\sum_{l=1}^{L}\theta_l\right) \geq \frac{1}{L}\sum_{l=1}^{L}f(\theta_l) \geq \frac{1}{L}\sum_{l=1}^{L}f\left(\left(1 - \frac{L-1}{L}\sum_{p=1}^{P}\xi_p\alpha_p\right)\theta_l\right). \tag{A48}$$

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 all public goods provided by the central government are subject to the uniformity constraint ($\mathcal{D} = \varnothing$), and that preferences are homogeneous across regions ($\alpha_p^l = \alpha_p$ for all l). Then public goods provision in region l is

$$g_{l,p,t}^D = x_{p,t}^D \exp\left(\hat{\eta}_{l,p}^D\right) = \beta_{l,p}^D b^D \left(1 - \rho_l^D\right) \exp\left(\hat{\eta}_{l,p}^D\right) \tag{A49}$$

under decentralization, and

$$g_{l,p,t}^C = \frac{1}{L} x_{p,t}^C \exp\left(\hat{\eta}_p^C\right) = \beta_p^C b^D \left(1 - \rho^C\right) \exp\left(\hat{\eta}_p^C\right) \tag{A50}$$

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 $\partial \mathbb{E} \exp\left(\hat{\eta}_{l,p}^D\right) / \partial \theta_l > 0$. Since $g_{l,p,t}^C$ is identical for all regions, it follows that

$$\mathbb{E}\left(g_{l,p}^C - g_{l,p}^D\right) > \mathbb{E}\left(g_{m,p}^C - g_{m,p}^D\right) \Leftrightarrow \mathbb{E} g_{l,p}^D < \mathbb{E} g_{m,p}^D \Leftrightarrow \theta_l < \theta_m. \tag{A51}$$

Similarly for residents' welfare

$$\begin{aligned} &\mathbb{E}\left(u_l^C - u_l^D\right) > \mathbb{E}\left(u_m^C - u_m^D\right) \Leftrightarrow \mathbb{E} u_l^D < \mathbb{E} u_m^D \Leftrightarrow \\ &\quad \sum_{p=1}^{P} \alpha_p \left(1 - \xi_p\right) \left[\log\left(1 - \rho_l^D\right) + \mathbb{E}\hat{\eta}_{l,p}^D\right] < \sum_{p=1}^{P} \alpha_p \left(1 - \xi_p\right) \left[\log\left(1 - \rho_l^D\right) + \mathbb{E}\hat{\eta}_{l,p}^D\right] \Leftrightarrow \\ &\quad \theta_l < \theta_m \end{aligned} \tag{A52}$$

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

$$\rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta_l\right)^{-1}, \; \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\alpha_p\theta_l, \; \beta_{l,p}^D = \alpha_p, \tag{A53}$$

and

$$\mathbb{E} u_l^D = \log b^D + \log\left(1 - \rho_l^D\right) + \sum_{p=1}^{P} \alpha_p^l \left(\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D\right). \tag{A54}$$

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

A.6. Proof of Proposition 5

Suppose that preferences are homogeneous, information is heterogeneous, and there are no externalities ($\alpha_p^l = \alpha_p$for all l, $\theta_l \neq \theta_m$ for some $l \neq m$, and $\xi_p = 0$ for all p). Under decentralization, welfare in region l is

$$\mathbb{E}u_l^D = \log b + \log\left(1 - \rho_l^D\right) + \sum_{p=1}^{P} \alpha_p \left(\mathbb{E}\hat{\eta}_{l,p}^D + \log \beta_{l,p}^D\right), \tag{A55}$$

for

$$\rho_l^D = \left(1 + 2\tilde{\delta}\phi\theta_l\right)^{-1}, \ \mathbb{E}\hat{\eta}_{l,p}^D = \phi\sigma^2\alpha_p\theta_l, \text{ and } \beta_{l,p}^D = \alpha_p. \tag{A56}$$

Aggregate welfare is

$$W^D = \log b + \sum_{p=1}^{P} \alpha_p \log \alpha_p + \phi\sigma^2\bar{\theta}\sum_{p=1}^{P}\alpha_p^2 + \frac{1}{L}\sum_{l=1}^{L}\log\frac{2\tilde{\delta}\phi\theta_l}{1+2\tilde{\delta}\phi\theta_l} \tag{A57}$$

Under centralization, welfare in region l is

$$\mathbb{E}u_l = \log(Lb) + \log\left(1-\rho^C\right) + \sum_{p=1}^{P}\alpha_p\mathbb{E}\hat{\eta}_p^C + \sum_{p\in\mathcal{U}}\alpha_p\log\beta_p^C + \sum_{p\in\mathcal{D}}\alpha_p\log\beta_{l,p}^C \tag{A58}$$

for

$$\rho^C = \left(1+2\tilde{\delta}\phi\bar{\theta}\right)^{-1}, \ \mathbb{E}\hat{\eta}_p^C = \phi\sigma^2\alpha_p\bar{\theta}, \ \beta_p^C = \frac{\alpha_p}{L} \text{ for } p\in\mathcal{U}, \ \beta_{l,p}^C = \frac{\alpha_p}{L}\frac{\theta_l}{\bar{\theta}} \text{ for } p\in\mathcal{D}. \tag{A59}$$

Aggregate welfare is

$$W^C = \log b + \sum_{p=1}^{P}\alpha_p\log\alpha_p + \phi\sigma^2\bar{\theta}\sum_{p=1}^{P}\alpha_p^2 + \log\frac{2\tilde{\delta}\phi\bar{\theta}}{1+2\tilde{\delta}\phi\bar{\theta}} + (1-\alpha_{\mathcal{U}})\left(\frac{1}{L}\sum_{l=1}^{L}\log\theta_l - \log\bar{\theta}\right). \tag{A60}$$

Then aggregate welfare is greater under centralization if

$$\alpha_{\mathcal{U}} \geq \frac{\log\left(1+2\tilde{\delta}\phi\bar{\theta}\right) - \frac{1}{L}\sum_{l=1}^{L}\log\left(1+2\tilde{\delta}\phi\theta_l\right)}{\log\bar{\theta} - \frac{1}{L}\sum_{l=1}^{L}\log\theta_l} \equiv \bar{\alpha}_{\mathcal{U}} \in (0,1). \tag{A61}$$

Region l gains from centralization if

$$\log\frac{1+2\tilde{\delta}\phi\theta_l}{1+2\tilde{\delta}\phi\bar{\theta}} - \alpha_{\mathcal{U}}\log\frac{\theta_l}{\bar{\theta}} \geq \phi\sigma^2\sum_{p=1}^{P}\alpha_p^2\left(\theta_l - \bar{\theta}\right). \tag{A62}$$

The function

$$f(x) \equiv \log\left(1+2\tilde{\delta}\phi x\right) - \alpha_{\mathcal{U}}\log x \tag{A63}$$

has a unique minimum

$$f'(x) = \frac{2\tilde{\delta}\phi}{1+2\tilde{\delta}\phi x} - \frac{\alpha_{\mathcal{U}}}{x} \geq 0 \Leftrightarrow x \geq \frac{\alpha_{\mathcal{U}}}{2\tilde{\delta}\phi(1-\alpha_{\mathcal{U}})}, \tag{A64}$$

and convexity

$$f''(x) = -\left(\frac{2\tilde{\delta}\phi}{1+2\tilde{\delta}\phi x}\right)^2 + \frac{\alpha_{\mathcal{U}}}{x^2} > 0 \Leftrightarrow \alpha_{\mathcal{U}} > \left(\frac{2\tilde{\delta}\phi x}{1+2\tilde{\delta}\phi x}\right)^2 \tag{A65}$$

Hence if

$$\alpha_{\mathcal{U}} = \frac{2\tilde{\delta}\phi\bar{\theta}}{1+2\tilde{\delta}\phi\bar{\theta}} = 1-\rho^C \tag{A66}$$

the left-hand side of condition A62 is a convex function of $\theta_l \in [0,1]$ with minimum at $\theta_l = \bar{\theta}$. This also proves that

$$\bar{\alpha}_{\mathcal{U}} < 1-\rho^C. \tag{A67}$$

Moreover, let

$$m = \arg\min_{l=1,\ldots,L} \left\{\theta_l : \theta_l > \bar{\theta}\right\}. \tag{A68}$$

Then for

$$\sigma^2 \leq \frac{1}{\phi\sum_{p=1}^{P}\alpha_p^2\left(\theta_m-\bar{\theta}\right)}\left\{\log\frac{1+2\tilde{\delta}\phi\theta_m}{1+2\tilde{\delta}\phi\bar{\theta}} - \frac{2\tilde{\delta}\phi\bar{\theta}}{1+2\tilde{\delta}\phi\bar{\theta}}\log\frac{\theta_m}{\bar{\theta}}\right\} \equiv \bar{\sigma}^2 > 0, \tag{A69}$$

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] Adserà, Alícia, Carles Boix, and Mark Payne. 2003. “Are You Being Served? Political Accountability and the Quality of Government.” *Journal of Law, Economics, and Organization* 19(2): 445–90

[2] Alesina, Alberto, Ignazio Angeloni, and Federico Etro. 2005. “International Unions.” *American Economic Review* 95(3): 602–15.

[3] Alesina, Alberto, Ignazio Angeloni, and Ludger Schuknecht. 2005. “What Does the European Union Do?” *Public Choice* 123(3): 275–319.

[4] Alesina, Alberto, and Enrico Spolaore. 1997. “On the Number and Size of Nations.” *Quarterly Journal of Economics* 112(4): 1027–56.

[5] Bardhan, Pranab, and Dilip Mookherjee. 2000. “Capture and Governance at Local and National Levels.” *American Economic Review* 90(2): 135–9.

[6] ——. 2006. “Decentralization, Corruption, and Government Accountability.” In *International Handbook on the Economics of Corruption*, edited by Susan Rose-Ackerman, 161–88. Cheltenham: Edward Elgar.

[7] Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. “Migration and hedonic valuation: The case of air quality.” *Journal of Environmental Economics and Management* 58(1): 1–14.

[8] Becker, Randy and Vernon Henderson. 2000. “Effects of Air Quality Regulations on Polluting Industries.” *Journal of Political Economy* 108(2): 379–421.

[9] 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.

[10] Besley, Timothy, and Anne Case. 1995. “Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition.” *American Economic Review* 85(1): 25–45.

[11] 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.

[12] Besley, Timothy, and Michael Smart. 2007. “Fiscal Restraints and Voter Welfare.” *Journal of Public Economics* 91: 755–73.

[13] Blanchard, Olivier, and Andrei Shleifer. 2001. “Federalism With and Without Political Centralization: China Versus Russia.” *IMF Staff Papers* 48(4): 171–9.

[14] 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.

[15] Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. “Robust Inference With Multiway Clustering.” *Journal of Business & Economic Statistics* 29(2): 238–49.

[16] Campa, Pamela. 2013. "Press and Leaks: Do Newspapers Reduce Toxic Emissions?" Mimeo, Stockholm University.

[17] Charron, Nicholas, Lewis Dijkstra, and Victor Lapuente. 2013. "Regional Governance Matters: A Study on Regional Variation in Quality of Government within the EU." *Regional Studies*, forthcoming.

[18] 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.

[19] ——. 2005. "Does Air Quality Matter: Evidence from the Housing Market." *Journal of Political Economy.* 116: 376-424.

[20] Currie, Janet and Schmieder, Johannes. 2009. "Fetal Exposures to Toxic Releases and Infant Health." *American Economic Review Papers and Proceedings* 99(2): 177–183.

[21] 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.

[22] ——. 2007. "The Determinants of Corruption in Italy: Regional Panel Data Analysis." *European Journal of Political Economy* 23: 379–96.

[23] Dincecco, Mark. 2011. *Political Transformations and Public Finances.* Cambridge, UK: Cambridge University Press.

[24] Dincecco, Mark, and Gabriel Katz. 2012. "State Capacity and Long-Run Performance." Mimeo, IMT Lucca.

[25] Durante, Ruben, Giovanna Labartino, and Roberto Perotti. 2011. "Academic Dynasties: Decentralization and Familism in the Italian Academia." NBER Working Paper No 17572.

[26] 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.

[27] 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.

[28] 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.

[29] Gennaioli, Nicola, and Ilia Rainer. 2007a. "The Modern Impact of Precolonial Centralization in Africa." *Journal of Economic Growth* 12(3): 185–234.

[30] ——. 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.

[31] Gennaioli, Nicola, and Hans-Joachim Voth. 2011. "State Capacity and Military Conflict." CEPR Discussion Paper No. 8699.

[32] Gentzkow, Matthew. 2006. "Television and voter turnout." *Quarterly Journal of Economics* 121(3): 931–72.

[33] Gentzkow, Matthew, Jesse Shapiro, and Michael Sinkinson. 2011. "The Effect of Newspaper Entry and Exit on Electoral Politics." *American Economic Review* 101: 2980–3018.

[34] Glaeser, Edward L., Giacomo A. M. Ponzetto, and Jesse M. Shapiro. 2005. "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values." *Quarterly Journal of Economics* 120(4): 1283–330.

[35] 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.

[36] Graber, Doris A. 1984. *Processing the News: How People Tame the Information Tide.* New York, NY: Longman.

[37] Gray, Wayne B. and Ronald J. Shadbegian. 2004. "'Optimal' Pollution Abatement: Whose Benefits Matter, and How Much?" *Journal of Environmental Economics and Management* 47(3): 510–34.

[38] 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.

[39] ——. 2004. "Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?" *Journal of Environmental Economics and Management.* 47: 585-611.

[40] Greenstone, Michael, John List, and Chad Syverson. 2012 "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing." Mimeo, MIT.

[41] Henderson, J. Vernon. 1996. "Effects of air quality regulation." *American Economic Review* 86: 789–813.

[42] Holmström, Bengt. [1982] 1999. "Managerial Incentive Problems: A Dynamic Perspective." *Review of Economic Studies* 66(1): 169–82.

[43] Kahn, Matthew E. 1996. "New evidence on trends in vehicle emissions." *Rand Journal of Economics* 27(1): 183–96.

[44] ——. 1997. "Particulate Pollution Trends in the United States." *Regional Science and Urban Economics* 27(1): 87–107.

[45] Lindbeck, Assar, and Jörgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice* 52: 273–97.

[46] List, John, and Craig Gallet. 1999. "The Environmental Kuznets Curve: Does One Size Fit All?" *Ecological Economics* 31: 409–423.

[47] List, John, and Shelby Gerking. 2000. "Regulatory Federalism and Environmental Protection in the United States." *Journal of Regional Sciences.* 40: 453-471.

[48] List, John, and Daniel Sturm. 2006. "How Elections Matter: Theory and Evidence from Environmental Policy." *Quarterly Journal of Economics.* 121: 1249-1281.

[49] Lockwood, Ben. 2002. "Distributive Politics and the Costs of Centralization." *Review of Economic Studies* 69: 313–37.

[50] ——. 2006. "The Political Economy of Decentralization." In *Handbook of Fiscal Federalism*, edited by Ehtisham Ahmad and Giorgio Brosio, 33–60. Cheltenham: Edward Elgar.

[51] Luechinger, Simon. 2009. "Valuing Air Quality Using the Life Satisfaction Approach." *Economic Journal* 119(536): 482–515.

[52] Nelson, Randy A, Tom Tietenberg, and Michael R Donihue. 1993. "Differential Environmental Regulation: Effects on Electric Utility Capital Turnover and Emissions." *Review of Economics and Statistics* 75 (2): 368–73.

[53] Oates, Wallace E. 1972. *Fiscal Federalism.* New York, NY: Harcourt Brace Jovanovich.

[54] ——. 1999. "An Essay on Fiscal Federalism." *Journal of Economic Literature* 37: 1120–49.

[55] Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy.* Cambridge, MA: MIT Press.

[56] Ponzetto, Giacomo A. M. 2011. "Heterogeneous Information and Trade Policy." CEPR Discussion Paper No. 8726.

[57] Poole, Keith T., and Howard Rosenthal. 1985. "A Spatial Model For Legislative Roll Call Analysis." *American Journal of Political Science* 29(2): 357–84.

[58] Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton, NJ: Princeton University Press.

[59] ——. 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York, NY: Simon & Schuster.

[60] Ryan, Stephen P. 2012. "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica* 80(3): 1019–61.

[61] Seabright, Paul. 1996. "Accountability and Decentralization in Government: An Incomplete Contracts Model." *European Economic Review* 40: 61–89.

[62] Snyder, James M., and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118(2): 335–408.

[63] Strömberg, David. 2004. "Radio's Impact on Public Spending." *Quarterly Journal of Economics* 99: 189–221.

[64] 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.

[65] U.S. Department of Commerce. 1971. "Environmental Quality Control." *State and Local Governments Special Studies.* U.S. Bureau of the Census.

[66] U.S. Department of Commerce. 1980. "Environmental Quality Control." *State and Local Governments Special Studies.* U.S. Bureau of the Census.

[67] U.S. Senate. 1970. "Hearings of the subcommitte on air and water pollution of the Committe on Public Works." U.S Government Printing Office.

[68] Wallis, John J. 2000. "American Government Finance in the Long Run: 1790 to 1990." *Journal of Economic Perspectives*, 14(1): 61-82.

[69] ——. 2006. "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.

[70] 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 A1 – Information and the Effects of the CAA on SO_2 Emissions

Dependent Variable: SO_2 Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	-1.641	-0.155	-0.158	-0.092	0.430	-0.102	0.392
× after 1970	(1.063)	(0.925)	(0.904)	(0.975)	(1.223)	(0.884)	(1.179)
Newspaper Circ.	0.319*	0.389*	0.390*	0.424**	0.517**	0.357**	0.431***
× t since 1970	(0.181)	(0.204)	(0.196)	(0.205)	(0.224)	(0.154)	(0.149)
Income (log)		-0.767	-0.535	-0.776	-0.337	-0.740	-0.198
× after 1970		(0.542)	(0.510)	(0.554)	(0.627)	(0.530)	(0.601)
Income (log)		-0.099	-0.130*	-0.104	-0.018	-0.115**	-0.048
× t since 1970		(0.077)	(0.077)	(0.076)	(0.098)	(0.054)	(0.058)
Pop. Density (log)		-0.048	-0.099	-0.046	-0.019	-0.039	-0.053
× after 1970		(0.054)	(0.060)	(0.055)	(0.071)	(0.050)	(0.067)
Pop. Density (log)		0.010	0.017**	0.012	-0.001	0.005	-0.002
× t since 1970		(0.008)	(0.008)	(0.008)	(0.013)	(0.006)	(0.010)
Polluting Mfg. %			3.278**				2.108*
× after 1970			(1.256)				(1.189)
Polluting Mfg. %			-0.441*				0.056
× t since 1970			(0.231)				(0.178)
DW-Nominate				0.066			0.001
× after 1970				(0.196)			(0.174)
DW-Nominate				0.036			-0.004
× t since 1970				(0.038)			(0.021)
Midwest					0.428**		0.394**
× after 1970					(0.185)		(0.186)
Midwest					-0.022		0.007
× t since 1970					(0.022)		(0.020)
South					0.600***		0.553***
× after 1970					(0.174)		(0.165)
South					0.029		0.038**
× t since 1970					(0.026)		(0.017)
West					0.454*		0.410
× after 1970					(0.245)		(0.256)
West					-0.026		0.002
× t since 1970					(0.040)		(0.032)
SO_2 Intensity (log)						0.080	0.032
× after 1970						(0.057)	(0.054)
SO_2 Intensity (log)						-0.049***	-0.050***
× t since 1970						(0.006)	(0.007)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.939	0.941	0.942	0.941	0.943	0.946	0.949

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Independent variables are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

Table A2 – Information and the Effects of the CAA on NO_x Emissions

Dependent Variable: NO_x Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	0.041	0.121	0.120	0.150	0.557	0.165	0.552
× after 1970	(0.262)	(0.356)	(0.338)	(0.368)	(0.394)	(0.348)	(0.369)
Newspaper Circ.	0.160**	0.190**	0.190**	0.187**	0.189*	0.164**	0.157**
× t since 1970	(0.070)	(0.087)	(0.087)	(0.088)	(0.097)	(0.063)	(0.062)
Income (log)		0.020	0.087	0.016	0.055	0.142	0.138
× after 1970		(0.196)	(0.194)	(0.203)	(0.252)	(0.186)	(0.254)
Income (log)		-0.005	-0.006	-0.004	-0.032	-0.078***	-0.081***
× t since 1970		(0.031)	(0.032)	(0.030)	(0.043)	(0.022)	(0.024)
Pop. Density (log)		-0.015	-0.030	-0.014	-0.009	-0.008	-0.013
× after 1970		(0.019)	(0.021)	(0.021)	(0.029)	(0.020)	(0.036)
Pop. Density (log)		-0.003	-0.003	-0.003	0.001	-0.008***	-0.008*
× t since 1970		(0.004)	(0.004)	(0.004)	(0.005)	(0.002)	(0.004)
Polluting Mfg. %			0.939***				0.640*
× after 1970			(0.343)				(0.337)
Polluting Mfg. %			-0.013				0.132
× t since 1970			(0.089)				(0.083)
DW-Nominate				0.031			0.004
× after 1970				(0.060)			(0.065)
DW-Nominate				-0.004			-0.004
× t since 1970				(0.015)			(0.010)
Midwest					0.010		-0.007
× after 1970					(0.043)		(0.040)
Midwest					0.001		0.016*
× t since 1970					(0.009)		(0.009)
South					0.115**		0.084
× after 1970					(0.044)		(0.053)
South					-0.006		0.014
× t since 1970					(0.011)		(0.010)
West					0.105*		0.091
× after 1970					(0.063)		(0.063)
West					0.013		0.018
× t since 1970					(0.017)		(0.013)
NO_x Intensity (log)						0.071**	0.042
× after 1970						(0.027)	(0.032)
NO_x Intensity (log)						-0.043***	-0.050***
× t since 1970						(0.009)	(0.009)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.973	0.973	0.973	0.973	0.974	0.976	0.976

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Independent variables are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

Table A3 – SO_2 Emissions: Alternative Controls for Economic Conditions

Dependent Variable: SO_2 Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	-0.158	-0.251	-0.164	-0.073	-0.183	-0.257	-0.264
× after 1970	(0.904)	(0.935)	(0.962)	(0.968)	(0.964)	(0.955)	(0.967)
Newspaper Circ.	0.390*	0.419**	0.389*	0.391*	0.458**	0.447**	0.486**
× t since 1970	(0.196)	(0.191)	(0.200)	(0.202)	(0.200)	(0.211)	(0.201)
Income (log)	-0.535	-0.647	-0.756	-0.254	-0.205	-0.785	-0.710
× after 1970	(0.510)	(0.598)	(0.595)	(0.580)	(0.584)	(0.696)	(0.799)
Income (log)	-0.130*	-0.135	-0.098	-0.089	-0.119	0.073	-0.006
× t since 1970	(0.077)	(0.083)	(0.072)	(0.082)	(0.078)	(0.108)	(0.089)
Density (log)	-0.099	-0.087	-0.047	-0.031	-0.045	-0.099	-0.162
× after 1970	(0.060)	(0.073)	(0.056)	(0.051)	(0.055)	(0.074)	(0.097)
Density (log)	0.017**	0.022**	0.011	0.011	0.019**	0.031**	0.034**
× t since 1970	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.012)	(0.015)
Polluting Mfg. %	3.278**						1.598
× after 1970	(1.256)						(1.652)
Polluting Mfg. %	-0.441*						-0.230
× t since 1970	(0.231)						(0.278)
Manufacturing %		0.719					0.571
× after 1970		(0.809)					(0.968)
Manufacturing %		-0.217					0.024
× t since 1970		(0.153)					(0.183)
Utilities %			0.952				-8.200
× after 1970			(10.782)				(11.314)
Utilities %			0.057				-0.216
× t since 1970			(2.053)				(1.969)
Fossil Fuels (log)				0.375***	0.321**	-0.200	-0.229
× after 1970				(0.130)	(0.125)	(0.371)	(0.450)
Fossil Fuels (log)				0.007	0.040	0.183*	0.162*
× t since 1970				(0.034)	(0.028)	(0.094)	(0.087)
Coal %					0.354		-0.004
× after 1970					(0.279)		(0.303)
Coal %					-0.215***		-0.181**
× t since 1970					(0.040)		(0.076)
Motor Gas. %						-3.280	-3.332
× after 1970						(2.162)	(2.584)
Motor Gas. %						1.000*	0.660
× t since 1970						(0.574)	(0.578)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.942	0.941	0.940	0.942	0.944	0.943	0.945

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. Manufacturing %, Polluting Manufacturing %, and Utilities % are shares of state GDP. Polluting manufacturing industries are: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). Fossil Fuels is the log of consumption relative to income. Coal % and Motor Gasoline % are shares of fossil-fuel consumption.

Table A4 – NO_x Emissions: Alternative Controls for Economic Conditions

Dependent Variable: NO_x Intensity of Income (log), 1960 to 1981	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	0.120	0.102	0.163	0.139	0.130	0.161	0.237
× after 1970	(0.338)	(0.344)	(0.347)	(0.347)	(0.353)	(0.345)	(0.320)
Newspaper Circ.	0.190**	0.189**	0.189**	0.182**	0.187**	0.176**	0.182**
× t since 1970	(0.087)	(0.087)	(0.084)	(0.068)	(0.073)	(0.069)	(0.068)
Income (log)	0.087	0.044	-0.033	0.138	0.142	0.201	0.136
× after 1970	(0.194)	(0.203)	(0.186)	(0.201)	(0.202)	(0.200)	(0.183)
Income (log)	-0.006	-0.003	-0.003	-0.056*	-0.058*	-0.073	-0.088**
× t since 1970	(0.032)	(0.032)	(0.027)	(0.030)	(0.030)	(0.044)	(0.037)
Density (log)	-0.030	-0.023	-0.018	-0.011	-0.012	-0.003	-0.026
× after 1970	(0.021)	(0.025)	(0.019)	(0.019)	(0.020)	(0.024)	(0.030)
Density (log)	-0.003	-0.004	-0.003	-0.005*	-0.004	-0.007	-0.009*
× t since 1970	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.005)
Polluting Mfg. %	0.939***						0.961*
× after 1970	(0.343)						(0.497)
Polluting Mfg. %	-0.013						0.284**
× t since 1970	(0.089)						(0.120)
Manufacturing %		0.140					0.020
× after 1970		(0.254)					(0.256)
Manufacturing %		0.009					-0.054
× t since 1970		(0.054)					(0.054)
Utilities %			-4.499				-7.437*
× after 1970			(3.438)				(3.745)
Utilities %			0.146				0.606
× t since 1970			(0.696)				(0.400)
Fossil Fuels (log)				0.086***	0.082**	0.154	0.126
× after 1970				(0.032)	(0.034)	(0.104)	(0.120)
Fossil Fuels (log)				-0.038***	-0.035***	-0.056	-0.075**
× t since 1970				(0.012)	(0.012)	(0.036)	(0.031)
Coal %					0.029		-0.066
× after 1970					(0.088)		(0.111)
Coal %					-0.015		-0.032
× t since 1970					(0.019)		(0.025)
Motor Gas. %						0.386	0.218
× after 1970						(0.590)	(0.666)
Motor Gas. %						-0.104	-0.146
× t since 1970						(0.184)	(0.165)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.973	0.973	0.973	0.975	0.975	0.975	0.976

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. Manufacturing %, Polluting Manufacturing %, and Utilities % are shares of state GDP. Polluting manufacturing industries are: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). Fossil Fuels is the log of consumption relative to income. Coal % and Motor Gasoline % are shares of fossil-fuel consumption.

Table A5 – SO_2 Emissions: Alternative Controls for Political Conditions

Dependent Variable: SO_2 Intensity of Income (log), 1960 to 1981							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	-0.092	1.707*	-0.132	0.320	0.272	-0.206	1.678
× after 1970	(0.975)	(0.988)	(0.906)	(0.897)	(0.930)	(0.977)	(1.249)
Newspaper Circ.	0.424**	0.484**	0.397*	0.375*	0.421*	0.381*	0.513**
× t since 1970	(0.205)	(0.234)	(0.208)	(0.216)	(0.220)	(0.198)	(0.241)
Income (log)	-0.776	-0.188	-0.727	-0.502	-0.641	-0.830	-0.452
× after 1970	(0.554)	(0.473)	(0.743)	(0.588)	(0.564)	(0.540)	(0.583)
Income (log)	-0.104	-0.069	-0.085	-0.106	-0.089	-0.068	-0.064
× t since 1970	(0.076)	(0.083)	(0.083)	(0.080)	(0.080)	(0.082)	(0.093)
Density (log)	-0.046	-0.136***	-0.052	-0.062	-0.057	-0.047	-0.140***
× after 1970	(0.055)	(0.050)	(0.054)	(0.053)	(0.053)	(0.053)	(0.051)
Density (log)	0.012	0.006	0.009	0.011	0.010	0.011	0.008
× t since 1970	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
DW-Nominate	0.066						-0.311
× after 1970	(0.196)						(0.240)
DW-Nominate	0.036						0.048
× t since 1970	(0.038)						(0.045)
White Pop. %		-2.741***					-3.789***
× after 1970		(0.726)					(1.175)
White Pop. %		-0.139					-0.065
× t since 1970		(0.129)					(0.148)
% R. in Office			-0.038				0.416
× after 1970			(0.281)				(0.300)
% R. in Office			-0.013				-0.041
× t since 1970			(0.030)				(0.033)
% R. Votes				-0.749*			-0.434
× after 1970				(0.410)			(0.728)
% R. Votes				0.022			0.294*
× t since 1970				(0.103)			(0.150)
Elect. Evenness					-0.586		0.721
× after 1970					(0.420)		(0.727)
Elect. Evenness					-0.043		-0.290
× t since 1970					(0.134)		(0.192)
Sen. Class 1+3						0.047	-0.041
× after 1970						(0.147)	(0.122)
Sen. Class 1+3						0.006	0.012
× t since 1970						(0.022)	(0.023)
Sen. Class 2+3						-0.009	-0.126
× after 1970						(0.115)	(0.130)
Sen. Class 2+3						0.034	0.036
× t since 1970						(0.025)	(0.024)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.941	0.943	0.941	0.941	0.942	0.941	0.945

Notes: Robust standard errors in parentheses, clustered by state; p = $10\%^{*}$, $5\%^{**}$, $1\%^{***}$ All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. The DW-Nominate score is the average for the state's U.S. senators. The share of Republicans in office is computed considering the governor, the leaders of the two state legislatures, the two U.S. senators, and the majority of the state delegation to the U.S. House of Representatives. The share of Republican votes is the average in gubernatorial elections from 1968 to 1972. Election evenness is the average of $1-|\%\text{ R. Votes } -0.5|$ in gubernatorial elections from 1968 to 1972.

Table A6 – NO_x Emissions: Alternative Controls for Political Conditions

Dependent Variable: NO_x Intensity of Income (log), 1960 to 1981

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newspaper Circ.	0.150	0.427	0.188	0.235	0.273	0.111	0.478
× after 1970	(0.368)	(0.365)	(0.338)	(0.341)	(0.336)	(0.364)	(0.375)
Newspaper Circ.	0.187**	0.188*	0.192**	0.217**	0.204**	0.173**	0.152*
× t since 1970	(0.088)	(0.102)	(0.087)	(0.095)	(0.091)	(0.076)	(0.089)
Income (log)	0.016	0.115	0.138	0.084	0.065	-0.006	0.156
× after 1970	(0.203)	(0.215)	(0.215)	(0.237)	(0.215)	(0.196)	(0.241)
Income (log)	-0.004	-0.006	-0.002	0.010	-0.001	-0.017	-0.003
× t since 1970	(0.030)	(0.030)	(0.030)	(0.026)	(0.034)	(0.029)	(0.026)
Density (log)	-0.014	-0.030	-0.026	-0.018	-0.018	-0.015	-0.030
× after 1970	(0.021)	(0.023)	(0.021)	(0.020)	(0.020)	(0.019)	(0.022)
Density (log)	-0.003	-0.003	-0.003	-0.004	-0.003	-0.003	-0.002
× t since 1970	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
DW-Nominate	0.031						0.118*
× after 1970	(0.060)						(0.060)
DW-Nominate	-0.004						0.000
× t since 1970	(0.015)						(0.018)
White Pop. %		-0.451*					-0.136
× after 1970		(0.266)					(0.283)
White Pop. %		0.004					0.047
× t since 1970		(0.048)					(0.063)
% R. in Office			-0.112*				-0.140**
× after 1970			(0.064)				(0.069)
% R. in Office			-0.003				-0.004
× t since 1970			(0.010)				(0.012)
% R. Votes				-0.181			0.108
× after 1970				(0.167)			(0.271)
% R. Votes				-0.043			-0.099
× t since 1970				(0.042)			(0.098)
Elect. Evenness					-0.210		-0.200
× after 1970					(0.142)		(0.264)
Elect. Evenness					-0.019		0.078
× t since 1970					(0.041)		(0.112)
Sen. Class 1+3						0.009	0.009
× after 1970						(0.041)	(0.038)
Sen. Class 1+3						0.015*	0.013
× t since 1970						(0.009)	(0.008)
Sen. Class 2+3						-0.015	-0.016
× after 1970						(0.032)	(0.034)
Sen. Class 2+3						0.004	0.004
× t since 1970						(0.010)	(0.009)
Observations	1056	1056	1056	1056	1056	1056	1056
R^2	0.973	0.973	0.973	0.973	0.973	0.973	0.974

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. All independent variables are 1970 values. The DW-Nominate score is the average for the state's U.S. senators. The share of Republicans in office is computed considering the governor, the leaders of the two state legislatures, the two U.S. senators, and the majority of the state delegation to the U.S. House of Representatives. The share of Republican votes is the average in gubernatorial elections from 1968 to 1972. Election evenness is the average of $1-|\%\text{ R. Votes } -0.5|$ in gubernatorial elections from 1968 to 1972.

Table A7 – SO_2 Emissions: Alternative Dependent Variables

	SO_2/cap.	SO_2/cap.	SO_2/sq.m.	SO_2/sq.m.	SO_2/$	SO_2/$
Income (log) at t					-0.720*	-0.690**
					(0.409)	(0.325)
Pop. Den. (log) at t					-0.820	-0.525
					(1.146)	(1.270)
Newspaper Circ.	-0.096	0.352	-0.048	0.402	-0.073	0.391
× after 1970	(0.881)	(1.146)	(0.870)	(1.133)	(0.880)	(1.154)
Newspaper Circ.	0.392*	0.443***	0.374*	0.449**	0.377*	0.442**
× t since 1970	(0.206)	(0.162)	(0.211)	(0.172)	(0.206)	(0.167)
Income (log)	-0.834	-0.221	-0.884*	-0.294	-0.856	-0.252
× after 1970	(0.528)	(0.575)	(0.513)	(0.555)	(0.511)	(0.548)
Income (log)	-0.076	-0.035	-0.099	-0.066	-0.101	-0.055
× t since 1970	(0.080)	(0.062)	(0.084)	(0.071)	(0.097)	(0.095)
Pop. Density (log)	-0.066	-0.074	-0.062	-0.066	-0.058	-0.063
× after 1970	(0.052)	(0.064)	(0.051)	(0.061)	(0.049)	(0.062)
Pop. Density (log)	0.007	-0.004	0.002	-0.009	0.004	-0.006
× t since 1970	(0.008)	(0.010)	(0.009)	(0.011)	(0.009)	(0.009)
Polluting Mfg. %		1.948*		1.901		1.973
× after 1970		(1.153)		(1.144)		(1.180)
Polluting Mfg. %		0.108		0.122		0.099
× t since 1970		(0.179)		(0.180)		(0.182)
DW-Nominate		0.015		0.049		0.028
× after 1970		(0.171)		(0.168)		(0.171)
DW-Nominate		-0.006		-0.004		-0.004
× t since 1970		(0.022)		(0.022)		(0.022)
Midwest		0.415**		0.414**		0.408**
× after 1970		(0.182)		(0.180)		(0.181)
Midwest		0.006		0.001		0.004
× t since 1970		(0.021)		(0.022)		(0.021)
South		0.559***		0.547***		0.551***
× after 1970		(0.159)		(0.158)		(0.159)
South		0.038**		0.035*		0.037*
× t since 1970		(0.018)		(0.019)		(0.018)
West		0.383		0.380		0.390
× after 1970		(0.246)		(0.241)		(0.249)
West		0.009		0.012		0.008
× t since 1970		(0.032)		(0.033)		(0.034)
SO_2 Intensity (log)		0.036		0.037		0.035
× after 1970		(0.054)		(0.053)		(0.054)
SO_2 Intensity (log)		-0.050***		-0.050***		-0.050***
× t since 1970		(0.007)		(0.007)		(0.007)
Observations	1056	1056	1056	1056	1056	1056
R^2						

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Dependent variables are logarithms. Independent variables interacted with the 1970 break are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

Table A8 – NO_x Emissions: Alternative Dependent Variables

	NO_x/cap.	NO_x/cap.	NO_x/sq.m.	NO_x/sq.m.	NO_x/$	NO_x/$
Income (log) at t					-1.101***	-1.007***
					(0.149)	(0.144)
Pop. Den. (log) at t					-1.031***	-0.933***
					(0.334)	(0.284)
Newspaper Circ.	0.179	0.497	0.227	0.536*	0.235	0.533*
× after 1970	(0.271)	(0.329)	(0.255)	(0.314)	(0.250)	(0.315)
Newspaper Circ.	0.193***	0.179***	0.175***	0.189***	0.175***	0.188***
× t since 1970	(0.071)	(0.055)	(0.058)	(0.052)	(0.055)	(0.052)
Income (log)	-0.047	0.105	-0.097	0.020	-0.105	0.025
× after 1970	(0.174)	(0.228)	(0.152)	(0.203)	(0.148)	(0.205)
Income (log)	0.017	-0.054**	-0.006	-0.077***	-0.004	-0.075***
× t since 1970	(0.025)	(0.026)	(0.019)	(0.029)	(0.022)	(0.028)
Pop. Density (log)	-0.033*	-0.037	-0.029*	-0.032	-0.031*	-0.032
× after 1970	(0.017)	(0.032)	(0.015)	(0.028)	(0.015)	(0.030)
Pop. Density (log)	-0.006*	-0.008*	-0.012***	-0.011**	-0.012***	-0.011**
× t since 1970	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Polluting Mfg. %		0.547*		0.549**		0.548**
× after 1970		(0.282)		(0.272)		(0.272)
Polluting Mfg. %		0.142		0.137		0.137
× t since 1970		(0.085)		(0.085)		(0.085)
DW-Nominate		0.016		0.049		0.047
× after 1970		(0.056)		(0.050)		(0.051)
DW-Nominate		-0.006		-0.003		-0.003
× t since 1970		(0.009)		(0.009)		(0.009)
Midwest		0.019		0.022		0.022
× after 1970		(0.037)		(0.034)		(0.034)
Midwest		0.011		0.003		0.004
× t since 1970		(0.008)		(0.008)		(0.008)
South		0.094**		0.087*		0.088*
× after 1970		(0.046)		(0.044)		(0.044)
South		0.008		0.003		0.003
× t since 1970		(0.009)		(0.009)		(0.009)
West		0.067		0.066		0.066
× after 1970		(0.051)		(0.046)		(0.047)
West		0.024*		0.026**		0.026**
× t since 1970		(0.012)		(0.012)		(0.012)
NO_x Intensity (log)		0.033		0.020		0.020
× after 1970		(0.028)		(0.025)		(0.026)
NO_x Intensity (log)		-0.036***		-0.028***		-0.028**
× t since 1970		(0.010)		(0.010)		(0.012)
Observations	1056	1056	1056	1056	1056	1056
R^2	0.971	0.973	0.995	0.996	0.977	0.979

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Dependent variables are logarithms. Independent variables interacted with the 1970 break are 1970 values. Polluting Manufacturing % is the share of value added represented by: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). The DW-Nominate score is the average for the state's U.S. senators.

Table A9 – Channels of Pollution Abatement

	Income	Pop. Den.	Poll. %	Mfg. %	Util. %	F. Fuel	Coal %	Gas. %
Newspaper Circ.	0.059	0.048	0.020	0.022	0.011	-0.147	0.223	-0.017
× after 1970	(0.158)	(0.042)	(0.014)	(0.038)	(0.007)	(0.297)	(0.151)	(0.038)
Newspaper Circ.	0.002	-0.018	0.000	-0.009	0.001	0.097*	0.041	-0.013
× t since 1970	(0.023)	(0.024)	(0.005)	(0.010)	(0.002)	(0.055)	(0.025)	(0.009)
Income (log)	-0.067	-0.050*	-0.009	-0.018	0.003	0.216	-0.031	-0.035
× after 1970	(0.047)	(0.028)	(0.007)	(0.013)	(0.003)	(0.172)	(0.096)	(0.029)
Income (log)	0.022**	-0.023**	0.001	0.012***	-0.000	-0.008	0.001	0.003
× t since 1970	(0.009)	(0.010)	(0.002)	(0.004)	(0.001)	(0.019)	(0.014)	(0.004)
Pop. Den. (log)	-0.018***	0.004	-0.000	-0.004***	0.001***	-0.004	-0.015	0.004
× after 1970	(0.005)	(0.003)	(0.001)	(0.001)	(0.000)	(0.018)	(0.010)	(0.003)
Pop. Den. (log)	-0.003***	-0.005***	0.000	0.000	0.000	-0.001	-0.003**	0.001**
× t since 1970	(0.001)	(0.001)	(0.002)	(0.001)	(0.000)			
Observations	1056	1056	912	912	912	1056	1056	1056
R^2	0.989	1.000	0.987	0.991	0.929	0.983	0.979	0.975

Notes: Robust standard errors in parentheses, clustered by state; p = 10%*, 5%**, 1%*** All regressions include time f.e., state f.e., and state-specific linear time trends. Independent variables interacted with the 1970 break are 1970 values. Income and Population Density are logarithms. Manufacturing %, Polluting Manufacturing %, and Utilities % are shares of state GDP. Polluting manufacturing industries are: Paper and allied products (SIC 26), Chemicals and allied products (SIC 28), Petroleum and coal products (SIC 29), Stone, clay, glass, and concrete (SIC 32), Primary metal industries (SIC 33). Fossil Fuels is the log of consumption relative to income. Coal % and Motor Gasoline % are shares of fossil-fuel consumption.

Ivie

Guardia Civil, 22 - Esc. 2, 1º
46020 Valencia - Spain
Phone: +34 963 190 050
Fax: +34 963 190 055

Website: www.ivie.es
E-mail: publicaciones@ivie.es