Hierarchical Accountability in Government: Theory and Evidence*

Razvan Vlaicu† Alexander Whalley‡

July 2013

Abstract

This paper develops a hierarchical agency model with informational asymmetries where a policymaker accountable to the voter through an intermediary behaves either as an insulated bureaucrat or as a pandering politician. We examine the model's predictions using U.S. city-level data and find a negative differential in popular police employment policies between indirectly-elected managers and directly-elected mayors. The policy effects vary with political and informational determinants of policymaker incentives and are robust to instrumentation by precipitation shocks that influenced early 20th century manager government adoptions for reasons obsolete today.

JEL Classification: D72, D73, H70.

Keywords: indirectly-elected policymaker, accountability, government incentives, pandering, city manager.

*We would like to thank Richard Akresh, Chris Berry, Richard Boylan, Ethan Bueno de Mesquita, Philip Curry, Jim Davies, George Deltas, Allan Drazen, Fernando Ferreira, Price Fishback, Lisa George, Gabriele Gratton, John Hatfield, Alex Hirsch, Shawn Kantor, Ethan Kaplan, Sumon Majumdar, Steeve Mongrain, Pablo Montagnes, Mattias Polborn, Gergely Ujhelyi, Emanuel Vespa, Joel Waldfogel, John Wallis, and seminar participants at IIES, New York University, Stockholm University, University of Chicago, University of Houston/Rice University, University of Illinois at Urbana, Canadian Public Economics Conference, Econometric Society NAWM, and Princeton Political Economy Conference. Erin Moody, Krystal Tapia, Michael Wall, and Marko Zivanovic provided excellent research assistance.

†University of Maryland, Department of Economics, and Department of Government and Politics, 3114 Tydings Hall, College Park, MD 20742, USA. E-mail: vlaicu@econ.umd.edu.

‡University of California - Merced, School of Social Sciences, Humanities & Arts, 5200 North Lake Road, Merced, CA 95343, USA; and NBER, 1050 Massachusetts Ave., Cambridge, MA 02138, USA. E-mail: awhalley@ucmerced.edu.

1 Introduction

Policymakers often operate under hybrid accountability forms, i.e., those that combine the pure appointment (bureaucrat) and pure election (politician) forms prevalent in the theoretical literature (e.g., Besley and Coate 2003, Maskin and Tirole 2004, Alesina and Tabellini 2007). A well-studied example in the empirical literature is the city manager, with responsibilities similar to an elected city mayor, but who is hired/fired through an elected city council. Unlike a politician a city manager is not answerable to voters; unlike a typical bureaucrat, he is not protected by civil service tenure.[1]

Do indirectly-accountable policymakers have bureaucrat- or politician-like incentives? Despite the significant research interest in city managers, the literature has remained inconclusive about which of "the two prominent but seemingly contradictory views of managers – that of 'professional helping hands' versus 'political hired hands'" is more compelling (Stillman 1977, p. 659). For instance, in Rauch (1995) the manager "is assumed to have lifetime tenure and can be terminated only for just cause. [The city manager] thus provides municipal government with [...] 'bureaucratic insulation' from the political process." (p. 970). By contrast, Deno and Mehay (1987) lean toward a politician interpretation: "Since city council members who implement programs and budget levels inconsistent with that preferred by the median voter also face the prospect of being replaced, a city manager appears constrained by the median voter's preferences as much as an elected mayor." (p. 628). This second view is formalized in the complete-information model of Persson, Roland, and Tabellini (1997) where in equilibrium voters hold the policymaker (executive) accountable to the same policy standard through an intermediary (legislature) as they would directly.[2]

In this paper we study indirectly-elected policymakers theoretically and empirically. We first develop a hierarchical agency (voter-intermediary-policymaker) model with informational asymmetries that generates political and informational conditions when the policymaker behaves like an insulated bureaucrat and conditions when he behaves like a pandering politician. The model's key insight is that while a popular policy signals preference congru-

[1] "If the manager is not responsive to the governing body, it has the authority to terminate the manager at any time." (ICMA 2007, p. 2). About half of U.S. cities are run by city managers. Other examples of this hybrid accountability form include prime ministers in parliamentary democracies, county executives in various developed countries, school superintendents in U.S. public school districts. In U.S. cities manager government was introduced during the reform movements of the Progressive Era (1890-1930) and was inspired by the corporate model where the CEO was indirectly and continuously accountable to the shareholders through an elected board of directors.

[2] One caveat is possible policymaker-intermediary collusion. In that case, an indirectly-elected policymaker has more discretion to pursue rent-seeking than a directly-elected policymaker. Persson, Roland, and Tabellini (2000) apply these ideas to fiscal policy in parliamentary and presidential regimes.

ence the intermediary's retention decision is an additional signal the voter uses to provide incentives. If policymaker congruence is high or the voter's policy expertise is low the retention signal dominates the policy signal and the voter always reelects an intermediary retaining the policymaker, in effect insulating the policymaker. Conversely, if the policy signal dominates the retention signal the voter gives the intermediary incentives to retain only popular policymakers, which in turn gives the policymaker the incentive to pander.

In our empirical application we study policymaking by U.S. city managers, indirectly-elected officials with the same major policy responsibilities as directly-elected mayors, i.e., writing the budget and hiring personnel. We argue that police officer employment is naturally susceptible to pandering incentives as crime has consistently ranked among the top two local policy issues in Gallup surveys of local attitudes since 1959 (see Gallup 2000), while civilian police employment, e.g., administrators, dispatchers, should not elicit such a clear popular preference. Across a number of specifications we find that on average managers employ fewer police officers per capita than mayors but a comparable number of police civilians per capita.

A central challenge to estimating institutional effects is that institutions may be endogenous to policymaking, for instance through unobserved voter preferences (Aghion, Alesina, and Trebbi 2004). To address potential endogeneity in accountability form we propose an instrument for manager government. The instrument is based on the observation that before the 1936 Flood Control Act transferred flood prevention from local governments to the Army Corps of Engineers cities often responded to flood-related infrastructure crises by adopting manager government because it facilitated the ascension of engineers into the top executive office. We document that pre-1936 precipitation shocks influenced early switches to manager government for reasons obsolete today and show that the police employment patterns noted above also appear in this IV setting.[3]

We further explore the theory model's incentive mechanisms and find that the officer employment differential increases in policymaker congruence and decreases in voters' crime priors, and is most pronounced in election years when the theory model predicts incentives should be sharper. We argue that these patterns are difficult to explain with alternative mechanisms, such as patronage motivations or policymaker type selection.[4]

Our paper is related to several strands of literature. The literature on how career concerns distort the incentives of asymmetrically informed experts includes Prendergast (1993), Mor-

[3]To our knowledge this is the first instrument for manager government in the literature, apart from Baqir (2002) who uses 30-year lags of city institutions to instrument for current city institutions.

[4]The existing evidence for electoral cycles focuses on fiscal policy, is mostly at the national or state level, and is generally weaker for developed countries (Drazen 2000).

ris (2001), Ashworth (2005), Prat (2005), Fox (2007), Levy (2007), Morelli and Van Weelden (2013), Che, Dessein, and Kartik (2013). We complement this literature by showing theoretically and empirically how institutional rules alleviate pandering.[5] In contract theory and corporate finance hierarchical agency models study optimal incentives through wage rather than political contracts (e.g., Strausz 1997, Park 2000), or strategic information transmission (Li 2010). Our main result echoes the finding in this literature that under asymmetric information an intermediary allows the principal to commit to a broader range of incentive structures. Weingast and Moran (1983), Besley and Coate (2003), Blanes i Vidal and Leaver (2011), Martinez-Bravo (2011), and Lim (2013) empirically examine appointed officials such as fixed-term or tenured regulators or judges. We focus on officials that can be replaced at will, e.g., through a no-confidence vote, and argue that their behavior can resemble either a bureaucrat or a politician depending on their environment. The empirical literature on U.S. city managers has not focused on pandering in policymaking instead finding that managers outspend mayors (Coate and Knight 2011), are more likely to privatize city services (Levin and Tadelis 2010), and reduce full-time city employment (Enikolopov 2012). We document differences in police employment, arguably more susceptible to pandering incentives and propose an instrument to deal with potential endogeneity in government form.[6]

2 Theory

At first sight, holding a policymaker accountable through an intermediary seems to disconnect policymaker choices from voter preferences. If the voter is fully informed he can, nevertheless, exploit the intermediary's reelection motivation to control policymaker moral hazard just as effectively as under direct accountability (Persson, Roland, and Tabellini 1997). In an incomplete information environment, however, the voter also faces an adverse selection problem which negatively impacts direct accountability by producing pandering incentives, i.e., the policymaker follows popular opinion by choosing popular policies, instead of using his expertise to choose optimal policies (Canes-Wrone, Herron, and Shotts 2001, Maskin and Tirole 2004). Even if the voter's best interest is to limit policymaker pandering a commitment problem renders him unable to do so: Ex-ante an uninformed voter may prefer to insulate the better-informed policymaker, but ex-post he is better off replacing an

[5]Pandering incentives are not confined to government settings. See Dasgupta and Prat (2006) for a model of fund manager conformism and Panova (2006) for a model of pandering by the media.

[6]Coate and Knight's (2011) citizen-candidate model attributes differences in spending not to incentives, but to voters electing different councilmember types. Two earlier papers (Baqir 2002, MacDonald 2008) found no spending differential between managers and mayors.

unpopular policymaker.[7]

When accountability to an imperfectly-informed voter distorts the policymaker's incentives in this way, delegating policymaker accountability to an informed intermediary seems justifiable. The case for delegation is less clear-cut, however, if the intermediary's preferences may differ from the voter's. In that case the voter needs to provide incentives for the intermediary to act in the voter's interest. To what extent can an informed intermediary help the voter insulate the policymaker from pandering incentives? If insulation is achievable, is it complete, allowing the policymaker bureaucrat-like discretion to act on his expertise, or partial, preserving some politician-like bending to public opinion?

To understand this hybrid accountability form we extend Maskin and Tirole's (2004) model by introducing an asymmetrically informed intermediary in the voter-policymaker relationship. For easy linking with our empirical application to police employment this section uses terminology specific to that particular policy domain. The model, however, applies in any setting featuring an indirectly-elected policymaker with asymmetric expertise.[8]

Model. At time t a policy choice x_t needs to be made from the set of policy alternatives $\{0, 1\}$. For instance, the policy issue can be police employment, in which case $x_t = 1$ can stand for "high police" and $x_t = 0$ for "low police." The policy's effect depends on the state of the world $s_t \in \{0, 1\}$ prevailing in that period. In the illustration state $s_t = 1$ can represent "high crime" and state $s_t = 0$ can represent "low crime." The state is i.i.d. across periods. Let $\pi = \mathbb{P}\{s_t = 1\}$ denote the probability of the "high crime" state. Assume $\pi > \frac{1}{2}$.[9]

There are three kinds of players: a voter, an intermediary, and a policymaker. The voter gets a unit of payoff when the policy matches the state, zero otherwise: $v(x_t, s_t) = \mathbf{1}\{x_t = s_t\}$. Notice that based on the crime prior π the voter prefers policy $x_t = 1$. Using Maskin and Tirole's (2004) terminology, we say that "high police" is the *popular* policy.[10]

[7] While closely related to the concept of populism (Acemoglu, Egorov, and Sonin 2012), pandering need not be associated with redistributive policy and affects public officials of all ideological stripes, not only those who are naturally inclined to the cause of the "common man," e.g., working-class politicians. Besley and Smart (2007) and Smart and Sturm (2011) notice the commitment issue in the context of fiscal rules and term limits, respectively. Appointed officials such as life-tenured judges are immune to pandering incentives (Maskin and Tirole 2004), while term limits (Smart and Sturm 2011) and informative media (Ashworth and Shotts 2010) can limit politician pandering.

[8] The agency approach to political accountability is surveyed in Besley (2006).

[9] Gallup's annual Crime Survey asks the question "Is there more crime in your area than there was a year ago, or less?" Since the survey started in 1973 the percentage of respondents who say "More" exceeded those that say "Less" except for three years 1998, 2000, 2001 (see Gallup 2010). Assuming the state is i.i.d. across periods means that past policy choices do not affect the voter's current crime prior. This seems to be a good approximation if police have only a short-term effect on crime, i.e., long-term crime is driven by more fundamental forces like economic inequality, economic growth, and race relations.

[10] Stucky (2005) reviews the literature on how the public's crime prior (also known as "fear of crime" in

The voter delegates policymaking to a policymaker (P) and delegates policymaker accountability to an intermediary (I). These two agents are both policy-motivated and office-motivated. Their policy motivation depends on their type, congruent or dissonant: $\theta_t^P, \theta_t^I \in \{C, D\}$. A congruent agent has the same preferences as the voter: $u(x_t, s_t|C) = \mathbf{1}\{x_t = s_t\}$. A dissonant agent has preferences opposite to the voter's: $u(x_t, s_t|D) = \mathbf{1}\{x_t = 1 - s_t\}$.[11]

The timing of the infinite-horizon game is as follows. (I) Every period the policymaker observes the state $s_t \in \{0, 1\}$ and chooses between the unpopular/popular policy: $x_t \in \{0, 1\}$. (II) Every period the intermediary observes policy and policymaker type (x_t, θ_t^P) and decides whether to replace/retain the policymaker: $y_t \in \{0, 1\}$. (III) Every *other* period the voter, having observed choices $[(x_{t-1}, y_{t-1}), (x_t, y_t)]$, but not types $\left(\theta_t^P, \theta_t^I\right)$ or the current state s_t, decides to replace/reelect the intermediary: $z_t \in \{0, 1\}$. An agent exiting at t is succeeded by an agent (challenger) whose type is a new draw from the type distribution. An exiting agent cannot run for office again.[12]

For expositional simplicity we assume that reelection incentives affect only election-period behavior (cf. Rogoff 1990, Shi and Svensson 2006).[13] Let $(\eta_t^j)_{t=1,2,\ldots}$ be sequences of i.i.d. preference shocks, for $j = P, I$ where $\eta_t^j \in \{C, D\}$ and $\mathbb{P}\left\{\eta_t^j = C\right\} = \gamma^j$ is the congruence probability. Assume $\gamma^j > \frac{1}{2}$.[14] An agent's type is either last period's or the current period's preference shock: $\theta_t^j \in \left\{\eta_{t-1}^j, \eta_t^j\right\}$ and $\mathbb{P}\left\{\theta_t^j = \eta_{t-1}^j\right\} = \lambda$. Assume $0 < \lambda < 1$. These assumptions have two implications: First, an agent's type is correlated over time between two adjacent periods, but uncorrelated between two non-adjacent periods. In particular, an agent's type at $t + 1$ (non-election period) is correlated with his type at t (election period), but uncorrelated with his type at $t - 1$ (non-election period).[15] Second, although initially policymaker and intermediary types are independent, correlation between their types can nevertheless develop through the intermediary's equilibrium retention decisions.

The policymaker cares about policies he himself chooses. The intermediary cares about the choices of the policymaker he is keeping in office. Formally, lifetime utility for a policymaker at the beginning of an electoral term, i.e., at $t - 1$, is $u(x_{t-1}, s_{t-1}|\theta_{t-1}^P) +$

the criminology literature) affects popular preference for police.

[11] The analysis would be similar if we assumed $u(x_t, s_t|D) = \mathbf{1}\{x_t = 0\}$.

[12] A key feature of hierarchical accountability is that the indirectly-elected policymaker can be replaced between elections. A city manager, for example, serves at the city council's pleasure. His "job tenure is only secure until the next council meeting" (Stillman 1977, p. 664). The model captures this feature by assuming that the intermediary can replace the policymaker in the middle of an electoral term.

[13] In Appendix A.2 we relax this assumption by allowing reelection incentives to affect both election and non-election period behavior (cf. Martinez 2009) and show that the qualitative predictions below are upheld.

[14] The parameter γ^j can be interpreted as the fraction of voters favoring state-matching policies. Then, an entering agent can be thought of as a random draw from the electorate.

[15] The type persistence probability is $\mathbb{P}\{\theta_t^j = \theta_{t-1}^j\} = \mathbb{P}\{\theta_t^j = \eta_{t-1}^j, \theta_{t-1}^j = \eta_{t-1}^j\} = \lambda\,(1 - \lambda)$, for any t.

$\sum_{i=t}^{\infty} \beta^{i-t+1} u(x_i, s_i|\theta_i^P) \prod_{\tau=t}^{i} y_{\tau-1}$; for an intermediary it is $[u(x_{t-1}, s_{t-1}|\theta_{t-1}^I)+\beta u(x_t, s_t|\theta_t^I)]+ \sum_{i=t}^{\infty} \beta^{2i-2t+2}[u(x_{2i-t+1}, s_{2i-t+1}|\theta_{2i-t+1}^I) + \beta u(x_{2i-t+2}, s_{2i-t+2}|\theta_{2i-t+2}^I)] \prod_{\tau=t}^{i} z_{2\tau-t}$. Here β is a time discount factor, with $0 < \beta < 1$.

As is standard our policymaker is a policy expert who knows whether the popular policy is optimal (the state). The intermediary, on the other hand, is a political expert, in the sense of Gailmard and Jenkins (2009), who knows the policymaker's preferences (his type). This assumption seems reasonable in settings where the intermediary has frequent contact with the policymaker.[16] The voter, on the other hand, knows neither the popular policy's optimality (moral hazard) nor the policymaker's preferences (adverse selection). Thus, in this model hierarchical accountability features two forms of the classic delegation tradeoff: by delegating policymaking to a policy expert the voter may benefit from more informed policy choices but may suffer from the policymaker's dissonance; by delegating policymaker accountability to a political expert the voter may benefit from more informed retention decisions but may suffer from the intermediary's dissonance.

We denote voter posterior beliefs about agent types by $\gamma_t^j = \mathbb{P}\left\{\theta_t^j = C|X_t, Y_t\right\}$, where $j = P, I$ and $(X_t, Y_t) = (x_\tau, y_\tau)_{\tau=1}^t$ is the public history at time t. Following standard practice, we refer to posterior beliefs about agent type as the agent's "reputation."

Equilibrium. The game described above is a principal-intermediary-agent model with incomplete information. We solve for its Perfect Bayesian Equilibrium, namely: (i) policymaker, intermediary, and voter strategies (x_t, y_t, z_t) that are sequentially rational; and (ii) voter beliefs $\left(\gamma_t^P, \gamma_t^I\right)$ that are consistent with agent strategies. We add the requirement that voter strategies be robust to a small fraction of non-strategic agents implying that in a pooling equilibrium both on- and off-equilibrium path beliefs are derived from sincere, i.e., preference-based, behavior.[17] We restrict attention to pure strategies. Throughout this section we assume that the discount factor β is large enough that both agent types seek to remain in office; the equilibrium will thus characterize agents' extrinsic *incentives*.[18]

Since the voter can only replace the policymaker through an intermediary, his problem is to reelect that intermediary that seems more in tune with the voter's preferences. Lemma 1 in Appendix A.1 shows that following a two-period electoral term $[t-1, t]$, after the voter has observed choices $[(x_{t-1}, y_{t-1}), (x_t, y_t)]$, the incumbent intermediary's public reputation

[16] That is probably the case for a city council since the policymaker they choose, the city manager, has to attend all city council meetings and sometimes committee meetings.

[17] See Maskin and Tirole (2004) for application of this refinement in a related setting.

[18] In Appendix A.3 we also solve the case where agents have no career concerns and instead follow their preferences; that equilibrium characterizes selection-driven behavior.

increases whenever $\hat{\gamma}_t^I(x_t, y_t) \equiv \mathbb{P}\left\{\eta_t^I = C | (x_t, y_t), \theta_t^I = \eta_t^I\right\} \geq \gamma^I$. In words, the voter evaluates congruence based on election-period behavior (x_t, y_t) under the assumption that the intermediary's type is based on current preferences.[19] The following proposition states the implications of this voter calculation for policymaker incentives. For contrast we also include the case of no intermediary, i.e., direct accountability, studied by Maskin and Tirole (2004).

Proposition 1 *For popular policy issues ($\pi > 1/2$) hierarchical accountability either insulates the policymaker pre-election, if $\gamma^P > \pi$, or creates policymaker pandering incentives pre-election, if $\gamma^P < \pi$; direct accountability always creates pre-election pandering incentives for the policymaker. Agents follow their preferences post-election; in particular, an intermediary retains only a policymaker of his own type.*

Proof. See Appendix A.1.

Under direct accountability the voter uses the *policy signal* x_t to learn about the policymaker's type. Because $\pi > 1/2$ a popular policy is more likely optimal thus the choice of a congruent policymaker, therefore it is a signal of congruence: $\hat{\gamma}_t^P(x_t = 1) = \frac{\gamma^P \pi}{\gamma^P \pi + (1-\gamma^P)(1-\pi)} > \gamma^P > \hat{\gamma}_t^P(x_t = 0) = \frac{\gamma^P(1-\pi)}{\gamma^P(1-\pi) + (1-\gamma^P)\pi}$. The voter's best response then is to reelect only after the popular policy, giving the policymaker incentives to pander, i.e., choose the popular policy regardless of its optimality.

Hierarchical accountability provides the voter an additional signal: the *retention signal* contained in the intermediary's retention decision y_t. Because the intermediary prefers to retain only policymakers of his own type, retention is a signal of intermediary congruence *if* the policymaker is more likely congruent. Thus, because for a popular policymaker $\hat{\gamma}_t^P(x_t = 1) > \gamma^P > 1/2$ retaining a popular policymaker signals intermediary congruence: $\hat{\gamma}_t^I(x_t = 1, y_t = 1) = \frac{\gamma^I \gamma^P \pi}{\gamma^I \gamma^P \pi + (1-\gamma^I)(1-\gamma^P)(1-\pi)} > \gamma^I$, and replacing a popular policymaker signals intermediary dissonance: $\hat{\gamma}_t^I(x_t = 1, y_t = 0) = \frac{\gamma^I\left(1-\gamma^P\right)(1-\pi)}{\gamma^I(1-\gamma^P)(1-\pi) + (1-\gamma^I)\gamma^P \pi} < \gamma^I$.

For an unpopular policymaker $\hat{\gamma}_t^P(x_t = 0) > \frac{1}{2}$ iff $\gamma^P > \pi$. That is, he may still be believed by voters to be more likely congruent if: (a) his initial reputation γ^P was high, or (b) the strength π of the policy signal is low. In these cases retaining an unpopular policymaker signals intermediary congruence: $\hat{\gamma}_t^I(x_t = 0, y_t = 1) = \frac{\gamma^I \gamma^P (1-\pi)}{\gamma^I \gamma^P (1-\pi) + (1-\gamma^I)(1-\gamma^P)\pi}$ ($> \gamma^I$ iff $\gamma^P > \pi$), and replacing him signals dissonance: $\hat{\gamma}_t^I(x_t = 0, y_t = 0) = \frac{\gamma^I\left(1-\gamma^P\right)\pi}{\gamma^I(1-\gamma^P)\pi + (1-\gamma^I)\gamma^P(1-\pi)}$ ($> \gamma^I$ iff $\gamma^P < \pi$). In other words, when the retention signal dominates the policy signal the voter gives the intermediary incentives to retain the policymaker without regard to his policy choices, in effect insulating the policymaker. Conversely, if the policy signal dominates the

[19] Since the voter's reelection decision at t depends only on (x_t, y_t), at $t-1$ the intermediary keeps only policymakers of his own type; this makes the intermediary's type in non-election-periods matter to the voter.

retention signal, then retaining an unpopular policymaker signals a dissonant intermediary. In this case the voter gives the intermediary incentives to retain only popular policymakers, which in turn gives the policymaker the incentive to pander.[20]

One may argue that another way in which indirectly-elected policymakers differ from directly-elected policymakers is that they have weaker motivations to deliver political patronage to various constituencies. While patronage could potentially affect popular policies, such as crime fighting or low taxes, it also typically influences *neutral policies* such as in-kind transfers or city jobs that may not elicit a clear popular preference.

Proposition 2 *For neutral policy issues* ($\pi = 1/2$) *policymaking outcomes do not vary with the accountability form in election periods or non-election periods.*

Proof. See Appendix A.1.

A balanced state prior silences the policy signal. This allows a voter exerting direct accountability to credibly insulate the policymaker since the voter cannot link a certain policy to a particular type. Under hierarchical accountability the equilibrium is also insulating because the retention signal is always stronger than the policy signal, $\gamma^P > \pi = 1/2$. These two effects imply that incentives on neutral issues should not differ by accountability form.

Propositions 1 and 2 show how the incentives of indirectly-elected policymakers may depend on their political and policy environment; observationally, this creates a negative election-period policy differential in popular policies between hierarchical and direct accountability when $\gamma^P > \pi$. The propositions are robust to a setup where the voter factors into his reelection decision both election-period and non-election-period policy.[21] The model can also be solved under the assumption that the accountability form's primary role is not to provide incentives but to select certain types of policymakers. In that scenario, since hierarchical selection is most effective at the end of the term (following intermediary selection) whereas direct selection at the beginning of the term (following the election), the election-year popular policy differential would be predicted to be positive rather than negative.[22] This allows us to distinguish between the two models empirically.

[20] These two types of intermediary behaviors are consistent with case studies of U.S. city councils that distinguish between council strategies that insulate the manager, termed "blind faith," versus council strategies that transmit popular preferences, termed "political." See Stillman (1977).

[21] In Appendix A.2 we solve the model under this relaxed assumption. We find that election-period pandering is driven by an intertemporal tradeoff: choosing a popular policy against preferences at the beginning of the term results in a sure loss, whereas postponing it for the end of the term results in a potential loss, i.e., because next period the state may be such that the policymaker prefers the popular policy. The same logic applies to the intermediary.

[22] See Appendix A.3 for details. The key insight of the selection model is that policymaker selection through an intermediary is more precise than directly because of the intermediary's superior information, but may backfire if intermediary selection by the voter is weak.

3 Empirical Application

The theory results in Propositions 1-2 imply that if policymaker behavior is driven by incentives we should observe the following patterns:

(H1) A negative policy differential should exist between hierarchically- and directly-accountable policymakers on popular policy issues. No such differential should exist on neutral policy issues.

(H2) The negative policy differential on popular issues should increase in policymaker congruence and decrease in the voter's crime prior.

(H3) The negative policy differential on popular issues should be larger in election years. No such differential should exist on neutral issues either in election or in non-election years.[23]

Our empirical application examines popular (police officer employment) and neutral (police civilian employment) policy issues in U.S. cities where chief executives are either indirectly-elected (managers) or directly-elected (mayors).

Data. Our sample consists of all U.S. cities with 1900 Census population over 17,500 residents. This sample selection criterion is not affected by government form choice, since manager reforms do not occur until the small city of Staunton, Virginia (pop. 7,289 in 1900) starts experimenting with it in 1908. After dropping Washington DC, since it has a federally-appointed city government until 1973, a number of 248 present-day cities satisfy this criterion. The sample period for our panel is 1960-2000.[24]

We point out two features of government form variation in our sample. First, the maps in Figures A1 and A2 display little geographic clustering in manager government. Second, Figure A3 shows that most manager charter adoptions in our sample occur before 1960. In fact, the majority of adoptions occurred before the federal takeover of flood control in 1936, and only ten occurred after 1960. We measure local policymaking behavior using police employment. This policy area has the attractive feature that it allows us to distinguish popular from neutral policy issues by disaggregating police employment into officer and civilian employees, according to the distinction made in the FBI's *Uniform Crime Reports*.[25]

[23] An additional prediction of the model is higher volatility in popular policies under hierarchical accountability because the policymaker is more responsive to the stochastic state. We explore this implication in Appendix Table A2. We thank an anonymous referee for this point.

[24] Two intervening annexations and one merger slightly alter the sample: Pittsburgh, PA annexed Allegheny, PA in 1907; Omaha, NE annexed South Omaha, NE in 1915; and West Hoboken, NJ merged with Union Hill, NJ to form Union City, NJ in 1925.

[25] *Census of Governments* city employment data has the advantage that it distinguishes between full-time and part-time city employees, but distinguishes between police officers and civilians only starting in 1977. We examine this alternative data in the Appendix.

We construct our instrument using 1900-2000 weather reports from the U.S. Historical Climatology Network's *Daily Temperature, Precipitation, and Snow Data*. This dataset contains daily readings for temperature extremes, rainfall, and snowfall collected from weather stations throughout the U.S. We match a city to the closest weather station based on geographic coordinates.[26] Our main precipitation measure is the yearly sum of rainfall and density-adjusted snowfall at the station level.[27] We define an extreme precipitation event as a year when precipitation exceeds the 99th percentile of the national 20th century yearly precipitation distribution. Our cross-sectional measure of precipitation shocks for a given city is the frequency of extreme precipitation events in a given period, referred to below as *precipitation shocks*. For instance, precipitation shocks in the local flood control period (1900-1936) are referred to as *LFC precipitation shocks*.

In addition to these key variables, we work with an extensive set of geographic, demographic, economic, crime, and policymaker controls. Appendix A.6 provides the complete list of variables, with details about their measurement and sources.

Table 1 reports descriptive statistics for our major variables by government form. Panel A shows that manager governments employ on average about 19% fewer officers per capita than mayor governments, and virtually the same number of civilians per capita. Interestingly, Panel B suggests that before the advent of the manager plan there were no major differences in either type of police employment. Other statistically significant differences are: manager cities have on average lower officer payroll, stricter officer education requirements, higher property crime (Panel A), higher policymaker salary (Panel C), higher incidence of Progressive institutions and Republican mayors (most manager cities maintain a honorary mayor position) (Panel D), and higher fraction of population college graduates (Panel E).[28]

Figure A4 provides a first look at how manager government affected police employment historically. Manager governments employ fewer officers from 1960 onward, but not in 1900; the same cannot be said for civilians. In the bottom panel we divide cities based on whether they experienced LFC precipitation shocks and find that cities hit by these shocks have lower officer employment after 1960, with the difference attenuating over time. Again, civilian

[26] The U.S. has 126 weather stations reporting in 1900. The median distance to the closest weather station for our sample cities is 47 miles. As the opening of new stations could be related to changes in local weather or local economy we keep every city matched to the same station over time.

[27] We adjust snowfall for water density by dividing it by ten, as suggested by the U.S. Department of Agriculture. See http://www.ak.nrcs.usda.gov/snow/data.

[28] If police reduces crime, less police under manager government implies higher crime. While Table 1 appears to contradict this potential implication in the case of violent crime Appendix Table A13 shows that after controlling for year effects, geographic, and demographic factors the correlation is positive.

employment does not follow this same pattern.[29]

Ordinary Least Squares. We first estimate a simple model of the form:

$$\log(Police_{i,t}) = \beta_1 Manager_{i,t} + \beta_2 X_{i,t} + \psi_t + \epsilon_{i,t} \quad (1)$$

where $Police_{i,t}$ is either officer or civilian police per capita for city i in year t, $Manager_{i,t}$ is a dummy variable indicating manager form of city government, $X_{i,t}$ is a set of controls, ψ_t's are year fixed effects, and $\epsilon_{i,t}$ is the error term. The coefficient β_1 measures the conditional difference in mean police employment between manager and mayor cities. According to hypothesis (H1) we expect $\beta_1 < 0$ for officer employment and $\beta_1 = 0$ for civilian employment.

This model allows us to account for measurable city geographic, demographic, economic, and political characteristics, as well as policymaker characteristics, and to control for national trends affecting local police employment. As government form changes infrequently during 1960-2000 and the observations are unlikely to be independent within a city we cluster the standard errors at the city level.[30]

Table 2 presents OLS estimates of the police employment differential β_1. Accounting only for year effects column (1) shows that manager governments employ significantly fewer officers. Adding division fixed effects (the U.S. has nine Census divisions), geographic and demographic controls to account for voter preferences in columns (2) and (3) reduces the point estimate to 7-8% but it remains statistically significant at the 5% level.[31]

Manager government may be related to other institutional and political factors that could affect policy. The descriptive statistics in Table 1 reflect the historical fact that manager charter reform often came on the heels of other urban reforms: non-partisan ballots, at-large elections, and civil service rules. The literature has sometimes packaged these reforms together with manager reform, distinguishing only between "traditional" and "reformed" cities (Stucky 2005).[32] Manager cities may also differ in police union strength and media penetration, factors that may affect policymaker congruence (γ^P) and voter crime prior (π). Adding these institutional and political controls in column (4) does little to alter the point estimate, although statistical significance decreases somewhat. Finally, column (5) shows

[29] Federal grants to fund police hiring through the Community Oriented Police Services (COPS) program established by Congress in 1994 may have reduced local discretion over police employment.

[30] See Appendix Table A10 Panel B for results using alternative standard error definitions.

[31] Voter or special interest preferences can affect institutional choice. Morelli and Van Weelden (2013) argue that pandering can depend on how divisive an issue is for the electorate and Aghion, Alesina, and Trebbi (2004) discuss how the distribution of voter preferences may affect the choice of institutions.

[32] Knoke (1982) attributes successful switches to manager to strong business interests, weak unions, high population mobility, small immigrant population, and small city size.

that the officer differential is robust to controls for policymaker salary, gender, and race.[33]

The second half of Table 2 changes the dependent variable from police officers to police civilians. We find little evidence of a similar statistically significant relationship between government form and police civilians. The point estimates in columns (6)-(10) change sign, are closer to zero, and not statistically significant. While the theory model considers popular and neutral issues in isolation, complementarities between officer and civilian police may well exist in practice. Strong complementaries should produce similar point estimates across the two employment measures.[34]

While the lower number of officers in manager cities may be consistent with managers' weaker patronage motivations (Enikolopov 2012) the lack of a similar pattern for civilians seems to belie this explanation. Moreover, when we look at the other two major city employee groups, firefighters and teachers, the pattern we find for police is again absent (see Appendix Table A3). Overall these patterns seem more consistent with pandering to voters' crime perceptions than with patronage to city employees.[35]

Using a different sample and exploiting switches to manager government Coate and Knight (2011) find that managers spend more than mayors. We find that manager governments employ fewer police officers. To reconcile these seemingly contradictory results we examine disaggregated city spending data in Appendix Table A4. The table first replicates Coate and Knight's (2011) finding in our sample. It then shows that manager governments pay higher officer salaries, in line with Edwards and Edwards (1982) and Deno and Mehay (1987), but have similar officer payrolls and police department spending overall. Taken together these results are consistent with both a spending effect driven by city councilmember selection (Coate and Knight 2011) and a police officer employment effect driven by policymaker incentives.[36]

---

[33] Ferraz and Finan (2011) and Gagliarducci and Nannicini (2013) find that higher salaries increase the productivity of Brazilian city council members and Italian mayors by attracting more competent types. Salary may be endogenous to policy outcomes. In Appendix Table A14 we find little evidence that policymakers who reduce crime have higher salaries. Gender and race have been shown to affect local government policy, e.g., Nye, Rainer, and Stratmann (2010), Ferreira and Gyourko (2011).

[34] In Appendix Table A1 we find that the officer differential is driven by full-time empoyment, consistent with Enikolopov (2012). In Appendix Table A2 we also document a positive officer employment *volatility* differential, consistent with the theory model and which is stronger for full-time employment.

[35] In fact, if the educational requirements and background checks for hiring a police officer are stricter than for a civilian any patronage effect should be more pronounced in civilian employment, all else equal, e.g., labor supply for these positions, state mandates of minimum forensic staff.

[36] Since the conceptual source of the officer differential is an informational distortion, it is useful to compare its magnitude with the impact of information on government policy from prior work. For instance, Stromberg's (2004) estimated 0.201 elasticity of federal unemployment relief spending with respect to radio penetration (see his Table II column IV) would imply a 20.1% upper bound, in absolute value, on the officer differential, corresponding to a change from zero radio penetration to full penetration. All the estimates of

Instrumental Variables. Even if the OLS specifications above control for all relevant confounds, estimates may still be biased by reverse causality. Despite the infrequency of actual changes in city institutions (e.g., 0.8% of 1,420 U.S. cities had changed their government form between 1980-90 according to Table 1 in Baqir 2002) city charters are endogenous by virtue of being subject to revision by popular referendum. Measurement error is another potential source of bias in OLS estimates leading to attenuation. As manager cities typically have an honorary mayor, city clerks in these cities sometimes mistakenly report a mayor form of government on ICMA survey forms (see Coate and Knight 2011).

To address concerns with bias in the OLS estimates we develop an IV approach. Our strategy is based on the observation that infrastructure crises often triggered early 20th century switches to manager government because it facilitated the ascension of engineers into the top executive office. Floods caused by extreme precipitation were one such crisis. Dayton, Ohio provides an illustration. In March 1913 after days of heavy rainfall the Great Miami River overflowed the city's levees causing a flood that destroyed over 20,000 homes. In the immediate aftermath local leaders sought to rebuild the flood control system with a large public works campaign. Expediency dictated the adoption of a manager charter so that an engineer could be appointed to lead the reconstruction effort.

After subsequent crises, e.g., the Great Mississippi Flood of 1927, the Northeast Flood of 1936, resulted in substantial losses in multiple local jurisdictions Congress passed the Flood Control Act in 1936 transferring responsibility for flood prevention and management to the Army Corps of Engineers. The hundreds of miles of levees and 375 major reservoirs built by the federal government after 1936 significantly weakened the link between heavy precipitation and the incidence of floods.[37] We thus instrument for manager government using the frequency of precipitation shocks in the local flood control (1900-1936) era, *LFC precipitation shocks*. Formally, the first stage model is:

$$\begin{aligned} Manager_{i,t} &= \alpha_1 LFC\_Precipitation\_Shocks_i + \alpha_2 Century\_Precipitation\_Shocks_i + \\ &\quad + \alpha_3 Median\_Precipitation_i + \alpha_4 X_{i,t} + \psi_t + \epsilon_{i,t} \qquad (2) \end{aligned}$$

where $LFC\_Precipitation\_Shocks_i$ is the frequency of precipitation shocks in the local flood control era for city i, $Century\_Precipitation\_Shocks_i$ is the frequency of city precipitation shocks during the 20th century, $Median\_Precipitation_i$ is median annual city

the officer differential in Table 2 are below this upper bound.

[37] Appendix Table A5 shows that the relationship between extreme precipitation and flood incidence is markedly stronger in the local flood control period (1900-1936).

precipitation during the 20th century, and the remaining variables are from equation (1). We include $Century\_Precipitation\_Shocks_i$ to make our exclusion restriction credible as previous research found that climate affects economic growth (Dell, Jones, and Olken 2012), crime (Jacob, Lefgren, and Moretti 2007), conflict (Miguel, Satyanath, and Sergenti 2004), and the origins of trust (Durante 2010).[38]

Our identification strategy requires that, conditional on typical local precipitation patterns: (i) cities hit by LFC precipitation shocks have the same average unobserved characteristics as spared cities, and (ii) LFC precipitation shocks affect police employment decades later only through their effect on government form. In the Appendix we provide supportive evidence for these identifying assumptions.[39]

Table 3 presents IV estimates of the officer and civilian employment differentials controlling for typical local precipitation patterns. In the odd-numbered columns the only additional controls are year effects. In the even-numbered columns we use within-Census-division variation and control for geographic factors. The first stage shows a strong relationship between LFC precipitation shocks and present-day manager government: the F-statistic exceeds the critical value of 10 below which finite-sample weak-instrument bias could be a concern (Bound, Jaeger, and Baker 1995). In the second stage we find significantly lower officer employment in manager cities, and mixed evidence of a civilian employment differential. The larger point estimates relative to the baseline OLS results suggest that measurement error in government form might be present in our sample (Coate and Knight 2011). Overall the IV results uphold the substantive conclusions derived from the OLS estimates.[40]

If the costs of changing government form are heterogeneous, the identified effects will be local to a subset of cities and potentially different from the population-wide treatment effect. For example, Acemoglu, Robinson, and Torvik (2013) argue that voters dismantle exogenously imposed checks and balances when politician rents are low and special interests are strong. One way to explore this issue is to examine how sensitive the IV estimates are

[38] In a related IV strategy using country-level data Bruckner and Ciccone (2011) exploit the fact that in non-democratic societies the cost of popular opposition to an authoritarian regime is lower during times of economic distress, rendering negative rainfall shocks democracy biased.

[39] In support of instrument validity and the exclusion restriction we find that: (i) LFC precipitation shocks are not correlated with early city characteristics (Appendix Table A6); (ii.a) in contrast to LFC precipitation shocks, federal flood control (1936-1960) era (FFC) precipitation shocks and LFC temperature shocks are not related to police employment today (Appendix Table A7); and (ii.b) LFC precipitation shocks are not related to other present-day city institutions (Appendix Table A8). Trends in city managers' educational backgrounds provide additional support for our identification strategy. More than 95 percent of city managers were engineers in 1918, 77 percent in 1934, and only 18 percent in 1977 (Stillman 1977).

[40] Baqir (2002) and Whalley (2013) also find little evidence that related forms of institutional endogeneity contaminate OLS results in the context of U.S. local governments.

to changes in instrument construction. If the effects of government form are heterogeneous and highly specific to cities induced to change government form by our proposed instrument, alternative instrument definitions would likely change the magnitude of the estimated effects. Appendix Table A9 shows that the IV results are robust to alternative instrument definitions, strengthening their external validity. Our OLS and IV results are also robust to a wide variety of sample alterations, inference procedures, and alternative outcome measures (e.g., scaling employment by number of crimes instead of by population), see Appendix Table A10.[41]

Political and Informational Mechanisms. According to our theory model a popular policy differential strictly speaking only exists for high-congruence low-crime-prior cities where voters can credibly commit to insulate the policymaker. Here we examine (H2) which states that the policy differential on popular issues should increase in policymaker congruence and decrease in voter crime prior.

We use the preclearance requirement of the 1965 Voting Rights Act (VRA) as a shock to policymaker congruence (γ^P). Section 4 of the VRA contains a formula to select jurisdictions required to obtain court or federal preclearance for any change to their local election laws.[42] Previous research has documented that enactment of the VRA created an influx of newly elected mayors and councilmembers increasing the representation of previously disenfranchised voters especially in areas with low voter registration (e.g., Cascio and Washington 2012, Shah, Ruhil, and Marschall 2013).[43] We expect that the increase in policymaker congruence in cities located in preclearance jurisdictions would lead to manager insulation which would be reflected in a more negative officer differential.

We use the "crack epidemic" that started to affect U.S. cities in the early 1980s as a shock to the voter crime prior (π). The emergence of crack cocaine and associated media attention heightened public perceptions of crime in urban areas (Blenko 1993). Importantly for identification the geographic spread of crack was driven by proximity to a few entry points and other topographic conditions, rather than by local enforcement activities (Evans, Garthwaite, and Moore 2012). Under hypothesis (H2) we expect the increase in voter crime

[41] An additional external validity question is whether the identified effects apply at other levels of government using hierarchical accountability; this is an interesting question for future research.

[42] According to the U.S. Justice Department: "The first element in the formula was that the state or political subdivision of the state maintained on November 1, 1964, a "test or device," restricting the opportunity to register and vote. The second element of the formula would be satisfied if the Director of the Census determined that less than 50 percent of persons of voting age were registered to vote on November 1, 1964, or that less than 50 percent of persons of voting age voted in the presidential election of November 1964." (http://www.justice.gov/crt/about/vot/ accessed in July 2013).

[43] This could have occured either through strengthening electoral competition (Besley, Persson, and Sturm 2010) or reducing the power of entrenched special interests (Olson 1983).

prior due to the crack epidemic to attenuate the negative officer differential toward zero.[44]

In Table 4 we add to equation (1) an interaction of $Manager_{i,t}$ with each mechanism. The results in columns (1)-(3) show that VRA preclearance amplifies the negative officer differential while the crack epidemic reduces it. Similar statistically significant interactions are not present in columns (4)-(6) for the civilian differential.[45]

Electoral Effects. Our theory model predicts that pandering incentives are strongest pre-election because the voter relies more heavily on recent information about policy choices. We measure election proximity by coding the year before the election date as an election year.[46] To estimate electoral effects we add to equation (1) the election indicator $Election_{i,t}$ and the interaction $Manager_{i,t} \times Election_{i,t}$, as well as city fixed effects. In this model $\beta_{Manager}$ measures the manager-mayor policy differential in non-election years, while $\beta_{Manager} + \beta_{Manager \times Election}$ measures the manager-mayor policy differential in election years. Hypothesis (H3) states that the election year employment differential is more negative than the non-election year differential for popular issues, which means $\beta_{Manager \times Election} < 0$ for officers, and that there are no employment differentials on neutral issues, i.e., civilians.

Table 5 column (1) shows that mayors employ 0.6% more officers in election years, suggesting pandering ahead of an election, although the estimate is statistically weak. To increase precision columns (2)-(4) restrict the sample to the subset of 174 cities with population over 50,000 in 1960.[47] The coefficient pattern from column (1) reappears in columns (2)-(3) and is now also statistically significant at conventional levels. The employment cycle could be confounded by a spending cycle although in Appendix Table A12 we find no differential spending cycle in our sample. The further addition of government spending in column (4) leaves the estimate of $\beta_{Manager \times Election}$ and its standard error virtually unchanged. The corresponding point estimates for civilians in columns (5)-(8) are all closer to zero and never

[44] Economists and other social scientists have provided evidence that the crack epidemic increased crime. For example, Grogger and Willis (2000) conclude that "Our results indicate that, in the absence of crack cocaine, the crime rate in 1991 would have remained below its previous peak in the early 1980s." (p. 528).

[45] The similar magnitudes between $\beta_{Manager \times VRAPreclearance}$ for officers and civilians suggest that VRA Preclearance may pick up factors other than policymaker congruence. The negative $\beta_{CrackEpidemic}$ for civilians could be explained by crime-driven urban decline, an effect that rising voter crime priors offsets for officers; economists have found effects of crime on urban depopulation in this period (e.g., Cullen and Levitt 1999). In Appendix Table A11 we rerun these specifications with alternative, although perhaps less plausibly exogenous, measures of policymaker congruence (no past police unionization) and voter crime prior (large change in city crime relative to county crime); the interaction estimates are similar to the ones in Table 4.

[46] Election dates are set in the city charter; special elections outside the regular election cycle, say to replace a mayor resigning before the end of the term, are infrequent and we ignore them.

[47] Levitt (1997) argues that electoral cycles should be more pronounced in larger cities. He exploits the correlation between election years and police hiring in a panel of 59 large U.S. cities with directly-elected mayors to identify the effect of police on crime.

statistically significant. These results are consistent with the incentive effects at the heart of the model (see Propositions 1-2 above) as opposed to a model of policymaker selection by preference type which would imply a positive election-year differential (see Appendix A.3).

4 Conclusion

This paper sheds light on the incentive effects of a hybrid accountability form that is pervasive at various levels of governments. Our principal-intermediary-agent model gives informational asymmetries a central role and shows that hierarchical accountability creates either bureaucrat or politician incentives depending on political and informational factors. It implies that in a setting with government expertise this structure rather than hindering democratic accountability increases the flexibility of voter control over policymaking incentives. The theory guides our empirical analysis of U.S. city managers in a salient area of local policy, law enforcement. The estimates for the main institutional effect and its mechanisms are consistent with the theoretically characterized incentives.

For simplicity we assumed that informational asymmetries are exogenous to the institutional environment, yet motivations for acquiring information by the voter (Larcinese 2009) as well as the policymaker (Stephenson 2011) may depend on accountability form: on one hand, the voter might be less interested in acquiring information if he cannot directly remove the policymaker; on the other hand, an indirectly-elected policymaker, who enjoys more policy discretion, may acquire more policy information. Thus, the informational asymmetry may be more severe under hierarchical accountability. Empirically, a more stringent test of the model would be to verify that voter and intermediary strategies adjust to the political and informational environment to provide optimal incentives for an indirectly-elected policymaker. This requires more detailed data on council elections and manager retention (see Enikolopov 2012 for an effort in this direction).

References

Acemoglu, D., J. Robinson, and R. Torvik (2013) "Why Do Voters Dismantle Checks and Balances?" *Review of Economic Studies* 80(3): 845-875.

Acemoglu, D., G. Egorov, and K. Sonin (2012) "A Political Theory of Populism," *Quarterly Journal of Economics* 128(2): 771-805.

Aghion, P., A. Alesina, and F. Trebbi (2004) "Endogenous Political Institutions," *Quar-*

terly Journal of Economics 119(2): 565-611.

Alesina, A. and G. Tabellini (2007) "Bureaucrats or Politicians? Part I: A Single Policy Task," *American Economic Review* 97(1): 169–179.

Ashworth, S. (2005) "Reputational Dynamics and Political Careers," *Journal of Law, Economics, and Organization* 21(2): 441-466

Ashworth, S. and K.W. Shotts (2010) "Does Informative Media Commentary Reduce Politicians' Incentives to Pander?" *Journal of Public Economics* 94(11-12): 838-847.

Baqir, R. (2002) "Districting and Government Overspending," *Journal of Political Economy* 110(6): 1318-1354.

Besley, T. (2006) *Principled Agents? The Political Economy of Good Government*, Oxford, UK: Oxford University Press.

Besley, T. and S. Coate (2003) "Elected Versus Appointed Regulators: Theory and Evidence," *Journal of the European Economic Association* 1(5): 1176-1206.

Besley, T., T. Persson, and D. Sturm (2010) "Political Competition, Policy, and Growth: Theory and Evidence from the US," *Review of Economic Studies* 77: 1329–1352.

Besley, T. and M. Smart (2007) "Fiscal Restraints and Voter Welfare," *Journal of Public Economics* 91(3-4): 755-773.

Blanes i Vidal, J. and C. Leaver (2011) "Are Tenured Judges Insulated from Political Pressure?" *Journal of Public Economics* 95(7-8): 570-586.

Blenko, S. (1993) *Crack and the Evolution of Anti-Drug Policy*, Westport, CT: Greenwood Press.

Bound, J., D.A. Jaeger, and R.M. Baker (1995) "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association* 90(430): 443-450.

Bruckner, M. and A. Ciccone (2011) "Rain and the Democratic Window of Opportunity," *Econometrica* 79(3): 923-947.

Canes-Wrone, B., M. Herron, and K. Shotts (2001) "Leadership and Pandering: A Theory of Executive Policymaking," *American Journal of Political Science* 45(3): 532-550.

Cascio, E. and E. Washington (2012) "Valuing the Vote: the Redistribution of Voting Rights and State Funds Following the Voting Rights Act of 1965," *Quarterly Journal of Economics* forthcoming.

Che, Y., W. Dessein, and N. Kartik (2013) "Pandering to Persuade," *American Economic Review* 103(1): 47-79.

Coate, S. and B. Knight (2011) "Government Form and Public Spending: Theory and

Evidence from U.S. Municipalities," *AEJ: Economic Policy* 3(3): 82-112.

Cullen, J.B. and S.D. Levitt (1999) "Crime, Urban Flight, and the Consequences for Cities," *Review of Economics and Statistics* 81(2): 159-169.

Dasgupta, A. and A. Prat (2006) "Financial Equilibrium with Career Concerns," *Theoretical Economics* 1(1): 67-93.

Dell, M., B. Jones, and B. Olken (2008) "Climate Shocks and Economic Growth: Evidence from the Last Half Century," *AEJ: Macroeconomics* 4(3): 66-95.

Deno, K. and S.L. Mehay (1987) "Municipal Management Structure and Fiscal Performance: Do City Managers Make a Difference?" *Southern Economic Journal* 53(3): 627-642.

Drazen, A. (2000) "The Political Business Cycle after 25 Years," *NBER Macroeconomic Annual* 15: 75-117.

Durante, R. (2010) "Risk, Cooperation, and the Economic Origins of Social Trust: An Empirical Investigation," working paper, Sciences Po.

Edwards, L. and F. Edwards (1982) "Public Unions, Local Government Structure, and the Compensation of Municipal Sanitation Workers," *Economic Inquiry* 29(3): 405-25.

Enikolopov, R. (2012) "Politicians, Bureaucrats and Targeted Redistribution: The Role of Career Concerns," working paper, New Economic School.

Evans, W., C. Garthwaite, and T.J. Moore (2012) "The White/Black Educational Gap, Stalled Progress, and the Long Term Consequences of the Emergence of Crack Cocaine Markets," working paper, Northwestern University.

Ferraz, C. and F. Finan (2011) "Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance," working paper, UC Berkeley.

Ferreira, F. and J. Gyourko (2011) "Does Gender Matter for Political Leadership? The Case of U.S. Mayors," working paper, University of Pennsylvania.

Fox, J. (2007) "Government Transparency and Policymaking," *Public Choice* 131: 23–44.

Gailmard, S. and J. Jenkins (2009) "Agency Problems, the 17th Amendment, and Representation in the Senate," *American Journal of Political Science* 53(2): 324-342.

Gagliarducci, S. and T. Nanicinni (2013) "Do Better Paid Politicians Perform Better?" *Journal of the European Economic Association* 11(2): 369-398.

Gallup, Inc. (2000) "Crime Tops List of Americans' Local Concerns," by Lydia Saad. http://www.gallup.com/poll/2800/crime-tops-list-americans-local-concerns.aspx.

Gallup, Inc. (2010) "Americans Still Perceive Crime as on the Rise," by Jeffrey M. Jones. http://www.gallup.com/poll/144827/americans-perceive-crime-rise.aspx.

Grogger, J. and M. Willis (2000) "The Emergence of Crack Cocaine and the Rise in

Urban Crime Rates," *The Review of Economics and Statistics* 82(4): 519-529.

ICMA (2007) *Council-Manager Form of Government*, Washington, DC: ICMA Press.

Jacob, B., L. Lefgren, and E. Moretti (2007) "The Dynamics of Criminal Behavior," *Journal of Human Resources* 42(3): 489-527.

Knoke, D. (1982) "The Spread of Municipal Reform: Temporal, Spatial, and Social Dynamics," *American Journal of Sociology*, 87(6): 1314-1339.

Larcinese V. (2009) "Information Acquisition, Ideology, and Turnout: Theory and Evidence From Britain," *Journal of Theoretical Politics* 21(2): 237-276.

Levin, J. and S. Tadelis (2010) "Contracting for Government Services: Theory and Evidence from U.S. Cities," *Journal of Industrial Economics* 58(3): 507-541.

Levitt, S. (1997) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review* 87(3): 270-290.

Levy, G. (2007) "Decision Making Procedures for Committees of Careerist Experts," *American Economic Review, Papers and Proceedings*, 97(2): 306-310.

Li, W. (2010) "Peddling Influence through Intermediaries," *American Economic Review* 100(3): 1136-1162.

Lim, C.S.H. (2013) "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges," *American Economic Review* 103(4): 1360-1397.

MacDonald, L. (2008) "The Impact of Government Structure on Local Public Expenditures," *Public Choice* 136(3-4): 457-473.

Martinez, L. (2009) "A Theory of Political Cycles," *Journal of Economic Theory* 144(3): 1166–1186.

Martinez-Bravo, M. (2011) "The Role of Local Officials in New Democracies: Evidence from Indonesia," working paper, Johns Hopkins University.

Maskin, E. and J. Tirole (2004) "The Politician and the Judge: Accountability in Government," *American Economic Review* 94(4): 1034-1054.

Miguel, E., S. Satyanath, and E. Sergenti, (2004) "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy* 112(4): 725-753.

Morris, S. (2001) "Political Correctness," *Journal of Political Economy* 109(2): 231-265.

Morelli, M. and R. Van Weelden (2013) "Ideology and Information in Policymaking," *Journal of Theoretical Politics* forthcoming.

Nye, J.V., I. Rainer, and T. Stratmann (2010) "Do Black Mayors Improve Black Employment Outcomes? Evidence from Large U.S. Cities," working paper, GMU.

Olson, M. (1983) "The South Will Fall Again: The South as Leader and Laggard in

Economic Growth," *Southern Economic Journal* 49(4): 917-932.

Panova, E. (2006) "A Model of Media Pandering," working paper, UQAM.

Park, C. (2000) "Monitoring and Structure of Debt Contracts," *Journal of Finance* 55(5): 2157-95.

Persson, T., G. Roland and G. Tabellini (1997) "Separation of Powers and Political Accountability," *Quarterly Journal of Economics* 112: 1163-1202.

Persson, T., G. Roland and G. Tabellini (2000) "Comparative Politics and Public Finance," *Journal of Political Economy* 108(6): 1121-1161.

Prat, A. (2005) "The Wrong Kind of Transparency," *American Economic Review* 95(3): 862-877.

Prendergast, C. (1993) "A Theory of 'Yes Men'," *American Economic Review* 83: 757-70.

Rauch, J.E. (1995) "Bureaucracy, Infrastructure, and Economic Growth: Evidence from U.S. Cities During the Progressive Era," *American Economic Review* 85(4): 968-979.

Rogoff, K. (1990) "Equilibrium Political Budget Cycles," *American Economic Review* 80(1): 21-36.

Shah, P., A. Ruhil, and M. Marschall (2013) "Are We There Yet? The Voting Rights Act and Black Representation on City Councils, 1981-2006," *Journal of Politics* forthcoming.

Shi, M. and Svensson, J. (2006) "Political Budget Cycles: Do They Differ Across Countries and Why?" *Journal of Public Economics* 90(8-9): 1367-1389.

Smart, M. and D. Sturm (2011) "Term Limits and Electoral Accountability," working paper, LSE.

Stephenson, M. (2011) "Information Acquisition and Institutional Design," *Harvard Law Review* 124: 1422-1483.

Stillman, R.J. (1977) "The City Manager: Professional Helping Hand, or Political Hired Hand?" *Public Administration Review* 37(6): 659-670.

Strausz, R. (1997) "Delegation of Monitoring in a Principal-Agent Relationship," *Review of Economic Studies* 64(3): 337-357.

Stromberg, D. (2004) "Radio's Impact of Public Spending," *Quarterly Journal of Economics* 119(1): 189-221.

Stucky, T.D. (2005) "Local Politics and Police Strength," *Justice Quarterly* 22: 139-169.

Weingast, B.R. and M.J. Moran (1983) "Bureaucratic Discretion or Congressional Control? Regulatory Policymaking by the FTC," *Journal of Political Economy* 91(5): 765-800.

Whalley, A. (2013) "Elected Versus Appointed Policymakers: Evidence from City Treasurers," *Journal of Law and Economics* 56(1): 39-81.

TABLE 1: Descriptive Statistics, By Government Form

Variable	Manager Cities mean (1)	Manager Cities std. dev. (2)	Mayor Cities mean (3)	Mayor Cities std. dev. (4)	Difference p-value (5)
Panel A: Police Department, 1960-2000					
Officers Per Capita	1.85	0.55	2.20	1.12	0.000
Civilians Per Capita	0.38	0.24	0.37	0.36	0.760
Police Dept. Spending per Capita	204	227	198	213	0.576
Officer Payroll per Capita (1977-2000)	74	30	90	53	0.000
Officer Avg. Pay (1977-2000)	36,841	10,096	36,685	10,986	0.855
Officer Educ. Min. (1993,1997,2000)	0.31	0.46	0.17	0.37	0.009
Police Unionization (1968)	0.37	0.43	0.31	0.44	0.266
Violent Crime Rate (1975-2000)	3.35	2.82	4.00	4.09	0.089
Property Crime Rate (1975-2000)	85.36	35.94	73.21	39.03	0.001
Panel B: Police Department, 1900					
Officers Per Capita	1.10	0.34	1.10	0.46	0.935
Civilians Per Capita	0.08	0.11	0.06	0.06	0.074
Arrests Per Capita	0.07	0.05	0.05	0.04	0.022
Panel C: Policymaker, 1992-1993,1995-2000					
Annual Salary	100,096	27,760	66,687	27,360	0.000
Male	0.88	0.33	0.84	0.37	0.262
White	0.80	0.40	0.77	0.42	0.526
Panel D: City Institutions, 1960-2000					
Non-Partisan Elections (1960)	0.85	0.35	0.53	0.50	0.000
Fraction At-Large Seats (1960)	0.83	0.34	0.55	0.43	0.000
Early Civil Service	0.69	0.46	0.70	0.46	0.886
Incumbent Mayor Republican	0.47	0.50	0.27	0.45	0.004
Local Newspaper Sales (1990-2000)	0.42	1.06	0.51	1.88	0.610
City Government Spending	2,115	2,584	2,208	2,681	0.548
Panel E: City Demographics, 1960-2000					
Population Size	146,198	184,417	252,521	712,626	0.074
Population Density	5,820	19,744	6,556	8,837	0.732
Fraction Non-White	0.19	0.15	0.20	0.20	0.900
Fraction College Graduate	0.15	0.08	0.13	0.08	0.020
Household Income	28,653	7,477	28,241	7,498	0.487
Number of Observations	3,622		6,546		10,168

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States as of 1900. Each entry in columns (1)-(4) presents the mean/std. dev. for the indicated variable and column (5) has p-values for the hypothesis that the difference between the means in columns (1) and (3) is zero, using standard errors clustered at the city level. The number of observations reflects the maximum sample size across all the reported variables. All monetary variables are expressed in 2000$.

TABLE 2: Government Form and Police Employment: OLS Estimates

Dependent Variable =	Log(Officers per Capita)					Log(Civilians per Capita)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Manager	-0.16***	-0.08**	-0.07***	-0.05*	-0.07**	0.09	0.02	0.03	-0.02	-0.06
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.07)	(0.07)	(0.06)	(0.07)	(0.08)
Additional Controls:										
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Demographic	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Institutional & Political	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Policymaker	No	No	No	No	Yes	No	No	No	No	Yes
Sample Years	Full	Full	Full	Full	1992+	Full	Full	Full	Full	1992+
Number of Clusters	248	248	248	248	236	248	248	248	248	235
Number of Observations	9,974	9,974	9,974	9,974	1,164	9,850	9,850	9,850	9,850	1,156

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6, except for columns (5) and (10) where the sample is 1992-1993, 1995-2000. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. Columns (1)-(10) report OLS estimates of β_1 from equation (1) in the text. Standard errors clustered at the city level reported in parentheses. Columns (1) and (6) only include year fixed effects as additional controls. Columns (2) and (7) also include Census division fixed effects and the following geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction, *Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil*. Columns (3) and (8) also include the following demographic controls: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and an indicator variable for *Population Density Missing*. Columns (4) and (9) also include controls for *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Police Unionization 1968*, and *Local Newspaper Sales 1990-2000,* as well as indicator variables for missing values for each variable. Columns (5) and (10) also control for *log(Policymaker Salary)*, *Policymaker Male,* and *Policymaker White.* Detailed reporting of these regressions is in Appendix Table A15. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE 3: Government Form and Police Employment: IV Estimates

Dependent Variable =	Manager		Log(Officers per Capita)		Manager		Log(Civilians per Capita)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager	–	–	-0.28**	-0.50**	–	–	0.69**	0.19
			(0.14)	(0.21)			(0.35)	(0.43)
LFC Precipitation Shocks	11.67***	9.79***	–	–	11.69***	9.85***	–	–
	(2.85)	(2.54)			(2.85)	(2.54)		
Century Precipitation Shocks	-6.90***	-9.86***	-0.27	-1.78	-6.95***	-9.87***	7.11***	2.23
	(2.07)	(2.29)	(0.87)	(1.54)	(2.07)	(2.30)	(1.82)	(2.36)
Median Precipitation	-0.54	1.41*	0.55**	0.38	-0.53	1.42*	-1.41**	-1.41
	(0.45)	(0.81)	(0.22)	(0.61)	(0.45)	(0.81)	(0.60)	(1.12)
Additional Controls:								
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	No	Yes	No	Yes	No	Yes	No	Yes
Excluded Instrument *F*-stat	16.76	14.88	–	–	16.83	14.99	–	–
[p-value]	[0.0001]	[0.0001]			[0.0001]	[0.0001]		
Sample Years	Full	Full	Full	Full	Full	Full	Full	Full
Number of Clusters	248	248	248	248	248	248	248	248
Number of Observations	9,974	9,974	9,974	9,974	9,850	9,850	9,850	9,850

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States as of 1900. Columns (3),(4) and (7),(8) report IV estimates of β_1 from equation (1) in the text. Columns (1),(2) and (5),(6) report first-stage estimates. Standard errors clustered at the city level reported in parentheses. The excluded instrument is *LFC Precipitation Shocks*. The models in odd-numbered columns only include year fixed effects as additional controls. The models in even-numbered columns also include Census division fixed effects and the following geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction*, Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil.* Detailed reporting of these regressions is in Appendix Table A16. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE 4: Government Form and Police Employment: Mechanisms

Dependent Variable =	Log(Officers per Capita)			Log(Civilians per Capita)		
	(1)	(2)	(3)	(4)	(5)	(6)
Manager	-0.03	-0.07**	-0.06**	0.00	-0.02	-0.01
	(0.03)	(0.03)	(0.03)	(0.07)	(0.08)	(0.08)
Manager × Voting Rights Act Preclearance	-0.22*	–	-0.23**	-0.31	–	-0.30
	(0.12)		(0.12)	(0.20)		(0.20)
Manager × Crack Epidemic	–	0.05**	0.06***	–	0.00	0.02
		(0.02)	(0.02)		(0.06)	(0.06)
Voting Rights Act Preclearance	0.11	–	0.11	0.04	–	0.04
	(0.08)		(0.08)	(0.17)		(0.18)
Crack Epidemic	–	-0.02	-0.02	–	-0.21***	-0.21***
		(0.03)	(0.03)		(0.07)	(0.07)
Additional Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	Yes	Yes	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	Yes	Yes	Yes	Yes
Institutional & Political	Yes	Yes	Yes	Yes	Yes	Yes
Sample Years	Full	Full	Full	Full	Full	Full
Number of Clusters	248	248	248	248	248	248
Number of Observations	9,974	9,974	9,974	9,850	9,850	9,850

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States as of 1900. The table reports OLS estimates of equation (1) in the text augmented with the indicated variables and interaction terms. Standard errors clustered at the city level reported in parentheses. The models include year and Census division fixed effects, and geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction, *Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil*, demographic controls: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and an indicator variable for *Population Density Missing*, institutional and political controls: *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Police Unionization 1968*, and *Local Newspaper Sales 1990-2000,* as well as indicator variables for missing values for each variable. Detailed reporting of these regressions is in Appendix Table A17. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE 5: Government Form and Police Employment: Electoral Effects

Dependent Variable =	Log(Officers per Capita)				Log(Civilians per Capita)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager	-0.078	-0.064	-0.044	-0.008	0.009	-0.040	-0.034	0.068
	(0.056)	(0.084)	(0.048)	(0.057)	(0.095)	(0.110)	(0.086)	(0.136)
Manager × Election	-0.009*	-0.015**	-0.013***	-0.013***	-0.001	-0.009	-0.007	-0.006
	(0.005)	(0.007)	(0.005)	(0.005)	(0.013)	(0.015)	(0.014)	(0.014)
Manager × Log(Spending per Capita)	–	–	–	-0.005	–	–	–	-0.015
				(0.007)				(0.017)
Election	0.006	0.010**	0.009**	0.009**	0.008	0.006	0.006	0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.009)	(0.008)	(0.008)
Log(Spending per Capita)	–	–	–	0.005	–	–	–	-0.003
				(0.004)				(0.008)
Additional Controls:								
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic	No	No	Yes	Yes	No	No	Yes	Yes
Sample Years	Full	Full	Full	Full	Full	Full	Full	Full
Sample Cities	All	Large	Large	Large	All	Large	Large	Large
Number of Clusters	248	174	174	174	248	174	174	174
Number of Observations	9,974	7,033	7,033	7,033	9,850	7,017	7,017	7,017

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year. In columns (1) and (5) the sample is the largest 248 self-governing cities in the United States as of 1900. In the remaining columns the sample is the 174 cities that also have at least 50,000 residents in 1960. The table reports OLS estimates of equation (1) in the text augmented with the indicated variables and interaction terms. Standard errors clustered at the city level reported in parentheses. Columns (1)-(2),(5)-(6) only include year and city fixed effects as additional controls. Columns (3) and (7) also include demographic controls: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and an indicator variable for *Population Density Missing*. Columns (4) and (8) also include an indicator variable for *Missing City Spending*. Detailed reporting of these regressions is in Appendix Table A18. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

Appendix [For Online Publication]

Section A.1 contains formal proofs of Propositions 1-2 in the text. Sections A.2-A.3 provide additional theory results. Sections A.4-A.5 present supplementary empirical results. Sections A.6-A.7 contain data definitions and sources, followed by Figures A1-A4 and Tables A1-A18.

A.1 Incentives Model

Assumption Ia *Given per-period i.i.d. preference shocks* $\left(\eta_t^j\right)_{t=1}^{\infty}$ *such that* $\eta_t^j \in \{C, D\}$ *with* $\mathbb{P}\left\{\eta_t^j = C\right\} = \gamma^j$*, agent* j*'s type at* $t+1$ *is* $\theta_{t+1}^j \in \left\{\eta_t^j, \eta_{t+1}^j\right\}$ *with* $\mathbb{P}\left\{\theta_{t+1}^j = \eta_t^j\right\} = \lambda$.

Assumption IIa *The discount factor* β *satisfies:* $\frac{\beta(1-\pi)}{1-\beta} > 1$.

Lemma 1 *Under Assumption Ia, after a history of observable actions* $(X_t, Y_t) = (x_\tau, y_\tau)_{\tau=1}^t$*, incumbent* $j = P, I$ *reputation depends on period* t *actions:* $\gamma_{t+1}^j \equiv \mathbb{P}\left\{\theta_{t+1}^j = C \middle| (X_t, Y_t)\right\} \geq \gamma^j$ *iff* $\hat{\gamma}_t^j(x_t, y_t) \equiv \mathbb{P}\left\{\eta_t^j = C \middle| (x_t, y_t), \theta_t^j = \eta_t^j\right\} \geq \gamma^j$.

Proof of Lemma 1. At $t+1$ incumbent j's type may be based on a past preference shock: $\mathbb{P}\left\{\theta_{t+1}^j = C \middle| (X_t, Y_t)\right\} = \mathbb{P}\left\{\theta_{t+1}^j = C, \theta_{t+1}^j = \eta_t^j \middle| (X_t, Y_t)\right\} + \mathbb{P}\left\{\theta_{t+1}^j = C, \theta_{t+1}^j = \eta_{t+1}^j \middle| (X_t, Y_t)\right\}$ $= \mathbb{P}\left\{\eta_t^j = C \middle| (X_t, Y_t)\right\}\lambda + \mathbb{P}\left\{\eta_{t+1}^j = C\right\}(1-\lambda)$. Then voter beliefs $\mathbb{P}\left\{\eta_t^j = C \middle| (X_t, Y_t)\right\} = \frac{\mathbb{P}\left\{\eta_t^j = C, (X_t, Y_t)\right\}}{\mathbb{P}\{(X_t, Y_t)\}} = \frac{\mathbb{P}\left\{\eta_t^j = C, (X_t, Y_t) | \theta_t^j = \eta_{t-1}^j\right\}\lambda + \mathbb{P}\left\{\eta_t^j = C, (X_t, Y_t) | \theta_t^j = \eta_t^j\right\}(1-\lambda)}{\mathbb{P}\left\{(X_t, Y_t) | \theta_t^j = \eta_{t-1}^j\right\}\lambda + \mathbb{P}\left\{(X_t, Y_t) | \theta_t^j = \eta_t^j\right\}(1-\lambda)}$. When $\theta_t^j = \eta_{t-1}^j$ observed actions (X_t, Y_t) are chosen based on preference shocks at $t-1$ and earlier, so η_t^j is uncorrelated with (X_t, Y_t). When $\theta_t^j = \eta_t^j$ observed actions (X_{t-1}, Y_{t-1}) are chosen based on preference shocks at $t-1$ or earlier and so are uncorrelated with θ_t^j, η_t^j. Then $\mathbb{P}\left\{\eta_t^j = C \middle| (X_t, Y_t)\right\} = \frac{\mathbb{P}\left\{\eta_t^j = C\right\}\mathbb{P}\left\{(X_t, Y_t) | \theta_t^j = \eta_{t-1}^j\right\}\lambda + \mathbb{P}\{(X_{t-1}, Y_{t-1})\}\mathbb{P}\left\{\eta_t^j = C, (x_t, y_t) | \theta_t^j = \eta_t^j\right\}(1-\lambda)}{\mathbb{P}\left\{(X_t, Y_t) | \theta_t^j = \eta_{t-1}^j\right\}\lambda + \mathbb{P}\{(X_{t-1}, Y_{t-1})\}\mathbb{P}\left\{(x_t, y_t) | \theta_t^j = \eta_t^j\right\}(1-\lambda)}$ which because $\mathbb{P}\left\{\eta_{t+1}^j = C\right\} = \mathbb{P}\left\{\eta_t^j = C\right\} = \gamma^j$ implies that $\mathbb{P}\left\{\theta_{t+1}^j = C \middle| (X_t, Y_t)\right\} \geq \gamma^j$ iff $\frac{\mathbb{P}\left\{\eta_t^j = C, (x_t, y_t) | \theta_t^j = \eta_t^j\right\}}{\mathbb{P}\left\{(x_t, y_t) | \theta_t^j = \eta_t^j\right\}} \geq \gamma^j$. ■

Proposition 1a *Under Assumptions Ia and IIa for popular policy issues* $(\pi > 1/2)$ *hierarchical accountability either insulates the policymaker pre-election, if* $\gamma^P > \pi$*, or creates policymaker pandering incentives pre-election, if* $\gamma^P < \pi$*; direct accountability always creates policymaker pandering incentives pre-election. Agents follow their preferences post-election; in particular, an intermediary retains only a policymaker of his own type.*

Proof of Proposition 1a. For direct accountability, i.e., no intermediary and every other period it is the voter who decides whether to reelect the policymaker, Lemma 1 implies that the voter reelects the incumbent policymaker at the end of electoral term $[t-1, t]$ iff $\hat{\gamma}_t^P(x_t) \geq \gamma^P$. Thus, at $t-1$ the incumbent follows his preferences. Suppose the equilibrium has separating election-period strategies: $x_t(s_t|C) \neq x_t(s_t|D)$ for some state s_t. Then, either

$\hat{\gamma}_t^P(1) > \gamma^P > \hat{\gamma}_t^P(0)$ or $\hat{\gamma}_t^P(0) > \gamma^P > \hat{\gamma}_t^P(1)$, so the voter's best response is to condition reelection on the election-period policy x_t that improves the policymaker's reputation. By Assumption IIa both types policymaker prefer to override their preferences and comply at t, because the expected payoff from doing so is at least $\sum_{\tau=1}^{\infty} \beta^{2\tau-1}[1+\beta(1-\pi)] = \frac{\beta[1+\beta(1-\pi)]}{1-\beta^2}$ so $x_t(s_t|C) = x_t(1-s_t|C) = x_t(s_t|D) = x_t(1-s_t|D)$ and election-period strategies cannot be separating. With pooling election-period strategies both on- and off-equilibrium path beliefs are pinned down by policymaker preferences, due to the assumption of a small fraction of non-strategic agents. Preference-based reputations are $\hat{\gamma}_t^P(1) = \frac{\gamma^P\pi}{\gamma^P\pi+(1-\gamma^P)(1-\pi)} > \gamma^P > \hat{\gamma}_t^P(0) = \frac{\gamma^P(1-\pi)}{\gamma^P(1-\pi)+(1-\gamma^P)\pi}$, which cause the voter to reelect iff $x_t = 1$. The unique equilibrium then has election-period pooling on the popular policy.

For hierarchical accountability Lemma 1 implies that the voter reelects the incumbent intermediary at the end of electoral term $[t-1,t]$ iff $\hat{\gamma}_t^I(x_t, y_t) \geq \gamma^I$. Thus, at $t-1$ the incumbent intermediary follows his preferences, namely retains only policymakers of his own type. Then, the policymaker also follows his preferences at $t-1$. Suppose the equilibrium has separating election-period strategies: $y_t(x_t|C) \neq y_t(x_t|D)$ for some policy x_t. Then, either $\hat{\gamma}_t^I(x_t,1) > \gamma^I > \hat{\gamma}_t^I(x_t,0)$ or $\hat{\gamma}_t^I(x_t,0) > \gamma^I > \hat{\gamma}_t^I(x_t,1)$, so given x_t the voter's best response is to condition reelection on the retention decision y_t that improves the intermediary's reputation. Overriding the preferred retention decision at t in order to get reelected is preferred by both intermediary types because that yields positive future payoffs whereas following preferences yields zero. Thus, $y_t(x_t|C) = y_t(x_t|D)$ for $x_t = 0,1$ and election-period strategies cannot be separating. With pooling election-period strategies both on- and off-equilibrium path beliefs are pinned down by agent preferences, due to the assumption of a small fraction of non-strategic agents. Preference-based reputations are given by $\hat{\gamma}_t^I(1,1) = \frac{\gamma^I\gamma^P\pi}{\gamma^I\gamma^P\pi+(1-\gamma^I)(1-\gamma^P)(1-\pi)}(> \gamma^I)$; $\hat{\gamma}_t^I(1,0) = \frac{\gamma^I(1-\gamma^P)(1-\pi)}{\gamma^I(1-\gamma^P)(1-\pi)+(1-\gamma^I)\gamma^P\pi}(< \gamma^I)$; $\hat{\gamma}_t^I(0,1) = \frac{\gamma^I\gamma^P(1-\pi)}{\gamma^I\gamma^P(1-\pi)+(1-\gamma^I)(1-\gamma^P)\pi}$ $(> \gamma^I$ iff $\gamma^P > \pi)$; $\hat{\gamma}_t^I(0,0) = \frac{\gamma^I(1-\gamma^P)\pi}{\gamma^I(1-\gamma^P)\pi+(1-\gamma^I)\gamma^P(1-\pi)}$ $(> \gamma^I$ iff $\gamma^P < \pi)$. Therefore, if $x_t = 1$ the voter's best response is to reelect the intermediary iff the intermediary retained the policymaker; if $x_t = 0$ it is also to reelect the intermediary iff the intermediary retained the policymaker, when $\gamma^P > \pi$, but to reelect the intermediary iff the intermediary replaced the policymaker, when $\gamma^P < \pi$. The unique equilibrium can thus take two forms: at t the intermediary retains regardless of x_t, or at t the intermediary retains iff $x_t = 1$. In the first type of equilibrium the policymaker follows his preferences in an election period; in the second type of equilibrium, by the argument in the previous paragraph and Assumption IIa, both policymaker types choose the popular policy in an election period. ■

Proposition 2a *Under Assumptions Ia and IIa for neutral policy issues* $(\pi = 1/2)$ *policy-*

making outcomes do not vary with the accountability form in an election or a non-election period.

Proof of Proposition 2a. The claim follows by substituting $\pi = 1/2$ in the voter's equilibrium beliefs in the Proof of Proposition 1. Under direct accountability election-period policy x_t does not signal policymaker type θ_t^P. Assuming that when indifferent between the incumbent and the challenger the voter reelects the incumbent, the voter reelects the incumbent policymaker regardless of policy. This insulates the policymaker in an election period: $x_t(s_t|C) = s_t, x_t(s_t|D) = 1-s_t$. Under hierarchical accountability the retention signal is still informative, since $\gamma^P > 1/2$, ruling out separating election-period strategies. Since only the case $\gamma^P > \pi$ is now possible, based on Proposition 1 the unique pooling equilibrium is policymaker election-period insulation. ■

A.2 Extension: Incentives Along the Electoral Term

In this section we explore the robustness of the incentives model results in Appendix A.1 by relaxing the constraint implicit in Assumption Ia that reelection incentives affect only election-period behavior (as in Rogoff 1990, Shi and Svensson 2006). Here we use the following assumption that allows reelection incentives to affect both election and non-election period behavior (as in Martinez 2009).

Assumption Ib *Given per-electoral-term i.i.d. preference shocks* $\left(\eta_{2\tau-1}^j\right)_{\tau=1}^{\infty}$ *such that* $\eta_{2\tau-1}^j \in \{C, D\}$ *with* $\mathbb{P}\left\{\eta_{2\tau-1}^j = C\right\} = \gamma^j$*, agent* j*'s type at* $[t+1, t+2]$ *is* $\theta_{t+1}^j = \theta_{t+2}^j \in \left\{\eta_{t-1}^j, \eta_{t+1}^j\right\}$ *with* $\mathbb{P}\left\{\theta_{t+1}^j = \eta_{t-1}^j\right\} = \lambda$.

Lemma 2 *Under Assumption Ib, after a history of observable actions* $(X_t, Y_t) = (x_\tau, y_\tau)_{\tau=1}^t$ *incumbent* $j = P, I$ *reputation depends on* $[t-1, t]$ *actions:* $\gamma_{t+1}^j \equiv \mathbb{P}\left\{\theta_{t+1}^j = C|\left(X_t, Y_t\right)\right\} \geq \gamma^j$ *iff* $\hat{\gamma}_t^j\left[\left(x_{t-1}, y_{t-1}\right), \left(x_t, y_t\right)\right] \equiv \mathbb{P}\left\{\eta_{t-1}^j = C|\left(x_{t-1}, y_{t-1}\right), \left(x_t, y_t\right), \theta_{t-1}^j = \eta_{t-1}^j\right\} \geq \gamma^j$.

Proof of Lemma 2. After t, $\mathbb{P}\left\{\theta_{t+1}^j = C|\left(X_t, Y_t\right)\right\} = \mathbb{P}\left\{\theta_{t+1}^j = C, \theta_{t+1}^j = \eta_{t-1}^j|\left(X_t, Y_t\right)\right\} + \mathbb{P}\left\{\theta_{t+1}^j = C, \theta_{t+1}^j = \eta_{t+1}^j|\left(X_t, Y_t\right)\right\} = \mathbb{P}\left\{\eta_{t-1}^j = C|\left(X_t, Y_t\right)\right\}\lambda + \mathbb{P}\left\{\eta_{t+1}^j = C\right\}(1-\lambda)$. Then voter beliefs about the preference shock of electoral term $[t-1, t]$ are $\mathbb{P}\left\{\eta_{t-1}^j = C|\left(X_t, Y_t\right)\right\} = \frac{\mathbb{P}\left\{\eta_{t-1}^j = C, (X_t, Y_t)\right\}}{\mathbb{P}\{(X_t, Y_t)\}} = \frac{\mathbb{P}\left\{\eta_{t-1}^j = C, (X_t, Y_t)|\theta_{t-1}^j = \eta_{t-3}^j\right\}\lambda + \mathbb{P}\left\{\eta_{t-1}^j = C, (X_t, Y_t)|\theta_{t-1}^j = \eta_{t-1}^j\right\}(1-\lambda)}{\mathbb{P}\left\{(X_t, Y_t)|\theta_{t-1}^j = \eta_{t-3}^j\right\}\lambda + \mathbb{P}\left\{(X_t, Y_t)|\theta_{t-1}^j = \eta_{t-1}^j\right\}(1-\lambda)}$. If $\theta_{t-1}^j = \eta_{t-3}^j$ history (X_t, Y_t) is based on the preference shock at $[t-3, t-2]$ and earlier, so η_{t-1}^j is uncorrelated with (X_t, Y_t). If $\theta_{t-1}^j = \eta_{t-1}^j$ history (X_{t-2}, Y_{t-2}) is based on preference shocks at $[t-3, t-2]$ or earlier and so are uncorrelated with $\theta_{t-1}^j, \eta_{t-1}^j$. Then $\mathbb{P}\left\{\eta_{t-1}^j = C|\left(X_t, Y_t\right)\right\} = \frac{\mathbb{P}\left\{\eta_{t-1}^j = C\right\}\mathbb{P}\left\{(X_t, Y_t)|\theta_{t-1}^j = \eta_{t-3}^j\right\}\lambda + \mathbb{P}\{(X_{t-2}, Y_{t-2})\}\mathbb{P}\left\{\eta_{t-1}^j = C, (x_{t-1}, y_{t-1}), (x_t, y_t)|\theta_{t-1}^j = \eta_{t-1}^j\right\}(1-\lambda)}{\mathbb{P}\left\{(X_t, Y_t)|\theta_{t-1}^j = \eta_{t-3}^j\right\}\lambda + \mathbb{P}\{(X_{t-2}, Y_{t-2})\}\mathbb{P}\left\{(x_{t-1}, y_{t-1}), (x_t, y_t)|\theta_{t-1}^j = \eta_{t-1}^j\right\}(1-\lambda)}$ which together with the previous derivation and because $\mathbb{P}\left\{\eta_{t+1}^{j}=C\right\}=\mathbb{P}\left\{\eta_{t-1}^{j}=C\right\}=\gamma^{j}$ implies that $\mathbb{P}\left\{\theta_{t+1}^{j}=C|\left(X_{t},Y_{t}\right)\right\}\geq\gamma^{j}$ iff $\frac{\mathbb{P}\left\{\eta_{t-1}^{j}=C,(x_{t-1},y_{t-1}),(x_{t},y_{t})|\theta_{t-1}^{j}=\eta_{t-1}^{j}\right\}}{\mathbb{P}\left\{(x_{t-1},y_{t-1}),(x_{t},y_{t})|\theta_{t-1}^{j}=\eta_{t-1}^{j}\right\}}\geq\gamma^{j}$. ■

Proposition 1b *Under Assumptions Ib and IIa for popular policy issues* ($\pi>1/2$) *hierarchical accountability either insulates the policymaker pre-election, if* $\gamma^{P}\left(1-\pi\right)^{2}>\left(1-\gamma^{P}\right)\pi^{2}$, *or creates policymaker pandering incentives pre-election with probability* $\gamma^{I}(1-\pi)+\left(1-\gamma^{I}\right)\pi$, *if* $\gamma^{P}\left(1-\pi\right)^{2}<\left(1-\gamma^{P}\right)\pi^{2}$*; direct accountability creates policymaker pandering incentives pre-election with probability* $\gamma^{P}(1-\pi)+\left(1-\gamma^{P}\right)\pi$*. Agents follow their preferences post-election; in particular, an intermediary retains only a policymaker of his own type.*

Proof of Proposition 1b. For direct accountability Lemma 2 implies that the voter reelects the incumbent policymaker at the end of electoral term $[t-1,t]$ iff $\hat{\gamma}_{t}^{P}\left(x_{t-1},x_{t}\right)\geq\gamma^{P}$. Suppose in equilibrium policymaker types play separating strategies. Then, because $\pi>1/2$, there is one pair of policies (x_{t-1},x_{t}) that is chosen more often by the dissonant type. The voter's best response is to not reelect when observing that policy pair. In a state pair (s_{t-1},s_{t}) where the dissonant type prefers that policy pair forgoing it at t yields at least $\sum_{\tau=1}^{\infty}\beta^{\tau}\left(1-\pi\right)=\frac{\beta(1-\pi)}{1-\beta}$, whereas forgoing reelection yields 1 at t. Thus, by Assumption IIa, policymaker strategies have to be pooling. In a pooling equilibrium both on- and off-equilibrium path voter beliefs are pinned down by policymaker preferences, due to the assumption of a small fraction of non-strategic agents. Then, $\hat{\gamma}_{t}^{P}\left(1,1\right)=\frac{\gamma^{P}\pi^{2}}{\gamma^{P}\pi^{2}+(1-\gamma^{P})(1-\pi)^{2}}(>\gamma^{P})$; $\hat{\gamma}_{t}^{P}\left(1,0\right)=\hat{\gamma}_{t}^{P}\left(0,1\right)=\gamma^{P}$; $\hat{\gamma}_{t}^{P}\left(0,0\right)=\frac{\gamma^{P}(1-\pi)^{2}}{\gamma^{P}(1-\pi)^{2}+(1-\gamma^{P})\pi^{2}}(<\gamma^{P})$. Assuming that when indifferent between the incumbent and the challenger the voter reelects the incumbent, the voter reelects when at least one of the last two observed policies is popular. A reelection-seeking policymaker will then choose the popular policy at least once. When s_{t-1} is such that the incumbent prefers the popular policy, he chooses $x_{t-1}=1$ and then x_{t} according to his period t preferences. When s_{t-1} is such that the incumbent prefers the unpopular policy, choosing $x_{t-1}=1$ and then x_{t} according to period t preferences yields β during the term whereas choosing x_{t-1} according to period $t-1$ preferences and then $x_{t}=1$ yields $1+\beta\pi$ for a congruent type and $1+\beta\left(1-\pi\right)$ for a dissonant type. Thus, for both types it is optimal to comply later in the term. Pandering in an election period occurs when the popular policy went against policymaker period $t-1$ preferences, that is with probability $\gamma^{P}(1-\pi)+\left(1-\gamma^{P}\right)\pi$. By Assumption IIa at t both policymaker types prefer to override their election-period preferences and comply with the voter's reelection rule, because it yields $\sum_{\tau=1}^{\infty}\beta^{2\tau-1}\left\{\pi\left(1+\beta\right)+\left(1-\pi\right)\left(1+\beta\pi\right)\right\}>\sum_{\tau=1}^{\infty}\beta^{\tau}\pi=\frac{\beta\pi}{1-\pi}$ for a congruent type and $\sum_{\tau=1}^{\infty}\beta^{2\tau-1}\left\{\left(1-\pi\right)\left(1+\beta\right)+\pi\left[1+\beta\left(1-\pi\right)\right]\right\}>\sum_{\tau=1}^{\infty}\beta^{\tau}\left(1-\pi\right)=\frac{\beta(1-\pi)}{1-\beta}$ for a dissonant type.

For hierarchical accountability Lemma 2 implies that the voter reelects the incumbent intermediary at the end of electoral term $[t-1,t]$ iff $\hat{\gamma}_t^I\left[(x_{t-1},y_{t-1}),(x_t,y_t)\right] \geq \gamma^I$. Because $\gamma^P > 1/2$, given (x_{t-1},x_t) there is one pair of retention decisions (y_{t-1},y_t) that is chosen more often by the dissonant type, which makes the voter not reelect when observing that pair of intermediary reactions. But then the dissonant type would not choose that course of action in the first place because pursuing reelection yields a positive expected payoff. Thus, intermediary strategies have to be pooling. In a pooling equilibrium both on- and off-equilibrium path voter beliefs are pinned down by intermediary preferences, due to the assumption of a small fraction of non-strategic agents. Then, after two popular policies $\hat{\gamma}_t^I\left[(1,1),(1,1)\right] = \frac{\gamma^I\gamma^P\pi^2}{\gamma^I\gamma^P\pi^2+(1-\gamma^I)(1-\gamma^P)(1-\pi)^2}(> \gamma^I)$; $\hat{\gamma}_t^I\left[(1,0),(1,1)\right] = \gamma^I$; $\hat{\gamma}_t^I\left[(1,0),(1,0)\right] = \frac{\gamma^I\left(1-\gamma^P\right)^2(1-\pi)^2}{\gamma^I(1-\gamma^P)^2(1-\pi)^2+(1-\gamma^I)(\gamma^P)^2\pi^2}(< \gamma^I)$; one popular and one unpopular policy $\hat{\gamma}_t^I\left[(0,1),(1,1)\right] = \hat{\gamma}_t^I\left[(1,1),(0,1)\right] = \frac{\gamma^I\gamma^P\pi(1-\pi)}{\gamma^I\gamma^P\pi(1-\pi)+(1-\gamma^I)(1-\gamma^P)(1-\pi)\pi}(> \gamma^I)$; $\hat{\gamma}_t^I\left[(0,0),(1,0)\right] = \hat{\gamma}_t^I\left[(1,0),(0,0)\right] = \frac{\gamma^I\left(1-\gamma^P\right)^2(1-\pi)\pi}{\gamma^I(1-\gamma^P)^2(1-\pi)\pi+(1-\gamma^I)(\gamma^P)^2\pi(1-\pi)}(< \gamma^I)$; $\hat{\gamma}_t^I\left[(0,0),(1,1)\right] = \frac{\gamma^I\left(1-\gamma^P\right)\gamma^P\pi^2}{\gamma^I(1-\gamma^P)\gamma^P\pi^2+(1-\gamma^I)\gamma^P(1-\gamma^P)(1-\pi)^2}(> \gamma^I)$; $\hat{\gamma}_t^I\left[(1,0),(0,1)\right] = \frac{\gamma^I\left(1-\gamma^P\right)\gamma^P(1-\pi)^2}{\gamma^I(1-\gamma^P)\gamma^P(1-\pi)^2+(1-\gamma^I)\gamma^P(1-\gamma^P)\pi^2}(< \gamma^I)$; after two unpopular policies $\hat{\gamma}_t^I\left[(0,1),(0,1)\right] = \frac{\gamma^I\gamma^P(1-\pi)^2}{\gamma^I\gamma^P(1-\pi)^2+(1-\gamma^I)(1-\gamma^P)\pi^2}(> \gamma^I$ iff $\gamma^P(1-\pi)^2 > \left(1-\gamma^P\right)\pi^2$ if $\gamma^P > p)$; $\hat{\gamma}_t^I\left[(0,0),(0,1)\right] = \gamma^I$; $\hat{\gamma}_t^I\left[(0,0),(0,0)\right] = \frac{\gamma^I\left(1-\gamma^P\right)^2\pi^2}{\gamma^I(1-\gamma^P)^2\pi^2+(1-\gamma^I)(\gamma^P)^2(1-\pi)^2}(> \gamma^I$ iff $\gamma^P < p)$, and we set $\hat{\gamma}_t^I\left[(x_{t-1},1),(x_t,0)\right] = \hat{\gamma}_t^I\left[(x_{t-1},0),(x_t,1)\right]$ to preserve symmetry, even though $\left[(x_{t-1},1),(x_t,0)\right]$ are not consistent with the preferences of either type of intermediary. Assuming that when indifferent between the incumbent and the challenger the voter reelects the incumbent, the voter reelects (i) after two popular policies, iff the intermediary retains at least once; (ii) after one popular and one unpopular policy, iff the intermediary retains after the popular policy, and (iii) after two unpopular policies, iff the intermediary retains at least once, when $\gamma^P > p$, regardless of the intermediary's reaction, when $\gamma^P(1-\pi)^2 > \left(1-\gamma^P\right)\pi^2$, or iff the intermediary replaces at least once, when $\gamma^P(1-\pi)^2 < \left(1-\gamma^P\right)\pi^2$. It remains to show that the intermediary prefers to comply with the voter's flexible reelection rule at the end of the term rather than at the beginning of the term.

When $\gamma^P(1-\pi)^2 > \left(1-\gamma^P\right)\pi^2$ the voter's strategy is to reelect the intermediary if there is at least one retention, and in the case of only one popular policy, the retention has to follow the popular policy. If the intermediary faces a policymaker of the same type, retaining in both periods is both consistent with intermediary preferences and guarantees reelection. If however the intermediary faces a policymaker of the opposite type, he could either retain at the beginning of the term or at the end. Retaining at $t-1$ and replacing at t produces a zero payoff at t, because the policymaker will follow his preferences at t; replacing at $t-1$ and retaining at t produces a positive expected payoff at t. Thus, the intermediary always retains at the end of the term. When $\gamma^P(1-\pi)^2 < (1-\gamma^P)\pi^2$ the voter's strategy is to reelect the intermediary if there is at least one retention after popular or one replacement after unpopular, and in the case of only one popular policy, the retention has to follow the popular policy. If the intermediary faces a popular same type policymaker or an unpopular opposite type policymaker, acting on preferences in both periods results in reelection. In the opposite cases, the intermediary can either replace an unpopular same type policymaker at $t-1$ yielding $\gamma^P + \beta\left[\pi + \beta\gamma^P \mathbb{P}\left\{\theta^P_{t+1} = \theta^I_{t+1} | \theta^P_t = \theta^I_t = C\right\} + \beta\left(1-\gamma^P\right)\mathbb{P}\left\{\theta^P_{t+1} = \theta^I_{t+1}\right\}\right] + V_{t+3}$ and $\left(1-\gamma^P\right) + \beta\left[(1-\pi) + \beta\gamma^P \mathbb{P}\left\{\theta^P_{t+1} = \theta^I_{t+1}\right\} + \beta\left(1-\gamma^P\right)\mathbb{P}\left\{\theta^P_{t+1} = \theta^I_{t+1} | \theta^P_t = \theta^I_t = D\right\}\right] + V_{t+3}$ for a congruent and dissonant type respectively, or retain at $t-1$ and retain again at t iff he is popular yielding $\pi + \beta\left[\mathbb{P}\left\{\theta^P_{t+1} = \theta^I_{t+1} | \theta^P_t = \theta^I_t = C\right\} + \beta\pi\right] + V_{t+3}$ for a congruent type and $(1-\pi) + \beta\left[\mathbb{P}\left\{\theta^P_{t+1} = \theta^I_{t+1} | \theta^P_t = \theta^I_t = D\right\} + \beta(1-\pi)\right] + V_{t+3}$ for a dissonant type; the intermediary can either retain a popular opposite type policymaker at $t-1$ and replace at t, yielding zero at t, or replace at $t-1$ and retain at t iff the incoming policymaker is popular, yielding a positive expected payoff at t. If the type persistence probability $\lambda(1-\lambda)$ is large enough, the strategy of complying with the voter's reelection rule at the end of the term has a higher payoff for both intermediary types. The frequency of election-period pandering then is $\gamma^I\left[\gamma^P(1-\pi) + \left(1-\gamma^P\right)(1-\pi)\right] + \left(1-\gamma^I\right)\left[\gamma^P\pi + \left(1-\gamma^P\right)\pi\right]$ which equals $\gamma^I(1-\pi) + \left(1-\gamma^I\right)\pi$. ■

Proposition 2b *Under Assumptions Ib and IIa for neutral policy issues* ($\pi = 1/2$) *policymaking outcomes do not vary with the accountability form in an election or a non-election period.*

Proof of Proposition 2b. The claim follows by substituting $\pi = 1/2$ in the voter's equilibrium beliefs in the Proof of Proposition 1b. Under direct accountability policies (x_{t-1}, x_t) do not signal type θ^P_{t-1}. Assuming that when indifferent between the incumbent and the challenger the voter reelects the incumbent, the voter reelects the incumbent policymaker regardless of policies chosen in the electoral term $[t-1, t]$. This insulates the policymaker throughout the electoral term. Under hierarchical accountability the retention signal is still informative, since $\gamma^P > 1/2$, ruling out separating equilibria. Since only the case $\gamma^P(1-\pi)^2 > \left(1-\gamma^P\right)\pi^2$ is now possible, based on Proposition 1b the unique pooling equilibrium is policymaker insulation in both halves of an electoral term. ■

A.3 Selection Model

Assumption IIc *The discount factor β satisfies: $\beta = 0$.*

Proposition 1c *Under Assumptions Ia and IIc for popular policy issues ($\pi > 1/2$) hierarchical accountability performs policymaker selection through intermediary selection. The increase in intermediary congruence post-election is $\gamma^I_{t+1} - \gamma^I = \lambda (1-\lambda) \gamma^I \left(1-\gamma^I\right) \left(2\gamma^P - 1\right)$, when $\gamma^P > \pi$, and $\gamma^I_{t+1} - \gamma^I = \lambda (1-\lambda) \gamma^I \left(1-\gamma^I\right) (2\pi - 1)$, when $\gamma^P < \pi$. The increase in policymaker congruence is $\gamma^P_{t+1} - \gamma^P = \lambda (1-\lambda) \gamma^P \left(1-\gamma^P\right) \left(2\gamma^I - 1\right)$ post-election; pre-election it is $\gamma^P_{t+2} - \gamma^P = \lambda (1-\lambda) \gamma^P \left(1-\gamma^P\right) \left(2\gamma^I_{t+1} - 1\right)$. Direct accountability performs policymaker selection post-election only. The increase in policymaker congruence post-election is $\gamma^P_{t+1} - \gamma^P = \lambda (1-\lambda) \gamma^P \left(1-\gamma^P\right) (2\pi - 1)$.*

Proof of Proposition 1c. By Assumption IIc both policymaker types as well as intermediary types follow their preferences rather than comply with a reelection or retention rule. As a consequence, their reputations rise and fall according to $\hat{\gamma}^j_t(x_t, y_t)$ in Lemma 1. Under direct accountability the voter reelects only policymakers that chose the popular policy in the most recent election period, by Proposition 1a. Selection occurs when the policymaker's type persists between election period t and non-election period $t+1$. Then, $\gamma^P_{t+1} - \gamma^P = \mathbb{P}\left\{\theta^P_{t+1} = \theta^P_t = \eta^P_t\right\} \mathbb{P}\left\{x_t = 1\right\} \left[\hat{\gamma}^P_t(1) - \gamma^P\right]$ which, using Proposition 1a, equals $\lambda (1-\lambda) \gamma^P \left(1-\gamma^P\right) (2\pi - 1)$. Under hierarchical accountability the voter reelects only intermediaries that have retained the policymaker in the election period, if $\gamma^P > \pi$, and reelects only intermediaries that have retained a popular policymaker and replaced an unpopular policymaker in the election period, if $\gamma^P < \pi$. Intermediary selection occurs when the policymaker's type persists between election period t and non-election period $t+1$. Then, $\gamma^I_{t+1} - \gamma^I = \mathbb{P}\left\{\theta^I_{t+1} = \theta^I_t = \eta^I_t\right\} \sum_{(x_t, y_t)} \mathbb{P}\left\{x_t, y_t\right\} \left[\hat{\gamma}^I_t(x_t, y_t) - \gamma^I\right]$, where in the summation $(x_t, y_t) \in \{(0,1), (1,1)\}$ for $\gamma^P > \pi$, and $(x_t, y_t) \in \{(0,0), (1,1)\}$ for $\gamma^P < \pi$. Using the equilibrium beliefs from Proposition 1a, $\gamma^I_{t+1} - \gamma^I = \lambda (1-\lambda) \gamma^I \left(1-\gamma^I\right) \left(2\gamma^P - 1\right)$ for $\gamma^P > \pi$, and $\gamma^I_{t+1} - \gamma^I = \lambda (1-\lambda) \gamma^I \left(1-\gamma^I\right) (2\pi - 1)$ for $\gamma^P < \pi$. Policymaker selection occurs every period as a result of the intermediary's retention decision. Post-election it is $\gamma^P_{t+1} - \gamma^P = \mathbb{P}\left\{\theta^P_{t+1} = \theta^P_t = \eta^P_t\right\} \mathbb{P}\left\{y_t = 1\right\} \left[\frac{\gamma^P \gamma^I}{\gamma^P \gamma^I + (1-\gamma^P)(1-\gamma^I)} - \gamma^P\right] = \lambda (1-\lambda) \gamma^P \left(1-\gamma^P\right) \left(2\gamma^I - 1\right)$. The expression for pre-election selection $\gamma^P_{t+2} - \gamma^P$ follows by replacing γ^I with γ^I_{t+1}. Thus $\gamma^P_{t+1} > \gamma^P_{t+2} = \gamma^P$ under direct voter selection while $\gamma^P_{t+1} < \gamma^P_{t+2}$ under hierarchical selection. Because $\pi > 1/2$ a congruent policymaker is more likely to choose the popular policy, creating a positive election-period popular policy differential between hierarchical and direct accountability. ■

Proposition 2c *Under Assumptions Ia and IIc for neutral policy issues* $(\pi = 1/2)$ *hierarchical accountability performs policymaker selection through intermediary selection. The increase in intermediary congruence post-election is* $\gamma^I_{t+1} - \gamma^I = \lambda(1-\lambda)\gamma^I\left(1-\gamma^I\right)\left(2\gamma^P - 1\right)$. *The increase in policymaker congruence is* $\gamma^P_{t+1} - \gamma^P = \lambda(1-\lambda)\gamma^P\left(1-\gamma^P\right)\left(2\gamma^I - 1\right)$ *post-election; pre-election it is* $\gamma^P_{t+2} - \gamma^P = \lambda(1-\lambda)\gamma^P\left(1-\gamma^P\right)\left(2\gamma^I_{t+1} - 1\right)$. *Direct accountability performs no policymaker selection.*

Proof of Proposition 2c. For direct accountability the result follows by substituting $\pi = 1/2$ into the direct selection formula of Proposition 1c. For hierarchical accountability only the case $\gamma^P > \pi$ is now possible. Then $\gamma^I_{t+1} - \gamma^I = \lambda(1-\lambda)\gamma^I\left(1-\gamma^I\right)\left(2\gamma^P - 1\right)$ and the formulas for policymaker selection are those derived in Proposition 1c. ■

A.4 Additional Robustness Checks

Tables A1-A14 report estimation results that further probe the robustness and sensitivity of the main text's Tables 2-5 results.

Alternative Outcome Measures. Table A1 checks sensitivity to alternative measures of police employment. While in the main analysis we used *Uniform Crime Reports* counts of police employees here we use *Census of Governments* counts that disaggregate police employment into full- and part-time categories starting in 1977. The results in panel B reveal that manager governments on average employ fewer full-time police officers. Similar differentials do not exist for either part-time officers or civilian full-time or part-time employment.

An additional testable prediction of our theory model is that police employment should be more volatile in manager cities because managers match employment to the stochastic state when insulated from public opinion.[1] In Table A2 we regress the log of the absolute deviation from city mean employment, scaled by city mean employment, on manager government and additional controls. We find that variation in full-time police employment is larger under manager government, consistent with managers being more responsive to the stochastic state on popular policy issues. This result is also consistent with the finding in Table A1 that the officer differential is driven by full-time officers.[2]

In Table A3 we examine whether an employment differential is also observed in two other major areas of city government services, fire protection and public education.[3] We find little

[1] We thank an anonymous reviewer for suggesting this additional hypothesis.

[2] The difference in sample size for part-time officers between Tables A1 and A2 is due to the fact that 46 cities report a single observation for part-time officers throughout the sample period. These single within-city observations have zero deviations from city mean, which after taking the log become missing values.

[3] We thank an anonymous reviewer for this suggestion.

evidence of a robust differential for firefighters and teachers. For non-firefighters and non-teachers employees we would also expect no differences by government form, as employment in these areas should not elicit as clear a popular preference as for police officers. An exception to this pattern occurs for non-firefighters. This correlation may be due to efforts by city managers, particularly in Midwest states, to streamline personnel by using cross-trained police officers to support single-trained firefighters (this type of organization is called a "hybrid public safety model," see Miller 2009 for a review).

Government Employment and Government Spending. Prior work indicates that manager governments spend more than mayor governments (Coate and Knight 2011), yet we find that manager governments employ fewer police officers on average. In Table A4 we seek to reconcile our results with the prior literature. Four of the five specifications show that managers spend more than mayors, as in Coate and Knight (2011), but there is no significant differential in police department spending. We find mixed evidence of a differential in police officer payrolls and some evidence that managers pay police officers more, consistent with higher minimum educational qualifications used in manager governments (see Table 1) as well as with previous literature (Edwards and Edwards 1982, Deno and Mehay 1987). In sum, Table A4 suggests that the lack of a police spending differential may reflect a quantity-quality tradeoff in police employment, with managers employing fewer but more highly educated officers who receive higher compensation.

Instrumental Variables: Precipitation Shocks and Flood Risk. We use precipitation to measure the incidence of flood-related crises because comprehensive local data on the occurrence of floods does not exist in the Local Flood Control (1900-1936) period and flood occurrence may partially reflect the choice of flood control technology. To provide a first check on our identification strategy we examine whether the relationship between extreme precipitation and flood incidence was stronger before 1936 as the historical record would suggest. We use a measure of flood frequency at the local level based on soil surveys by USDA soil scientists.[4] This index is based on local flood frequency class as determined through soil survey fieldwork. Table A5 examines the relationship between soil indicated flood frequency and three candidate measures of precipitation shocks: the fraction of years that city precipitation is in the top 1st, 5th, and 10th percentile of the national precipitation distribution for the century. In principle any of these measures could represent a promising instrument. To avoid concerns about weak instruments leading to finite sample bias in our

[4] The data are based on maps of annual flood frequency regions averaged at the county level.

IV estimates we seek the strongest possible predictor of flood frequency.[5]

All LFC precipitation shock measures predict flood frequency. However, comparing the different measures indicates that the shocks based on the 1 percent definition have the most power in explaining flood frequency. Furthermore, when we add Census division fixed effects and geographic controls in columns (4)-(6) the shocks based on the 5 and 10 percent definitions have substantially less power in explaining the occurrence of floods. In contrast, the second half of the table shows that century precipitation shocks have little relationship with flood risk. This suggests that flood control technology post-1936 had become more effective in reducing flood risk (Arnold 1988).

Instrumental Variables: Validity and Exclusion Restriction. We next estimate models that probe the validity and exclusion restriction of our IV strategy. One threat to validity would be that the Flood Control Act was passed in 1936 precisely because politically powerful cities particularly suffered in the Great Flood of 1936. If these cities also employ more police today regardless of government form we would estimate a negative effect of manager government even if no effect existed. To shed light on this issue we examine whether cities affected by LFC shocks were already different in 1900 before the shocks occurred. The results in Table A6 show that cities experiencing precipitation shocks in the local flood control era are very similar across a number of city observables early in the 20th century to cities that do not.

In Table A7 we examine whether other climate shocks that have local effects but are less likely to generate infrastructure crises during the local flood control era have similar effects on manager government adoption and police employment. We present the results of reduced form models of the relationship between climate shocks and manager government, and between climate shocks and police officer employment. If our exclusion restriction is valid we would expect that the more recent Federal Flood Control (1936-1960) era (FFC) precipitation shocks would not predict either manager government or police officer employment. The results in columns (2) and (7) confirm this expectation. During the Local Flood Control era cities were often also responsible for providing water to city residents and engineering expertise was also a key input into effective water supply. Thus, cities experiencing negative precipitation shocks (droughts) are also more likely to adopt manager government.

[5] Our analysis is necessarily based on the use of national percentiles to define extreme precipitation years. If we instead used city-specific percentiles to define extreme precipitation years we would obtain essentially the same fraction of years above a given percentile cutoff for every city. This is one way in which our precipitation shock cohort approach differs from the within-country rainfall approach of Bruckner and Ciccone (2011).

We find some evidence of this effect in column (3). Lastly, we would not expect extreme temperature shocks during the local flood control era to affect either government form or police employment today.[6]

Another potential threat to our exclusion restriction is that LFC shocks might lead to changes with persistent effects on policymaking, regardless of government form. We test for clear violations of our exclusion restriction by examining whether LFC shocks lead to the adoption of other political institutions that also affect policy. The results in Table A8 reveal little relationship between LFC shocks and the three other Progressive reforms commonly associated with manager reform. These findings may be expected as these other institutions have little to do with flood control policy.

Instrumental Variables: Alternative Instruments. As precipitation shocks are by definition a relatively rare event our IV estimates may be local to the set of cities induced by LFC shocks to adopt manager government. To probe the external validity of our results in Table A9 we examine whether our estimates change with alternative versions of the instrument. We consider three different types of instrument specifications.[7]

First, instead of using the fraction of years with a positive precipitation shock we instead use only whether a city has any positive precipitation shocks during the local flood control era. We examine two candidate definitions for a positive precipitation shock, a year with precipitation ever in the top 1st percentile or top 5th percentile of the national distribution. Second, we choose different time periods when positive precipitation shocks would be manager biased. Lastly, we use negative precipitation shocks as city governments were frequently involved in the supply of water to city residents and thus droughts increased the demand for engineering skills in government in a fashion similar to flood risk (see Table A7).[8] The results in Table A9 suggest that our central results above change little when using these alternative instruments. In columns (1) to (5) the negative effect of manager government on police officer employment remains statistically significant at the 5% level in all the specifications (with the exception of column (4) where significance is at the 10% level). In columns (6) to (10) the estimates of the manager coefficient for civilian employment are less stable and

[6] In unreported analysis we have estimated the models in column (4) adding geographic controls and found a statistically insignificant relationship between LFC hot shocks and manager government.

[7] While there are many candidates for alternative instruments we focus on the ones that have sufficient power in explaining manager form of government, i.e., first-stage F-statistic close to or exceeding 10, to mitigate finite sample bias issues.

[8] We obtain additional first stage strength by adding water availablility controls: *Presence of Very Large River*, *Presence of Large River*, *Presence of Small River*, *Located on the Coast*, and *Fraction of City Area in Water (1900).*

none indicate a statistically significant negative manager effect.

Sensitivity to Sample and Inference Procedure. Table A10 reports sensitivity checks for the parsimonious OLS and IV specifications, i.e., the model that does not include Census division fixed effects and geographic controls. Panel A checks the sensitivity of our results to minor sample alterations: excluding extremely large/small cities, dropping dependent variable outliers, dropping Census divisions with few observations, excluding post-1994 years, when the federal government intervened more forcefully in local police employment through COPS grants, dropping cities with the most years of missing precipitation data, and dropping cities far away from weather stations. Overall, the OLS and the IV results maintain their prior patterns. Panel B varies the inference procedure. We examine three sets of alternative standard errors: clustering on both city and year, clustering on weather station (as this is the unit of observation for the weather variables), and accounting for spatial correlation (Conley 1999). While the OLS estimates are unchanged for both officers and civilians, and the IV estimates for officers are stable, the IV estimates for civilians remain positive and sometimes lose their statistical significance, failing to support a negative civilian employment differential similar to the officer differential.[9]

Table 2 scaled the number of police employees by population, however a more relevant scaling factor may be the crime level. A high officer-crime ratio may also better capture deviations from policy optimality, by providing a measure of "excess police." Table A10 Panel C presents estimates with employment scaled by three different crime measures. In both OLS and IV specifications manager cities have fewer officers per crime than mayor cities (though scaling by violent crime leads to imprecise estimates), but not fewer civilians per crime.

Alternative Tests for Mechanisms. In Table 4 we proxied policymaker congruence by whether the city is under VRA preclearance and the voter crime prior by the timing of the crack epidemic. In Table A11 we use the absence of past police unionization as an alternative proxy for policymaker congruence. Booth and Vespa (2013) argue that police unions have strong effects on city policy, especially in the area of police wages. One can thus view a lack of unionization as reduced interest group influence on city government. We also use an alternative measure for voter crime priors, namely the change in the city's crime

[9] To implement Conley standard errors we allow for a spatial dependence of up to 0.5 degrees latitude and longitude, which corresponds to about 65 miles, quite close to the 58 miles mean distance of a city from the nearest weather station. We do not include year effects in this model as the unbalanced nature of the panel impedes the estimation of the Conley procedure. This slight change accounts for the difference between the point estimates in row (9) and previous ones.

rate relative to its county's crime rate.[10] The estimates using these alternative proxies are similar to the ones in Table 4 and consistent with (H2). In Table A12 we examine whether electoral cycles in government spending differ by government form. The results show little evidence of a differential electoral cycle in government spending by government form in our sample, lending further support to the results in Table 5 that suggest that differential police employment cycles are not simply due to changes in government spending.

Crime Rates. In Table A13 we examine whether manager governments are associated with more crime, possibly due to their lower average officer employment. We find no evidence of a differential in violent crime rates in either OLS or IV specifications. While the OLS specification in column (3) shows a property rate differential, when we instrument for government form in column (4) it is no longer statistically significant and the point estimate is very close to zero. Recent work by Enikolopov (2011) indicates that managers face implicit incentive to increase city population growth. In Table A14 we examine whether similar incentives are in place for crime levels and whether they might differ between managers and mayors. The results reveal little relationship between salary and crime rates for either type of policymaker.

A.5 Detailed Reporting of Baseline Tables

Tables A15-A18 report the full set of coefficient estimates of the main text's Tables 2-5.

A.6 Data

i) Police Employment, Spending, and Crime

Officers per Capita: Number of sworn police officers, per 1,000 residents. *Sources:* Uniform Crime Reports (1960-2000).
Civilians per Capita: Number of civilian (non-sworn) police employees, per 1,000 residents. *Sources:* Uniform Crime Reports (1960-2000).
Police Department Spending per Capita: Total police department expenditure, per resident, in 2000$. *Sources:* Census of Governments, City Government Finances (1960-2000).
Officer Payroll per Capita: Total payroll for full-time police officers, per resident, in 2000$. *Sources:* Census of Governments, City Government Finances (1977-2000).

[10] We thank an anonymous reviewer for suggesting this alternative measure.

Officer Avg. Income: Average pay for a full-time police officer, in 2000$. *Sources:* Census of Governments, City Government Finances (1977-2000).

Officer Educ. Min.: Dummy variable equal to 1 if the police department maintains a high school minimal educational requirement for officers, and 0 otherwise. *Sources:* Law Enforcement Management and Administrative Statistics, (1993, 1997, and 2000).

Police Unionization 1968: Fraction of the police department unionized in 1968. *Source:* Municipal Year Book (1969).

Violent Crime Rate: Total number of murder, rape, and robbery crimes reported, per 1,000 residents. *Sources:* Uniform Crime Reports (1975-2000).

Property Crime Rate: Total number of motor vehicle theft, larceny, burglary, and assault crimes reported, per 1,000 residents. *Sources:* Uniform Crime Reports (1975-2000).

Total Crime Rate: Sum of violent and property crime rates. *Sources:* Uniform Crime Reports (1975-2000).

ii) 1900 City Outcomes

Sources: U.S. Census Bureau (1905, 1906).

Officers per Capita: Number of sworn police officers, per 1,000 residents.

Civilians per Capita: Number of civilian (non-sworn) employees, per 1,000 residents.

Arrests per Capita: Number of police department arrests, per 1,000 residents.

Miles of Paved Roads per Square Mile: Miles of paved road per square mile.

Miles of Sewers per Square Mile: Miles of sewers per square mile.

Population Size: Number of city residents.

iii) Policymaker Characteristics

Annual Salary: Salary of manager or mayor, in 2000$. *Sources:* ICMA Salaries of Municipal Officials (1992-1993, 1995-2000).

Male: Manager or mayor is male. *Sources:* ICMA Salaries of Municipal Officials (1992-1993,1995-2000).

White: Manager or mayor is white. *Sources:* ICMA Salaries of Municipal Officials (1992-1993,1995-2000).

iv) City Institutions

Manager: Dummy variable equal to 1 if the city has a manager form of government, and 0 otherwise. *Sources:* Municipal Year Book (1960-2000).

Election: Indicator variable equal to 1 in an election year, and 0 otherwise. If the election takes place before July 31 the previous year is coded as an election year. Election years are coded based on mayor elections and in case no mayor office exists in a manager city, based council elections. *Sources:* Municipal Year Book (various years), World Almanac and Book of Facts (various years), www.ourcampaigns.com, city charters, newspaper articles.

Non-Partisan Elections: Indicator variable equal to 1 if the city charter in effect in 1960 mandates non-partisan elections, and 0 otherwise. *Source:* Municipal Year Book (1960).

Fraction At-Large Seats: Fraction of city council seats elected at-large in 1960. *Source:* Municipal Year Book (1960).

Early Civil Service: Indicator variable equal to 1 if the city has a non-political civil service before 1937, and 0 otherwise. *Source:* Civil Service Assembly (1938).

City Government Spending Per Capita: Total city government expenditure, in thousands of 2000$, per capita. *Sources:* Census of Governments, City Government Finances (1960-2000).

Incumbent Mayor Republican: Takes a value of 1 if the incumbent mayor's party is Republican, and 0 otherwise. Most manager cities maintain a separate and honorary office of the mayor. *Sources:* Ferreira and Gyourko (2009).

Local Newspaper Sales: Average daily sales of local newspapers, per capita, 1990-2000. *Source:* George and Waldfogel (2006).

v) Other City Government Employment

Sources: U.S. Census Bureau, Census of Governments (1977-2000).

Full-Time Officers per Capita: Full-time police officers, per 1,000 residents.

Part-Time Officers per Capita: Part-time police officers, per 1,000 residents.

Full-Time Civilians per Capita: Full-time police civilian employees, per 1,000 residents.

Part-Time Civilians per Capita: Part-time police civilian employees, per 1,000 residents.

Firefighters per Capita: Total firefighters per 1,000 residents.

Non-Firefighters per Capita: Total fire department employees minus firefighters, per 1,000 residents.

Teachers per Capita: Elementary education teachers, per 1,000 residents.

Non-Teachers per Capita: Elementary education employees minus teachers, per 1,000 residents.

vi) Demographics

Population Size: Number of city residents, in thousands. Based on 1960, 1970, 1980, 1990, and 2000 Census of Population numbers, linearly interpolated in intercensal years. *Sources:* U.S. Census Bureau, Census of Population (various years).

Population Density: Number of city residents, in thousands, divided by square miles of city area. Population based on 1960, 1970, 1980, 1990, and 2000 Census of Population numbers, linearly interpolated in intercensal years. City area based on 1900 and 2000 Census numbers, linearly interpolated between 1900 and 2000. *Sources:* Authors' calculations.

Fraction Non-White: Fraction of city population that are non-white. Based on 1960, 1970, 1980, 1990, and 2000 Census of Population numbers, linearly interpolated in intercensal years. *Sources:* U.S. Census Bureau, Census of Population (various years).

Fraction College Graduate: Fraction of city population that are college graduates. Based on 1960, 1970, 1980, 1990, and 2000 Census of Population numbers, linearly interpolated in intercensal years. *Sources:* U.S. Census Bureau, Census of Population (various years).

Household Income: Median household income of city residents, in 2000$. Based on 1960, 1970, 1980, 1990, and 2000 Census of Population numbers, linearly interpolated in intercensal years. *Sources:* U.S. Census Bureau, Census of Population (various years).

vii) Mechanisms

Voting Rights Act Preclearance: Dummy variable equal to 1 if the city is identified by Section 5 of the Voting Rights Act as a covered jurisdiction under the formula of Section 4, and 0 otherwise. Covered jurisdictions have to obtain court or federal preclearance before enacting any change to their election regulations. The first element in the formula is that the state or political subdivision of the state maintained on November 1, 1964, a "test or device," restricting the opportunity to register and vote. The second element of the formula would be satisfied if the Director of the Census determined that less than 50 percent of persons of voting age were registered to vote on November 1, 1964, or that less than 50 percent of persons of voting age voted in the presidential election of November 1964. *Source:* U.S. Department of Justice.

Crack Epidemic: Dummy variable equal to 1 if crack cocaine arrived in the state in a prior year, based on cocaine-related deaths, and 0 otherwise. *Source:* Evans, Garthwaite, and Moore (2012).

No Police in Union 1968: Dummy variable equal to 1 if the fraction of the police department unionized in 1968 is zero, and 0 otherwise. *Source:* Municipal Year Book (1969).

High Δ Relative Crime Rate: Dummy variable equal to 1 if the change in city to county total crime rate from last year is above median. *Sources:* Authors' calculations based on data in Uniform Crime Reports (1975-2000).

viii) Weather

Sources: The U.S. Historical Climatology Network's *Daily Temperature, Precipitation, and Snow Data* contains daily readings for precipitation, snowfall, and temperature extremes collected from weather stations throughout the U.S. We construct yearly variables based on this dataset.

LFC Precipitation Shocks: Fraction Local Flood Control years (1900-1936) with annual precipitation in the top 1 percent of the national precipitation distribution.

FFC Precipitation Shocks: Fraction of Federal Flood Control years (1937-1960) with annual precipitation in the top 1 percent of the national precipitation distribution.

LFC Drought Shocks: Fraction Local Flood Control years (1900-1936) with annual precipitation in the bottom 1 percent of the national precipitation distribution.

Century Precipitation Shocks: Fraction of 1900-2000 years with annual precipitation in the top 1 percent of the national precipitation distribution.

Century Drought Shocks: Fraction of 1900-2000 years with annual precipitation in the bottom 1 percent of the national precipitation distribution.

Median Precipitation: Median annual precipitation from 1900-2000.

LFC Hot Shocks: Fraction Local Flood Control years (1900-1936) with annual high temperature in the top 1 percent of the national high temperature distribution.

LFC Cold Shocks: Fraction Local Flood Control years (1900-1936) with annual high temperature in the bottom 1 percent of the national high temperature distribution.

Century Hot Shocks: Fraction of 1900-2000 years with annual high temperature in the top 1 percent of the national high temperature distribution.

Century Cold Shocks: Fraction of 1900-2000 years with annual low temperature in the bottom 1 percent of the national low temperature distribution.

Median High Temperature: Median annual high temperature from 1900-2000.

Median Low Temperature: Median annual low temperature from 1900-2000.

ix) Geography

Sources: Fishback, Horrace, and Kantor (2005, 2006). The data are reported at the county level. We match it to our sample cities.

Elevation Minimum: The minimum elevation in the county.

Elevation Maximum: The maximum elevation in the county.

Latitude: Latitude of the county seat.

Longitude: Longitude of the county seat.

Distance to Nearest River: Minimum distance to a county with a small, large, or very large river.

Presence of Very Large River: County has a river that goes through more than 50 counties.

Presence of Large River: County has a river that goes through 21 to 50 counties.

Presence of Small River: County has a river that goes through 11 to 20 counties.

Presence of Swamp: County has a swamp.

Located on the Coast: County is located on the coast.

Percentage of Clay in the Soil: Estimated fraction of clay in the soil, based on contemporary surveys by USDA soil scientists.

Soil Indicated Flood Frequency Index: Average flood class of the county standardized to a variable with a mean of zero and a standard deviation of one, based on contemporary surveys by USDA soil scientists.

A.7 Appendix References

Arnold, J.L. (1988) *The Evolution of the 1936 Flood Control Act*, Fort Belvoir, VA: Office of History, U.S. Army Corps of Engineers.

Booth, R. and E. Vespa (2013) "Public Unions and Policy Persistence: Theory and Evidence from US Municipalities," working paper, UC Santa Barbara.

Civil Service Assembly (1938) "Civil Service Agencies in the United States: A 1937 Census," Pamphlet No. 11, January.

Conley, T.G. (1999) "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics* 92: 1-45.

Enikolopov, R. (2011) "Are Bureaucrats Really Paid Like Bureaucrats?" working paper, New Economic School.

Evans, W., C. Garthwaite, and T.J. Moore (2012) "The White/Black Educational Gap, Stalled Progress, and the Long Term Consequences of the Emergence of Crack Cocaine Markets," working paper, Northwestern University.

Federal Bureau of Investigation, *Law Enforcement Management and Administrative Statistics*.

Federal Bureau of Investigation, *Uniform Crime Reports.*

Ferreira, F. and J. Gyourko (2009) "Do Political Parties Matter? Evidence From U.S. Cities," *Quarterly Journal of Economics* 124(1): 399-422.

Fishback, P.V., W.C. Horrace, and S. Kantor (2005) "Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales During the Great Depression," *Journal of Economic History* 65(1): 6-71.

Fishback, P.V., W.C. Horrace, and S. Kantor (2006) "The Impact of New Deal Expenditures on Mobility During the Great Depression," *Explorations in Economic History* 43(2): 179-222.

George, L.M. and J. Waldfogel (2006) "The New York Times and the Market for Local Newspapers," *American Economic Review* 96(1): 435-447.

International County/City Management Association, *Municipal Year Book.*

International County/City Management Association, *Salaries of Municipal Officials.*

Miller, M. (2009) "All-Hazards Approach and Cross-Trained Responders," *Law and Order* 57(10): 58-61.

Newspaper Enterprise Association, *World Almanac and Book of Facts.*

U.S. Census Bureau (1905) "The Statistics of Cities Having a Population of Over 25,000: 1902 and 1903," Bulletin 20.

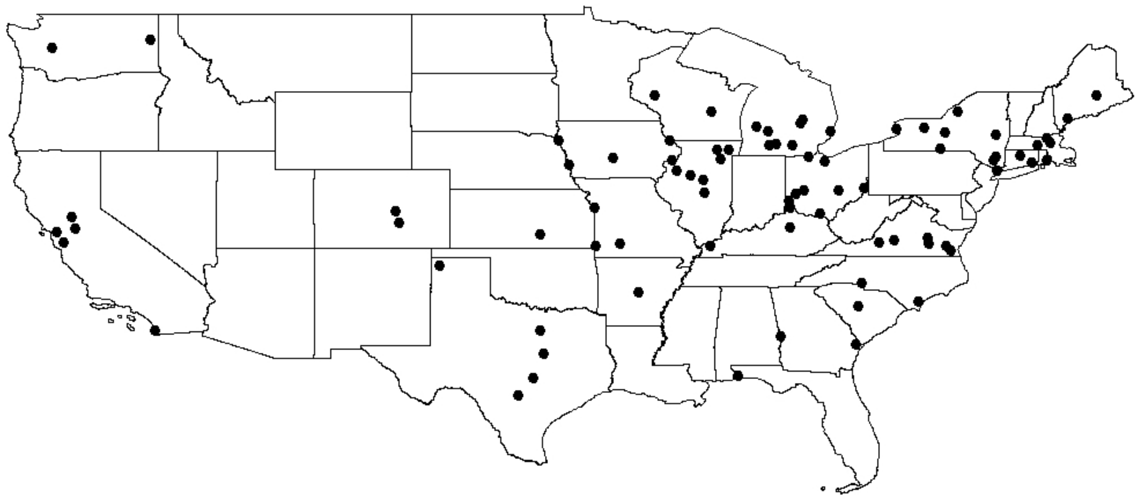
U.S. Census Bureau (1906) "The Statistics of Cities Having a Population 8,000 to 25,000: 1903," Bulletin 45.

U.S. Census Bureau, *Census of Population.*

U.S. Census Bureau, *Census of Governments.*

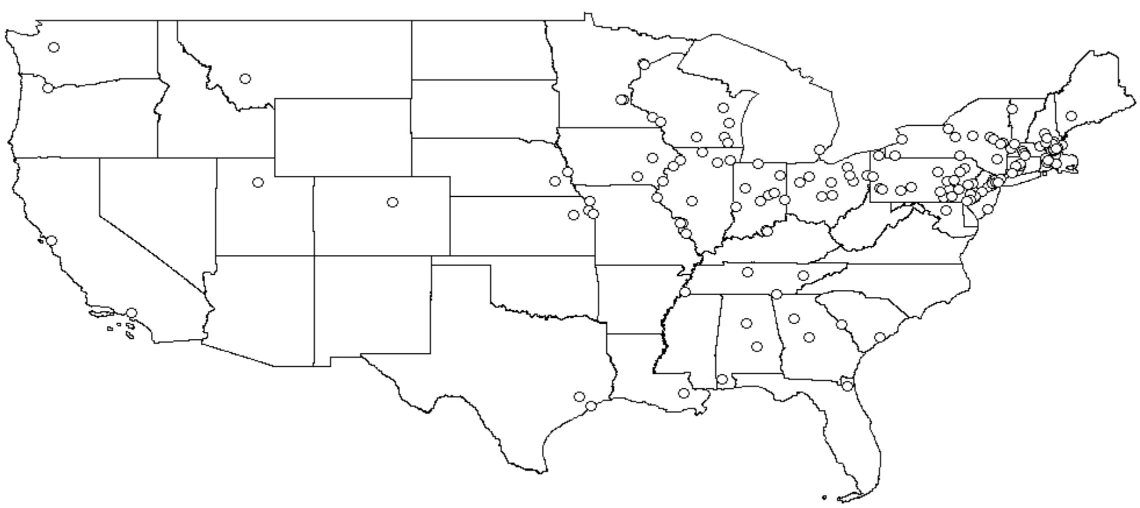
Williams, C.N., M.J. Menne, R.S. Vose, and D.R. Easterling (2006) "U.S. Historical Climatology Network Daily Temperature, Precipitation, and Snow Data," Oak Ridge National Laboratory.

FIGURE A1: Manager Cities, 1960



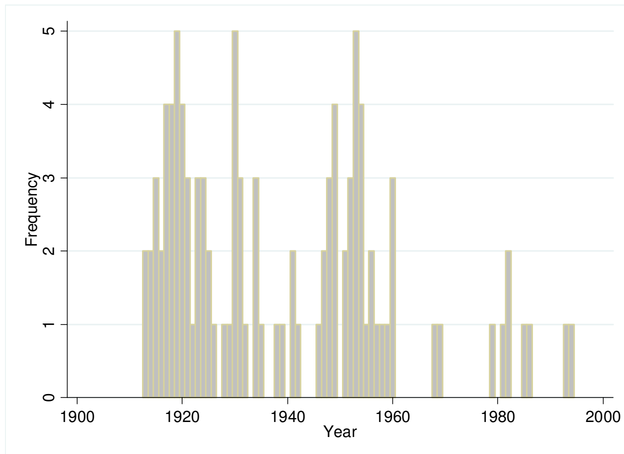
Notes: Authors' calculations for the sample of the 248 largest self-governing cities as of 1900. Manager government locations are plotted based on 1960 form of government. The map reflects state jurisdictional boundaries.

FIGURE A2: Mayor Cities, 1960



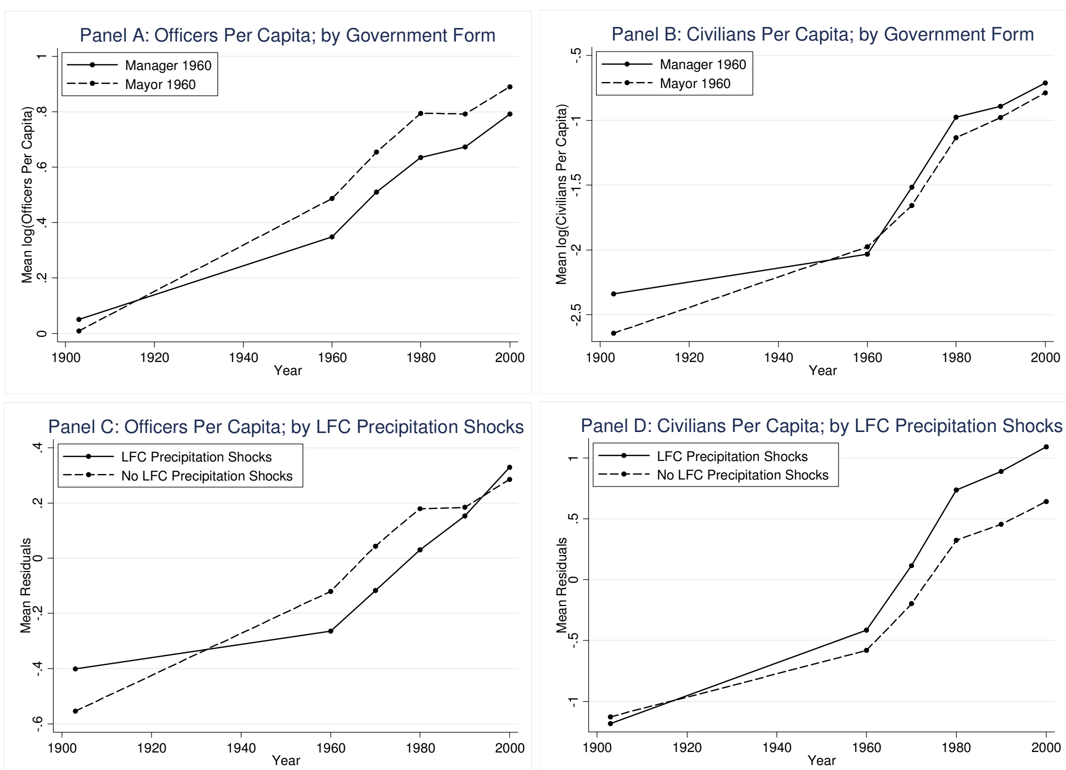
Notes: Authors' calculations for the sample of the 248 largest self-governing cities as of 1900. Mayor government locations are plotted based on 1960 form of government. The map reflects state jurisdictional boundaries.

FIGURE A3: Timing of Manager Charter Adoptions



Notes: Authors' tabulations using adoption data from the *Municipal Year Book*, newspapers, and city charters for the sample of the 248 largest self-governing cities as of 1900. Year of manager government adoption is based on year of charter approval by voters.

FIGURE A4: Police Employment, by Indicated Cohort



Notes: Authors' calculations for the sample of the 248 largest self-governing cities as of 1900. Panels C and D plot the mean residual from a regression of *log(Police Per Capita)* on *Century Precipitation Shocks* and *Median Precipitation 1900-2000*, where *Police* is either *Officers* or *Civilians*.

TABLE A1: Government Form and Police Employment: Levels

Dependent Variable =	Log(Officers per Capita)					Log(Civilians per Capita)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Total (Uniform Crime Reports, 1960-2000)									
Manager	-0.16***	-0.08**	-0.07***	-0.05*	-0.07**	0.09	0.02	0.03	-0.02	-0.06
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.07)	(0.07)	(0.06)	(0.07)	(0.08)
Number of Clusters	248	248	248	248	236	248	248	248	248	235
Number of Observations	9,974	9,974	9,974	9,974	1,164	9,850	9,850	9,850	9,850	1,156
	Panel B: Full-Time (Census of Governments, 1977-2000)									
Manager	-0.15***	-0.09***	-0.07***	-0.06**	-0.09***	0.04	-0.02	0.00	-0.05	-0.07
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.07)	(0.07)	(0.06)	(0.07)	(0.09)
Number of Clusters	246	246	246	246	230	245	245	245	245	230
Number of Observations	5,462	5,462	5,462	5,462	952	5,309	5,309	5,309	5,309	940
	Panel C: Part-Time (Census of Governments, 1977-2000)									
Manager	-0.69**	-0.20	-0.09	0.00	0.30	-0.40***	-0.05	-0.04	0.04	0.12
	(0.32)	(0.28)	(0.24)	(0.26)	(0.64)	(0.14)	(0.14)	(0.14)	(0.15)	(0.21)
Number of Clusters	148	148	148	148	61	245	245	245	245	200
Number of Observations	679	679	679	679	134	4,183	4,183	4,183	4,183	732
Additional Controls:										
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Demographic	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Institutional & Political	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Policymaker	No	No	No	No	Yes	No	No	No	No	Yes
Sample Years	Full	Full	Full	Full	1992+	Full	Full	Full	Full	1992+

Notes: Authors' calculations with city data as described in Appendix A.6. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. The table reports OLS estimates of β_1 from equation (1) in the text. Standard errors clustered at the city level in parentheses. Columns (2)-(5),(7)-(10) include Census division fixed effects and geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction, *Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, *Percentage of Clay in Soil*. Columns (3)-(5),(8)-(10) also include demographic controls: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and *Population Density Missing*. Columns (4)-(5),(9)-(10) also include controls for *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Police Unionization 1968*, and *Local Newspaper Sales 1990-2000,* as well as indicator variables for missing values for each variable. Columns (5),(10) also control for *log(Policymaker Salary)*, *Policymaker Male,* and *Policymaker White.* * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A2: Government Form and Police Employment: Volatility

Dependent Variable =	Log(Officers per Capita Coefficient of Variation)					Log(Civilians per Capita Coefficient of Variation)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Total (Uniform Crime Reports, 1960-2000)									
Manager	0.12**	0.06	0.06	0.03	0.12	0.02	0.06	0.05	0.07	0.14
	(0.06)	(0.06)	(0.06)	(0.06)	(0.14)	(0.06)	(0.06)	(0.06)	(0.06)	(0.14)
Number of Clusters	248	248	248	248	236	248	248	248	248	235
Number of Observations	9,974	9,974	9,974	9,974	1,164	9,850	9,850	9,850	9,850	1,156
	Panel B: Full-Time (Census of Governments, 1977-2000)									
Manager	0.17**	0.18**	0.21***	0.17**	0.21	-0.17**	-0.08	-0.08	-0.06	-0.06
	(0.07)	(0.07)	(0.07)	(0.07)	(0.15)	(0.08)	(0.09)	(0.09)	(0.10)	(0.14)
Number of Clusters	246	246	246	246	230	245	245	245	245	230
Number of Observations	5,462	5,462	5,462	5,462	952	5,309	5,309	5,309	5,309	940
	Panel C: Part-Time (Census of Governments, 1977-2000)									
Manager	0.18	-0.25	-0.44	-0.32	-0.40	0.26**	0.18	0.18	0.07	0.20
	(0.33)	(0.38)	(0.38)	(0.41)	(0.55)	(0.12)	(0.13)	(0.12)	(0.13)	(0.19)
Number of Clusters	102	102	102	102	55	243	243	243	243	200
Number of Observations	633	633	633	633	128	4,181	4,181	4,181	4,181	732
Additional Controls:										
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Demographic	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Institutional & Political	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Policymaker	No	No	No	No	Yes	No	No	No	No	Yes
Sample Years	Full	Full	Full	Full	1992+	Full	Full	Full	Full	1992+

Notes: Authors' calculations with city data as described in Appendix A.6. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. Dependent variable is calculated as: $\log[abs(Police_{i,t} - Avg\_Police_i) / Avg\_Police_i]$. OLS estimates with standard errors clustered at the city level in parentheses. Columns (1) and (6) only include year fixed effects. Columns (2) and (7) also include Census division fixed effects and geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction, *Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil*. Columns (3) and (8) also include demographic controls: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and *Population Density Missing*. Columns (4) and (9) also include controls for *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Police Unionization 1968*, and *Local Newspaper Sales 1990-2000,* as well as indicator variables for missing values for each variable. Columns (5) and (10) also include *log(Policymaker Salary)*, *Policymaker Male,* and *Policymaker White.* * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A3: Government Form and Other City Employees: Levels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Police (Census of Governments, 1977-2000)									
Dependent Variable =	Log(Officers per Capita)					Log(Civilians per Capita)				
Manager	-0.16***	-0.09***	-0.07***	-0.07***	-0.10***	-0.13*	-0.06	-0.05	-0.08	-0.11
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)
Number of Clusters	246	246	246	246	230	246	246	246	246	230
Number of Observations	5,462	5,462	5,462	5,462	952	5,364	5,364	5,364	5,364	945
	Panel B: Fire (Census of Governments, 1977-2000)									
Dependent Variable =	Log(Firefighters per Capita)					Log(Non-Firefighters per Capita)				
Manager	-0.12**	-0.02	-0.02	-0.01	-0.09	-0.24***	-0.24***	-0.22**	-0.20**	-0.24*
	(0.05)	(0.03)	(0.03)	(0.04)	(0.05)	(0.09)	(0.08)	(0.08)	(0.09)	(0.13)
Number of Clusters	246	246	246	246	230	239	239	239	239	218
Number of Observations	5,456	5,456	5,456	5,456	951	4,790	4,790	4,790	4,790	865
	Panel C: Education (Census of Governments, 1977-2000)									
Dependent Variable =	Log(Teachers per Capita)					Log(Non-Teachers per Capita)				
Manager	0.05	0.03	0.01	0.01	-0.01	0.11	0.06	0.04	0.00	0.06
	(0.04)	(0.05)	(0.04)	(0.04)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.11)
Number of Clusters	85	85	85	85	63	86	86	86	86	63
Number of Observations	1,604	1,604	1,604	1,604	240	1,605	1,605	1,605	1,605	240
Additional Controls:										
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Demographic	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Institutional & Political	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Policymaker	No	No	No	No	Yes	No	No	No	No	Yes
Sample Years	Full	Full	Full	Full	1992+	Full	Full	Full	Full	1992+

Notes: Authors' calculations with 1977-2000 city data as described in Appendix A.6, except for columns (5),(10) where the sample is 1992-1993, 1995-2000. The panel consists of the largest 248 self-governing cities in the United States as of 1900. Standard errors clustered at the city level. Columns (2)-(5),(7)-(10) include Census division fixed effects and the geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction, *Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil*. Columns (3)-(5),(6)-(10) also include: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and *Population Density Missing*. Columns (4)-(5),(9)-(10) also include *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Police Unionization 1968*, and *Local Newspaper Sales 1990-2000,* and missing values indicators. Columns (5) and (10) also include *log(Policymaker Salary)*, *Policymaker Male,* and *Policymaker White.* * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A4: Government Form and Government Spending: Reconciling with Coate and Knight (2011)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable =	Log(City Government Spending per Capita)					Log(Officer Payroll per Capita)				
Manager	0.28*	0.35**	0.36**	0.39**	-0.09	-0.14***	-0.05	-0.03	-0.02	-0.09**
	(0.15)	(0.15)	(0.16)	(0.18)	(0.20)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)
Sample Years	60-00	60-00	60-00	60-00	1992+	77-00	77-00	77-00	77-00	1992+
Number of Clusters	248	248	248	248	236	246	246	246	246	230
Number of Observations	10,168	10,168	10,168	10,168	1,164	5,462	5,462	5,462	5,462	952
Dependent Variable =	Log(Police Department Spending per Capita)					Log(Officer Average Pay)				
Manager	-0.01	0.04	0.05	0.04	0.06	0.01	0.04**	0.04**	0.03**	0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.09)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Sample Years	60-00	60-00	60-00	60-00	1992+	77-00	77-00	77-00	77-00	1992+
Number of Clusters	239	239	239	239	230	246	246	246	246	230
Number of Observations	9,706	9,706	9,706	9,706	1,124	5,462	5,462	5,462	5,462	952
Additional Controls:										
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Demographic	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Institutional & Political	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Policymaker	No	No	No	No	Yes	No	No	No	No	Yes

Notes: Authors' calculations with 1960-2000 or 1977-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. Standard errors clustered at the city level reported in parentheses. Columns (2)-(5),(7)-(10) include Census division fixed effects and geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction*, Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil*. Columns (3)-(5),(8)-(10) also include: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and *Population Density Missing*. Columns (4)-(5),(9)-(10) also include the following demographic controls: *Population Density, Fraction of Population Non-White, Fraction of Population College Graduate*, *Median Household Income,* and *Population Density Missing*. Columns (5),(10) also include controls for *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Low Police Unionization 1968*, and *High Local Newspaper Sales 1990-2000,* and missing values indicators. Columns (5),(10) also control for *log(Policymaker Salary)*, *Policymaker Male,* and *Policymaker White.** indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A5: Precipitation Shocks and Soil Indicated Flood Frequency: OLS Estimates

Dependent Variable =	Soil Indicated Flood Frequency Index					
Percentile x =	1%	5%	10%	1%	5%	10%
	(1)	(2)	(3)	(4)	(5)	(6)
LFC Precipitation Shocks (*x* Percentile)	31.79***	8.67***	8.79***	16.68**	1.72	3.09
	(6.91)	(2.08)	(1.48)	(7.26)	(2.63)	(1.99)
Century Precipitation Shocks (*x* Percentile)	-3.89	-3.35*	-4.85***	1.52	2.17	-0.38
	(5.09)	(1.79)	(1.36)	(5.83)	(2.42)	(2.06)
Median Precipitation	0.69	0.60	0.14	-2.55	-3.20*	-3.32
	(0.86)	(1.01)	(1.08)	(1.66)	(1.90)	(2.11)
Additional Controls:						
Division & Geography	No	No	No	Yes	Yes	Yes
Sample	Cross-Section	Cross-Section	Cross-Section	Cross-Section	Cross-Section	Cross-Section
Number of Observations	248	248	248	248	248	248

Notes: Authors' calculations with the city data described in Appendix A.6. The unit of observation is a city for the sample of the largest 248 self-governing cities in the United States as of 1900. Each column reports the results from one regression. Standard errors reported in parentheses. The models in columns (1)-(3) include no additional controls. The models in columns (4)-(6) also include Census division fixed effects and the following geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude,* a *Latitude-Longitude* interaction, *Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil*. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A6: Early (1900) City Characteristics and Precipitation Shocks: OLS Estimates

Dependent Variable =	Log (Officers per Capita)	Log (Civilians per Capita)	Log (Arrests per Capita)	Log (Miles of Paved Roads per Sq. Mile)	Log (Miles of Sewers per Sq. Mile)	Log (Population)	Distance to Nearest River
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LFC Precipitation Shocks	0.09	-3.73	7.11	-17.25	-12.23	-3.01	-185.85
	(2.60)	(7.20)	(6.26)	(10.73)	(9.44)	(5.69)	(276.73)
Century Precipitation Shocks	2.77	1.79	4.54	-11.22	-2.82	3.51	496.30*
	(1.92)	(5.04)	(4.85)	(7.54)	(5.93)	(4.13)	(192.60)
Median Precipitation	0.38	-0.35	-0.83	5.80***	1.38	-0.10	-102.85***
	(0.34)	(0.72)	(0.58)	(1.72)	(0.94)	(0.78)	(28.65)
Sample	Cross-Section	Cross-Section	Cross-Section	Cross-Section	Cross-Section	Cross-Section	Cross-Section
Number of Observations	248	172	246	246	244	248	248

Notes: Authors' calculations with the city data described in Appendix A.6. The unit of observation is a city for the sample of the largest 248 self-governing cities in the United States as of 1900. Standard errors reported in parentheses. Each column reports the results from one regression with no additional controls. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A7: Weather Shocks, Government Form, and Police Employment: OLS Estimates

Dependent Variable =	Manager					Log (Officers per Capita)				
Weather =	High Precip.		Low Precip.	High Temp.	Low Temp.	High Precip.		Low Precip.	High Temp.	Low Temp.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
LFC Weather Shocks	10.14***	10.18***	3.64*	-10.12**	1.70	-4.91**	-4.62**	0.01	2.91	-0.19
	(3.04)	(3.03)	(1.90)	(4.48)	(1.03)	(2.11)	(2.01)	(1.12)	(2.75)	(0.36)
FFC Weather Shocks	–	0.32	–	–	–	–	2.26	–	–	–
		(2.39)					(1.68)			
Century Weather Shocks	-6.03***	-6.37**	-4.46**	10.42**	-1.95*	2.68	0.26	0.49	-2.87	0.11
	(2.20)	(3.05)	(2.14)	(4.72)	(1.06)	(1.63)	(2.42)	(1.28)	(2.90)	(0.38)
Median Weather	-0.61	-0.61	-1.17***	1.39**	0.54	0.75***	0.74***	1.09***	-0.87***	0.19
	(0.46)	(0.46)	(0.38)	(0.51)	(0.58)	(0.24)	(0.24)	(0.21)	(0.27)	(0.34)
Sample	Cross-Sc.	Cross-Sc.	Cross-Sc.	Cross-Sc.	Cross-Sc.	Cross-Sc.	Cross-Sc.	Cross-Sc.	Cross-Sc.	Cross-Sc.
Number of Observations	248	248	248	248	248	243	243	243	243	243

Notes: Authors' calculations with the city data described in Appendix A.6. The unit of observation is a city for the sample of the largest 248 self-governing cities in the United States as of 1900. Each column reports the results from one regression. Standard errors reported in parentheses. The models include no additional controls. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A8: Precipitation Shocks and Other City Institutions: OLS Estimates

Dependent Variable =	Manager 1960	Non-Partisan Elections 1960	Fraction Seats At-Large 1960	Early Civil Service
	(1)	(2)	(3)	(4)
LFC Precipitation Shocks	10.14***	-2.79	2.76	-3.88
	(3.04)	(2.73)	(2.83)	(3.68)
Century Precipitation Shocks	-6.03***	7.67***	-0.86	1.65
	(2.20)	(2.02)	(2.19)	(2.63)
Median Precipitation	-0.61	-1.02***	0.27	-0.83**
	(0.46)	(0.35)	(0.41)	(0.34)
Sample	Cross-Section	Cross-Section	Cross-Section	Cross-Section
Number of Observations	248	247	247	248

Notes: Authors' calculations with the city data described in Appendix A.6. The unit of observation is a city from the sample of the largest 248 self-governing cities in the United States as of 1900. Standard errors reported in parentheses. The models include no additional controls. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A9: Government Form and Police Employment: Alternative IV Estimates

Specification=	Precipitation Shock Definition		Reform Era Definition		Drought IV	Precipitation Shock Definition		Reform Era Definition		Drought IV
Shock Definition=	Any Year Top 1 %	Any Year Top 5 %	1909-1936	1900-1929	Bottom 1 %	Any Year Top 1 %	Any Year Top 5 %	1909-1936	1900-1929	Bottom 1 %
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Panel A: Second Stage:					
	Dependent Variable = Log (Officers per Capita)					*Dependent Variable* = Log (Civilians per Capita)				
Manager	-0.33**	-0.44**	-0.27**	-0.30*	-0.63**	1.11**	-0.05	0.83**	0.74*	-0.70
	(0.14)	(0.20)	(0.13)	(0.16)	(0.28)	(0.44)	(0.35)	(0.37)	(0.39)	(0.48)
Century Weather Shocks	0.16**	0.15***	-0.26	-0.27	-0.11	0.16	0.30***	7.18***	7.13**	0.18
	(0.05)	(0.05)	(0.86)	(0.89)	(0.18)	(0.14)	(0.10)	(2.06)	(1.89)	(0.32)
Median Weather	-0.11	-0.23	0.56***	0.54**	-0.07	-0.57	-2.30***	-1.28**	-1.37**	-1.90**
	(0.26)	(0.32)	(0.22)	(0.23)	(0.36)	(0.87)	(0.63)	(0.64)	(0.64)	(0.59)
					Panel B : First Stage:					
					Dependent Variable = Manager					
Weather Shocks	0.50***	0.31***	12.04***	8.41***	5.01***	0.50***	0.31***	12.03***	8.42***	4.98***
	(0.12)	(0.07)	(3.09)	(2.39)	(1.54)	(0.12)	(0.08)	(3.09)	(2.38)	(1.54)
Century Weather Shocks	-0.10	-0.19**	-5.05***	-6.15***	-5.98***	-0.10	-0.19**	-5.08***	-6.20***	-5.95***
	(0.08)	(0.08)	(1.78)	(2.15)	(1.73)	(0.08)	(0.08)	(1.78)	(2.15)	(1.73)
Median Weather	-0.98**	-1.28**	-0.74	-0.57	-1.13***	-0.98**	-1.28**	-0.74*	-0.57	-1.13***
	(0.47)	(0.52)	(0.43)	(0.45)	(0.36)	(0.47)	(0.53)	(0.43)	(0.45)	(0.36)
<u>Additional Controls:</u>										
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	No	No	No	No	Partial	No	No	No	No	Partial
F-statistic	17.34	17.61	15.22	12.44	10.54	17.17	17.48	15.15	12.48	10.45
[p-value]	[0.0000]	[0.0000]	[0.0001]	[0.0005]	[0.0013]	[0.0000]	[0.0000]	[0.0001]	[0.0005]	[0.0001]
Sample Years	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full
Number of Clusters	248	248	248	248	248	248	248	248	248	248
Number of Observations	9,974	9,974	9,974	9,974	9,974	9,850	9,850	9,850	9,850	9,850

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. Standard errors clustered at the city level reported in parentheses. Columns (1)-(4) and (6)-(9) only include year fixed effects as additional controls. Columns (5) and (10) also include indicator variables for *Presence of Very Large River*, *Presence of Large River*, *Presence of Small River*, *Located on the Coast,* and *Fraction of City Area in Water (1900)* together with an indicator for missing values. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A10: Government Form and Police Employment: Robustness

Dependent Variable=	Log (Officers per Capita)		Log (Civilians per Capita)	
Model:	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Baseline	-0.16***	-0.28**	0.09	0.69**
	(0.03)	(0.14)	(0.07)	(0.35)
	Panel A: Alternative Samples			
(1) Drop Very Small and Very	-0.13***	-0.28**	0.14*	0.68*
Large Cities	(0.04)	(0.14)	(0.07)	(0.35)
(2) Drop Dependent Variable	-0.08***	-0.21*	0.18**	0.79**
Outliers	(0.03)	(0.11)	(0.07)	(0.34)
(3) Drop Low Government	-0.15***	-0.19*	0.12	0.80**
Concentration Census Divisions	(0.04)	(0.11)	(0.08)	(0.31)
(4) Drop post-COPS Program	-0.17***	-0.35**	0.10	0.70**
Years	(0.04)	(0.15)	(0.07)	(0.35)
(5) Drop High Missing	-0.18***	-0.26*	0.05	0.95**
Precipitation Cities	(0.04)	(0.15)	(0.07)	(0.43)
(6) Drop Far from Weather	-0.17***	-0.28*	0.13*	0.79**
Station Cities	(0.04)	(0.16)	(0.08)	(0.40)
	Panel B: Alternative Standard Error Construction and Inferences			
(7) Cluster on both City and	-0.16***	-0.28**	0.09	0.69**
Year	(0.03)	(0.14)	(0.07)	(0.34)
(8) Cluster on Weather Station	-0.16***	-0.28**	0.09	0.69
	(0.04)	(0.13)	(0.09)	(0.55)
(9) Conley Standard Errors	-0.16***	-0.29**	0.09	0.66
(No Year Fixed Effects)	(0.04)	(0.14)	(0.07)	(0.41)
	Panel C: Alternative Police Employment Scaling			
(10) Per Property Crime	-0.35***	-0.64**	-0.07	0.22
	(0.05)	(0.25)	(0.06)	(0.23)
(11) Per Violent Crime	-0.16	-1.02*	0.11	-0.16
	(0.10)	(0.53)	(0.10)	(0.40)
(12) Per Total Crime	-0.34***	-0.66**	-0.06	0.20
	(0.05)	(0.26)	(0.06)	(0.22)
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Division & Geography	No	No	No	No

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States as of 1900. The main entries report estimates of β_1 in equation (1) in the text. Standard errors clustered at the city level reported in parentheses (except in Panel B). See text for detailed descriptions of the specifications in panels A, B, and C. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A11: Government Form and Police Employment: Alternative Proxies for Mechanisms

Dependent Variable =	Log(Officers per Capita)			Log(Civilians per Capita)		
	(1)	(2)	(3)	(4)	(5)	(6)
Manager	0.04	-0.07**	0.02	0.01	-0.03	-0.01
	(0.03)	(0.03)	(0.03)	(0.09)	(0.07)	(0.09)
Manager × No Police in Union (1968)	-0.14***	–	-0.14***	-0.01	–	-0.01
	(0.05)		(0.05)	(0.13)		(0.13)
Manager × High Δ Relative Crime Rate	–	0.03*	0.03**	–	0.05	0.05
		(0.02)	(0.02)		(0.04)	(0.04)
No Police in Union (1968)	0.17*	–	0.17*	0.24	–	0.24
	(0.09)		(0.09)	(0.20)		(0.20)
High Δ Relative Crime Rate	–	0.00	0.00	–	0.00	0.00
		(0.01)	(0.01)		(0.02)	(0.02)
<u>Additional Controls:</u>						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Division & Geography	Yes	Yes	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	Yes	Yes	Yes	Yes
Institutional & Political	Yes	Yes	Yes	Yes	Yes	Yes
Sample Years	1977-2000	1977-2000	1977-2000	1977-2000	1977-2000	1977-2000
Number of Clusters	248	248	248	248	248	248
Number of Observations	5,576	5,576	5,576	5,553	5,553	5,553

Notes: Authors' calculations with 1977-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States as of 1900. Columns (1)-(6) report OLS estimates of equation (1) in the text augmented with the indicated variables and interaction terms. Standard errors clustered at the city level reported in parentheses. All models include year and Census division fixed effects, geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude, a Latitude-Longitude interaction, Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, and *Percentage of Clay in Soil*, demographic controls: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and an indicator variable for *Population Density Missing*, institutional and political controls: *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Low Police Unionization 1968*, and *High Local Newspaper Sales 1990-2000,* as well as indicator variables for missing values for each variable. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A12: Government Form and Government Spending: Electoral Effects

Dependent Variable =	Log(City Government Spending per Capita)		
	(1)	(2)	(3)
Election	0.039	0.058*	0.057*
	(0.024)	(0.033)	(0.033)
Manager	0.093	0.025	0.061
	(0.067)	(0.084)	(0.077)
Manager × Election	-0.008	0.007	0.004
	(0.042)	(0.055)	(0.054)
<u>Additional Controls:</u>			
Year Fixed Effects	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes
Demographic	No	No	Yes
Sample Years	1960-2000	1960-2000	1960-2000
Sample Cities	All	Large	Large
Number of Clusters	248	174	174
Number of Observations	10,168	7,134	7,134

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year. In column (1) the sample is the largest 248 self-governing cities in the United States as of 1900. In the remaining columns the sample is the 174 cities that also have at least 50,000 residents in 1960. Columns (1)-(3) report OLS estimates of equation (1) in the text augmented with *Election* and *Manager×Election* controls. Standard errors clustered at the city level reported in parentheses. Column (3) also includes: *Population Density, Fraction Non-White, Fraction College Graduate*, *Median Household Income,* and an indicator variable for *Population Density Missing***.** * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A13: Government Form and Crime Rates

Dependent Variable =	Log(Violent Crime Rate)		Log(Property Crime Rate)	
Model:	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Manager	0.06	0.34	0.14***	0.01
	(0.07)	(0.40)	(0.03)	(0.24)
<u>Additional Controls:</u>				
Year Fixed Effects	Yes	Yes	Yes	Yes
Division & Geography	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	Yes	Yes
Sample Years	1975-2000	1975-2000	1975-2000	1975-2000
Number of Clusters	248	248	248	248
Number of Observations	6,131	6,131	6,148	6,148

Notes: Authors' calculations with 1975-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. Each column reports the results from one regression. Standard errors clustered at the city level reported in parentheses. OLS models include Census division fixed effects and the following geographic controls: *Elevation Minimum*, *Elevation Maximum, Latitude, Longitude, a Latitude-Longitude interaction, Distance to Nearest River*, *Presence of Large River*, *Presence of Swamp*, *Located on Coast*, *Percentage of Clay in Soil*, as well as demographic controls: *Population Density, Fraction of Population Non-White, Fraction of Population College Graduate*, *Median Household Income,* and *Population Density Missing*. IV models also include *Century Precipitation Shocks* and *Median Precipitation*. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A14: Crime Rates and Policymaker Salary: OLS Estimates

Dependent Variable =	Log(Policymaker Salary)			
	(1)	(2)	(3)	(4)
Log(Violent Crime Rate)	0.01	0.01	–	–
	(0.04)	(0.03)		
Log(Property Crime Rate)	–	–	0.01	0.01
			(0.03)	(0.02)
Log(Violent Crime Rate) × Manager	0.02	0.01	–	–
	(0.03)	(0.03)		
Log(Property Crime Rate) × Manager	–	–	0.00	0.00
			(0.03)	(0.03)
Manager	0.53*	0.56*	0.53	0.57*
	(0.31)	(0.30)	(0.34)	(0.32)
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes
Demographic	No	Yes	No	Yes
Sample Years	1992+	1992+	1992+	1992+
Number of Clusters	233	233	233	233
Number of Observations	1,052	1,052	1,055	1,055

Notes: Authors' calculations with 1992-1993, 1995-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. Standard errors clustered at the city level reported in parentheses. The models include year and city fixed effects. The models in columns (2) and (4) also include the following demographic controls: *Population Density, Fraction of Population Non-White, Fraction of Population College Graduate*, *Median Household Income,* and *Population Density Missing*. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A15: Detailed Reporting of Table 2 in the Main Text

Dependent Variable =	Log(Officers per Capita)					Log(Civilians per Capita)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Manager	-0.16***	-0.08**	-0.07***	-0.05*	-0.07**	0.09	0.02	0.03	-0.02	-0.06
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.07)	(0.07)	(0.06)	(0.07)	(0.08)
Elevation Minimum	–	-0.01	-0.01	-0.02	-0.03	–	-0.05	-0.01	-0.01	0.00
		(0.04)	(0.02)	(0.02)	(0.02)		(0.06)	(0.05)	(0.05)	(0.07)
Elevation Maximum	–	-0.03***	-0.02**	-0.03**	-0.02*	–	-0.05**	-0.02	-0.01	-0.01
		(0.01)	(0.01)	(0.01)	(0.01)		(0.02)	(0.02)	(0.02)	(0.03)
Latitude	–	-0.05	0.36	0.03	-0.38	–	-1.27*	-0.64	-0.46	-1.32
		(0.43)	(0.31)	(0.35)	(0.36)		(0.96)	(0.83)	(0.85)	(1.23)
Longitude	–	0.23	0.09	0.20	0.31*	–	0.42	0.20	0.18	0.44
		(0.19)	(0.15)	(0.16)	(0.17)		(0.41)	(0.35)	(0.36)	(0.52)
Latitude × Longitude	–	0.00	0.00	0.00	-0.01	–	-0.01	-0.01	0.00	-0.01
		(0.00)	(0.00)	(0.00)	(0.00)		(0.01)	(0.01)	(0.01)	(0.01)
Distance to Nearest River	–	-0.32	-0.25	-0.16	-0.35	–	0.71	1.12	1.14	1.85
		(0.41)	(0.31)	(0.33)	(0.38)		(1.03)	(0.98)	(0.99)	(1.37)
Presence of Large River	–	0.11*	0.07	0.07	0.08*	–	-0.10	-0.10	-0.09	0.04
		(0.06)	(0.05)	(0.05)	(0.05)		(0.15)	(0.13)	(0.13)	(0.15)
Presence of Swamp	–	0.00	0.00	0.00	0.00	–	-0.01***	-0.01***	-0.01***	-0.01***
		(0.00)	(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)	(0.00)
Located on Coast	–	0.17***	0.12***	0.11***	0.09***	–	0.20***	0.13*	0.12*	0.00
		(0.04)	(0.03)	(0.03)	(0.03)		(0.08)	(0.07)	(0.07)	(0.09)
Percentage of Clay in Soil	–	-0.06***	-0.03	-0.03	-0.01	–	-0.08	-0.03	-0.03	0.05
		(0.03)	(0.02)	(0.02)	(0.02)		(0.05)	(0.05)	(0.05)	(0.06)
Division 2: Mid Atlantic	–	0.18**	0.07	0.07	0.05	–	0.03	-0.20	-0.28**	-0.53**
		(0.06)	(0.05)	(0.06)	(0.07)		(0.16)	(0.13)	(0.14)	(0.18)
Division 3: East North Central	–	0.20	0.07	0.02	0.00	–	0.16	-0.09	-0.04	-0.44
		(0.13)	(0.10)	(0.10)	(0.11)		(0.28)	(0.25)	(0.24)	(0.31)
Division 4: West North Central	–	0.17	0.07	0.02	-0.05	–	0.41	0.19	0.21	-0.31
		(0.19)	(0.15)	(0.15)	(0.14)		(0.39)	(0.35)	(0.34)	(0.42)
Division 5: South Atlantic	–	0.25**	0.03	-0.02	-0.01	–	0.08	-0.31	-0.30	-0.67*
		(0.12)	(0.10)	(0.10)	(0.11)		(0.29)	(0.25)	(0.24)	(0.36)
Division 6: East South Central	–	0.30**	0.13	0.08	0.04	–	0.45	0.14	0.19	-0.23
		(0.15)	(0.14)	(0.14)	(0.14)		(0.32)	(0.30)	(0.30)	(0.35)
						–				

Division 7: West South Central	–	0.58**	0.29	0.20	0.11	–	0.43	-0.03	0.10	-0.52
		(0.22)	(0.18)	(0.18)	(0.17)		(0.44)	(0.38)	(0.39)	(0.48)
Division 8: Mountain	–	1.05***	0.76***	0.74***	0.56**	–	1.37**	0.63	0.70	0.08
		(0.35)	(0.26)	(0.26)	(0.23)		(0.68)	(0.60)	(0.60)	(0.72)
Division 9: Pacific	–	0.95**	0.57**	0.50*	0.17**	–	0.81	0.15	0.26	-0.35
		(0.37)	(0.30)	(0.28)	(0.25)		(0.73)	(0.65)	(0.63)	(0.76)
Population Density	–	–	0.02***	0.02***	0.01	–	–	0.00	0.00	0.00
			(0.01)	(0.01)	(0.00)			(0.02)	(0.02)	(0.01)
Fraction Non-White	–	–	0.83***	0.81***	0.73***	–	–	1.45***	1.46***	1.40***
			(0.13)	(0.13)	(0.09)			(0.25)	(0.26)	(0.29)
Fraction College Graduate	–	–	1.04***	0.94***	0.96***	–	–	2.04***	1.81***	0.60
			(0.21)	(0.22)	(0.21)			(0.54)	(0.53)	(0.58)
Median Household Income	–	–	-0.12***	-0.11***	-0.13***	–	–	-0.11	-0.09	-0.03
			(0.03)	(0.03)	(0.02)			(0.07)	(0.07)	(0.06)
Incumbent Mayor Republican	–	–	–	0.02	0.03	–	–	–	-0.02	-0.14
				(0.02)	(0.03)				(0.05)	(0.09)
Non-Partisan Elections 1960	–	–	–	0.03	0.03	–	–	–	-0.08	-0.08
				(0.03)	(0.03)				(0.08)	(0.09)
Fraction At-Large Seats 1960	–	–	–	-0.07**	-0.06*	–	–	–	0.16**	0.04
				(0.03)	(0.03)				(0.08)	(0.10)
Early Civil Service	–	–	–	0.02	0.01	–	–	–	-0.03	-0.11
				(0.03)	(0.03)				(0.07)	(0.09)
Police Unionization 1968	–	–	–	0.00	0.00	–	–	–	0.00	0.00
				(0.00)	(0.00)				(0.00)	(0.00)
Local Newspaper Sales 1990-2000	–	–	–	0.01	0.00	–	–	–	-0.02	-0.04***
				(0.01)	(0.01)				(0.01)	(0.01)
Log(Policymaker Salary)	–	–	–	–	0.09***	–	–	–	–	0.16*
					(0.02)					(0.08)
Policymaker Male	–	–	–	–	0.08**	–	–	–	–	-0.02
					(0.04)					(0.09)
Policymaker White	–	–	–	–	-0.03	–	–	–	–	0.00
					(0.03)					(0.07)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Years	Full	Full	Full	Full	1992+	Full	Full	Full	Full	1992+
Number of Clusters	248	248	248	248	236	248	248	248	248	235
Number of Observations	9,974	9,974	9,974	9,974	1,164	9,850	9,850	9,850	9,850	1,156

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6, except for columns (5) and (10) where the sample is 1992-1993, 1995-2000. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States as of 1900. Columns (1)-(10) report an OLS estimate of β_1 from equation (1) in the text. Standard errors clustered at the city level reported in parentheses. Columns (1) and (6) do not include any additional controls. Columns (3) and (8) also include the *Population Density Missing*. Columns (4) and (9) also include controls for indicator variables for missing values of *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Police Unionization 1968*, and *Local Newspaper Sales 1990-2000.* * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A16: Detailed Reporting of Table 3 in the Main Text

Dependent Variable =	Manager		Log(Officers per Capita)		Manager		Log(Civilians per Capita)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager	–	–	-0.28**	-0.50**	–	–	0.69**	0.19
			(0.14)	(0.21)			(0.35)	(0.43)
LFC Precipitation Shocks	11.67***	9.79***	–	–	11.69***	9.85***	–	–
	(2.85)	(2.54)			(2.85)	(2.54)		
Century Precipitation Shocks	-6.90***	-9.86***	-0.27	-1.78	-6.95***	-9.87***	7.11***	2.23
	(2.07)	(2.29)	(0.87)	(1.54)	(2.07)	(2.30)	(1.82)	(2.36)
Median Precipitation	-0.54	1.41*	0.55**	0.38	-0.53	1.42*	-1.41**	-1.41
	(0.45)	(0.81)	(0.22)	(0.61)	(0.45)	(0.81)	(0.60)	(1.12)
Elevation Minimum	–	0.11**	–	0.03	–	0.11**	–	-0.07
		(0.05)		(0.04)		(0.05)		(0.08)
Elevation Maximum	–	0.02	–	-0.03	–	0.02	–	-0.06
		(0.03)		(0.02)		(0.03)		(0.03)
Latitude	–	2.28**	–	0.79	–	2.27**	–	-2.48
		(1.06)		(0.79)		(1.06)		(1.60)
Longitude	–	-1.06**	–	-0.14	–	-1.06**	–	0.97
		(0.49)		(0.35)		(0.49)		(0.72)
Latitude × Longitude	–	0.02**	–	0.01	–	0.02**	–	-0.02
		(0.01)		(0.01)		(0.01)		(0.02)
Distance to Nearest River	–	2.51***	–	0.81	–	2.51***	–	0.20
		(0.79)		(0.75)		(0.79)		(1.57)
Presence of Large River	–	0.00	–	0.12*	–	0.00	–	-0.10
		(0.10)		(0.07)		(0.10)		(0.16)
Presence of Swamp	–	0.00	–	0.00	–	0.00	–	0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Located on Coast	–	-0.12*	–	0.11**	–	-0.12*	–	0.21**
		(0.07)		(0.05)		(0.07)		(0.09)
Percentage of Clay in Soil	–	-0.02	–	-0.05	–	-0.02	–	-0.07
		(0.05)		(0.03)		(0.05)		(0.05)
Division 2: Mid Atlantic	–	-0.02	–	0.19**	–	-0.03	–	-0.11
		(0.14)		(0.09)		(0.14)		(0.16)
Division 3: East North Central	–	-0.03	–	0.21	–	-0.04	–	0.08
		(0.24)		(0.16)		(0.24)		(0.28)

Division 4: West North Central	–	0.05	–	0.23	–	0.04	–	0.30
		(0.31)		(0.22)		(0.31)		(0.39)
Division 5: South Atlantic	–	0.59	–	0.54**	–	0.58	–	-0.16
		(0.24)		(0.21)		(0.24)		(0.47)
Division 6: East South Central	–	0.16	–	0.38**	–	0.14	–	0.36
		(0.28)		(0.19)		(0.28)		(0.32)
Division 7: West South Central	–	0.42	–	0.81***	–	0.40	–	0.22
		(0.39)		(0.29)		(0.39)		(0.49)
Division 8: Mountain	–	-0.41	–	0.93**	–	-0.43	–	1.29*
		(0.52)		(0.37)		(0.53)		(0.71)
Division 9: Pacific	–	0.12	–	1.07***	–	0.10	–	0.69
		(0.62)		(0.40)		(0.62)		(0.75)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excluded Instrument *F*-stat	16.76	14.88	–	–	16.83	14.99	–	–
[p-value]	[0.0001]	[0.0001]			[0.0001]	[0.0001]		
Sample Years	Full	Full	Full	Full	Full	Full	Full	Full
Number of Clusters	248	248	248	248	248	248	248	248
Number of Observations	9,974	9,974	9,974	9,974	9,850	9,850	9,850	9,850

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States as of 1900. Each column reports the results from one regression. The first row of columns (3)-(4),(7)-(8) report IV estimates of β_1 from equation (1) in the text with the first stage given by columns (1)-(2),(5)-(6). Standard errors clustered at the city level in parentheses. The excluded instrument is *LFC Precipitation Shocks*. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A17: Detailed Reporting of Table 4 in the Main Text

Dependent Variable=	Log(Officers per Capita)			Log(Civilians per Capita)		
	(1)	(2)	(3)	(4)	(5)	(6)
Manager	-0.03	-0.07***	-0.06**	0.00	-0.02	-0.01
	(0.03)	(0.03)	(0.03)	(0.07)	(0.08)	(0.08)
Manager × Voting Rights Act Preclearance	-0.22*	–	-0.23**	-0.31	–	-0.30
	(0.12)		(0.12)	(0.20)		(0.20)
Manager × Crack Epidemic	–	0.05**	0.06***	–	0.00	0.02
		(0.02)	(0.02)		(0.06)	(0.06)
Voting Rights Act Preclearance	0.11	–	0.11	0.04	–	0.04
	(0.08)		(0.08)	(0.17)		(0.18)
Crack Epidemic	–	-0.02	-0.02	–	-0.21***	-0.21***
		(0.03)	(0.03)		(0.07)	(0.07)
Elevation Minimum	-0.02	-0.02	-0.02	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)
Elevation Maximum	-0.03**	-0.03**	-0.03**	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Latitude	0.15	0.04	0.16	-0.37	-0.43	-0.34
	(0.35)	(0.35)	(0.35)	(0.88)	(0.85)	(0.88)
Longitude	0.17	0.20	0.17	0.16	0.15	0.14
	(0.16)	(0.16)	(0.16)	(0.36)	(0.35)	(0.36)
Latitude × Longitude	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Distance to Nearest River	-0.30	-0.17	-0.31	0.99	1.11	0.97
	(0.33)	(0.33)	(0.33)	(1.00)	(0.98)	(1.00)
Presence of Large River	0.07	0.07	0.07	-0.10	-0.09	-0.11
	(0.05)	(0.05)	(0.05)	(0.13)	(0.13)	(0.13)
Presence of Swamp	0.00	0.00	0.00	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Located on Coast	0.11***	0.11***	0.11***	0.11*	0.12*	0.12*
	(0.03)	(0.03)	(0.03)	(0.06)	(0.07)	(0.07)
Percentage of Clay in Soil	-0.02	-0.03	-0.02	-0.02	-0.03	-0.02
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)
Division 2: Mid Atlantic	0.07	0.07	0.07	-0.28**	-0.25*	-0.26*
	(0.06)	(0.06)	(0.06)	(0.14)	(0.14)	(0.14)

Division 3: East North Central	0.03	0.02	0.03	-0.03	-0.02	-0.02
	(0.10)	(0.10)	(0.10)	(0.24)	(0.24)	(0.24)
Division 4: West North Central	0.03	0.02	0.03	0.22	0.21	0.22
	(0.14)	(0.15)	(0.14)	(0.34)	(0.34)	(0.34)
Division 5: South Atlantic	0.03	-0.01	0.04	-0.19	-0.29	-0.18
	(0.10)	(0.10)	(0.10)	(0.25)	(0.24)	(0.25)
Division 6: East South Central	0.09	0.09	0.10	0.21	0.19	0.21
	(0.14)	(0.14)	(0.14)	(0.29)	(0.30)	(0.29)
Division 7: West South Central	0.26	0.21	0.27	0.22	0.08	0.21
	(0.19)	(0.18)	(0.19)	(0.40)	(0.39)	(0.40)
Division 8: Mountain	0.78***	0.74***	0.78***	0.74	0.68	0.72
	(0.26)	(0.26)	(0.26)	(0.59)	(0.60)	(0.59)
Division 9: Pacific	0.54*	0.50*	0.54*	0.31	0.26	0.31
	(0.28)	(0.28)	(0.28)	(0.63)	(0.63)	(0.63)
Population Density	0.02***	0.02***	0.02***	0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Fraction Non-White	0.80***	0.81***	0.80***	1.47***	1.47***	1.47***
	(0.13)	(0.13)	(0.13)	(0.26)	(0.27)	(0.26)
Fraction College Graduate	0.88***	0.95***	0.88***	1.73***	1.82***	1.75***
	(0.22)	(0.22)	(0.22)	(0.51)	(0.52)	(0.51)
Median Household Income	-0.11***	-0.11***	-0.11***	-0.08	-0.09	-0.09
	(0.03)	(0.03)	(0.03)	(0.07)	(0.06)	(0.06)
Incumbent Mayor Republican	0.02	0.02	0.02	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)
Non-Partisan Elections 1960	0.03	0.03	0.03	-0.08	-0.08	-0.08
	(0.03)	(0.03)	(0.03)	(0.08)	(0.08)	(0.08)
Fraction At-Large Seats 1960	-0.07**	-0.07**	-0.07**	0.16**	0.15*	0.16**
	(0.03)	(0.03)	(0.03)	(0.08)	(0.08)	(0.08)
Early Civil Service	0.02	0.02	0.02	-0.05	-0.03	-0.04
	(0.03)	(0.03)	(0.03)	(0.07)	(0.07)	(0.07)
Police Unionization 1968	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Local Newspaper Sales 1990-2000	0.01	0.01	0.01	-0.02	-0.02	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Sample Years	Full	Full	Full	Full	Full	Full
Number of Clusters	248	248	248	248	248	248
Number of Observations	9,974	9,974	9,974	9,850	9,850	9,850

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States as of 1900. Each column reports the results from one regression. Columns (1)-(6) report OLS estimates of equation (1) in the text augmented with the additional controls indicated. Standard errors clustered at the city level in parentheses. Columns (2) and (5) also include *Population Density Missing*. Columns (3) and (6) also include controls for indicator variables for missing values of *Incumbent Mayor Republican*, *Non-Partisan Elections 1960*, *Fraction At-Large Seats 1960*, *Early Civil Service*, *Police Unionization 1968*, and *Local Newspaper Sales 1990-2000.* * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.

TABLE A18: Detailed Reporting of Table 5 in the Main Text

Dependent Variable =	Log(Officers per Capita)				Log(Civilians per Capita)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager	-0.078	-0.064	-0.044	-0.008	0.009	-0.040	-0.034	0.068
	(0.056)	(0.084)	(0.048)	(0.057)	(0.095)	(0.110)	(0.086)	(0.136)
Manager × Election	-0.009*	-0.015**	-0.013***	-0.013***	-0.001	-0.009	-0.007	-0.006
	(0.005)	(0.007)	(0.005)	(0.005)	(0.013)	(0.015)	(0.014)	(0.014)
Manager × Log(City Spending per Capita)	–	–	–	-0.005	–	–	–	-0.015
				(0.007)				(0.017)
Election	0.006	0.010**	0.009**	0.009**	0.008	0.006	0.006	0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.009)	(0.008)	(0.008)
Log(City Spending per Capita)	–	–	–	0.005	–	–	–	-0.003
				(0.004)				(0.008)
Population Density	–	–	0.001	0.001	–	–	-0.017**	-0.019**
			(0.004)	(0.004)			(0.009)	(0.008)
Fraction Non-White	–	–	0.027	0.027	–	–	-0.153	-0.145
			(0.139)	(0.139)			(0.350)	(0.347)
Fraction College Graduate	–	–	-0.569	-0.558	–	–	-1.425	-1.402
			(0.523)	(0.518)			(0.998)	(0.985)
Median Household Income	–	–	-0.075**	-0.075**	–	–	0.066	0.062
			(0.031)	(0.031)			(0.067)	(0.067)
<u>Additional Controls:</u>								
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Years	Full	Full	Full	Full	Full	Full	Full	Full
Sample Cities	All	Large	Large	Large	All	Large	Large	Large
Number of Clusters	248	174	174	174	248	174	174	174
Number of Observations	9,974	7,033	7,033	7,033	9,850	7,017	7,017	7,017

Notes: Authors' calculations with 1960-2000 city data as described in Appendix A.6. The unit of observation is a city-year. In columns (1) and (5) the sample is the largest 248 self-governing cities in the United States as of 1900. In the remaining columns the sample is the 174 cities that also have at least 50,000 residents in 1960. Each column reports the results from one regression. The main entries in columns (1)-(8) report OLS estimates of equation (1) in the text augmented with the indicated controls. Standard errors clustered at the city level reported in parentheses. Columns (1) and (5) do not include any additional controls. Columns (3) and (7) also include *Population Density Missing*. Columns (4) and (8) include a *Missing City Spending* indicator. * indicates significance at the 10 percent level, ** significance at the 5 percent level, and *** significance at the 1 percent level.